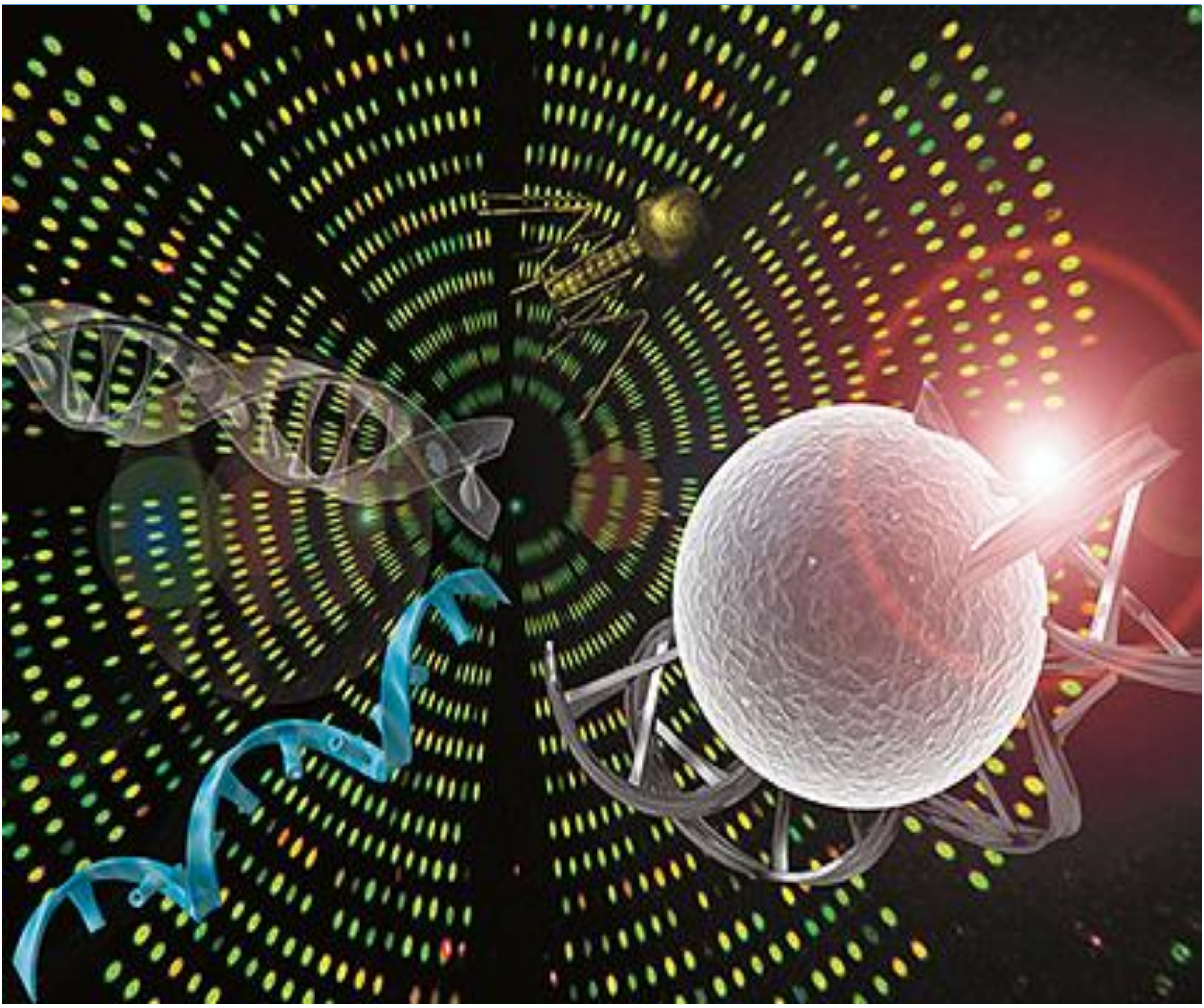
BHARATI VIDYAPEETH
(DEEMED TO BE UNIVERSITY), PUNE

FACULTY OF INTERDISCIPLINARY STUDIES
M.Sc. - BIOINFORMATICS
New Syllabus



BVDU-RGITBT-M.Sc. Bioinformatics Syllabus

2019

BHARATI VIDYAPEETH DEEMED TO BE UNIVERSITY
PUNE

REVISED SYLLABUS FOR
MASTER OF SCIENCE
M.Sc. IN BIOINFORMATICS

UNDER
FACULTY OF INTERDISCIPLINARY STUDIES

SYLLABUS OF SEM I – SEM IV UNDER
CHOICE BASED CREDIT SYSTEM

To be effective from Academic Year

2019-20

Bharati Vidyapeeth Deemed To Be University is a multidisciplinary, multicampus university having 32 Institutions imparting quality education in various disciplines. All programmes of the University are approved by UGC and respective statutory councils. BVDU has been re accredited for the third time with 'A+' grade by NAAC in 2017. UGC has accorded 12B Status [UGC ACT1956] to the university. Ministry of Human Resource and Development, Government of India has awarded "A" category to the University in 2012 based on parameters including innovative programs, research and infrastructure facilities. The University is a member of Association of Indian Universities [AIU] which has ranked BVDU among top 10 universities of India for International students' enrollment. BVDU is also a member of International Association of Universities.

Rajiv Gandhi Institute of IT and Biotechnology is a constituent unit of BVDU established in 2003. The Institute is approved by UGC to conduct graduate and post graduate courses in Biotechnology. The Institute has excellent infrastructure, state-of-the-art laboratories and competent faculty facilitating appropriate learning environment. The Institute offers one undergraduate and four postgraduate programmes in Biotechnology.

INTRODUCTION

The Master of Science (M.Sc.) in Bioinformatics is a full time post graduate program offered by **Bharati Vidyapeeth (Deemed to be University (BVDU))** in its constituent unit, Rajiv Gandhi Institute of IT and Biotechnology (RGITBT). The course was initiated in the year 2010 and was designed to facilitate empowerment of students to face cutting edge technological applications in bioinformatics sector. 21st century biology has been transformed into integrative biology due to paradigm shift from reductionist to holistic approach. The data explosion is one of the obvious causes of the transformation to holistic or systems biology mode. In order to have deeper insights in to biological systems, integration of the data from genome to phenome levels to generate dynamic models has surfaced as key area of bioinformatics. On its implementation for 8 years, the curriculum is being revised two times to embrace newer emerging disciplines and value added courses. The revised **M.Sc. Bioinformatics is a full time 108 credits** Program to be implemented in Rajiv Gandhi Institute of IT and Biotechnology from the academic year 2019-20. The feedback of students, alumni, faculty, employers and parents has a substantial contribution in designing of this curriculum.

OBJECTIVES

1. To impart deep knowledge of the discipline
2. Develop skills in relevant areas to enhance employment opportunities
3. Introduce emerging areas of NGS, Drug designing and data mining
4. Build interdisciplinary approach
5. Foster global competence among students
6. Inculcate social and moral values and sense of scientific responsibilities in students

ELIGIBILITY FOR ADMISSION TO THE COURSE

Candidates satisfying following criteria are eligible to apply for M.Sc. Bioinformatics Course

1. A student who has obtained Bachelor's degree with minimum **50%** in (i) B.Sc. (Biotechnology / Microbiology / Biochemistry / Zoology / Botany / Chemistry /Agriculture/ Mathematics /Statistics / Computer Science). (ii)B. Pharm (iii). MBBS (iv) BDS (v) BAMS (vi) BHMS (vii) B.E. in Computer Science or Information technology or Biomedical Engineering will be considered eligible for admission to M.Sc. Bioinformatics course or **45%** aggregate marks for **SC / ST** category respectively at graduate level university examination.
2. Subject to above conditions, the admission will be based on the merit at Entrance Examination conducted by Bharati Vidyapeeth (Deemed to be University).

DURATION OF THE COURSE

The course will be executed in four semesters. The medium of instruction and examination will be only English.

RULES FOR THE COURSE

1. The entire course is of 108 credits.
2. One credit for theory course is equivalent to 15 lectures/tutorials; while one credit for practical course is equivalent to 25 – 30 hrs. of lab /field work or demonstration.
3. The curriculum comprises of core, advanced and value added courses. The Core and Advanced Courses are compulsory where as Value Added are elective.
4. The Core Courses are aimed at providing fundamental knowledge of the discipline. The advanced courses are designed to develop manpower for bioinformatics industries as per current need. The Value Added Courses are intended to inculcate the thrust of new domains of bioinformatics within students.
5. The teaching schedule for the 3 credits and 2 credits theory courses will be 3 and 2 lectures per week respectively. All courses will have one tutorial fortnightly.
6. The respective elective course will be implemented only if more than 10 students enroll for that course.
7. All core courses will be evaluated by University Examination. The elective courses will be evaluated by University Examination and Continuous Assessment.

8. Two extra credits will be awarded to students if there is any significant outcome of their dissertation study. The research outcome in terms of publication in indexed national/International journal; filing of patent; or commercialization of technology will be considered for the award of credits.

RULES FOR EXAMINATION

A: Nature of Examination:

1. Each course will have 40% marks for internal assessment and 60% marks for semester-end examination.
2. The assessment for 1, 2 and 3 Credits courses will be as given in following table:

Table 1: Evaluation pattern for one to three Credit Courses

Course Credits	Marks for UE (60% Weightage)	Marks for IE (40% weightage)	Total Marks for evaluation
1	**15**	**10**	**25**
2	**30**	**20**	**50**
3	**45**	**30**	**75**
4	**60**	**40**	**100**

3. The duration of **60 Marks UE theory paper will be 2.00 Hrs; for 30 Marks 1.30 Hrs.** respectively.
4. The Internal Assessments (IA) will be conducted by the Institute and an end-of-the term University Examination (UE) conducted by the university. The UE will be based on the entire syllabus.
5. The performances at UE and IA will be combined to obtain the Grade Point Average (GPA) for the course.

STANDARD OF PASSING

A: Grading System: A **10-point absolute grading system** will be adapted for grading in each head of passing. The system will have seven grade points, the highest being 10. The grading system shall be as shown in Table-2 below. The performance indicators O, A+, A, B+, B, C, and D shall respectively mean Outstanding, Excellent, Very Good, Good, Average, Satisfactory, and Poor.

Table-2: The grading system under CBCS

Range of Marks (out of 100)	Grade Point	Grade
80 ≥*Marks* ≤ 100	10	O
70 ≥ *Marks* ≤ 80	9	A+
60 ≥*Marks* ≤ 70	8	A
55 ≥*Marks* ≤ 60	7	B+
50 ≥*Marks* ≤ 55	6	B
40 ≥*Marks* ≤ 50	5	C
***Marks* <40**	0	D

1. The grade point average (GPA) for a course shall be calculated by first finding the total marks **out of 100 for the course**. **The corresponding GP (as per the table-2) shall be the GPA for the course**.
2. Two kinds of performance indicators, namely, the Semester Grade Point Average (SGPA) and the Cumulative Grade Point Average (CGPA) shall be computed at the end of each term. The SGPA measures the cumulative performance of a learner in all the courses in a particular semester, while the CGPA measures the cumulative performance in all courses up to and including the current semester. The CGPA of a student when he/she completes the program is his/her final result.
3. The SGPA is calculated by the formula , $SGPA = \frac{\sum Ck \times GPAk}{\sum Ck}$ where *Ck* is the Credit-value assigned to a course and *GPAk* is the GPA obtained by the student in the course. In the above, the sum is taken over all the courses that the student has undertaken for the study during the semester, including those in which

he/she might have failed or those for which he/she remained absent. **The SGPA shall be calculated up to two decimal place accuracy.**

4. The CGPA is calculated by the formula , $\text{CGPA} = \frac{\sum C_k \times GPA_k}{\sum C_k}$ where C_k is the Credit-value assigned to a course and GPA_k is the GPA obtained by the student in the course. In the above, the sum is taken over all the courses that the student has undertaken for the study from the time of his/her enrolment and also the during the semester for which CGPA is calculated, including those in which he/she might have failed or those for which he/she remained absent. **The CGPA shall be calculated up to two decimal place accuracy.**

5. The CGPA, calculated after the minimum credits specified for the program are 'earned', will be the final result grace marks of 1, 2 or 3 may be awarded to a candidate at UE as per the university rules.

B: Standards of Passing and ATKT rules:

1. For all Core Courses, both UE and IA

Constitute separate heads-of-passing (HoP). In order to pass in such courses and to 'earn' the assigned credits

(a) the learner must obtain a minimum grade point of 5.0 (40% marks) at UE and also a minimum grade point of 5.0 (40% marks) at IA;

OR

(b) If he/she fails in IA, then also the learner passes in the course, **provided that a minimum of 25% is obtained in IA and GPA for the course is at least 6.0 (50%marks) in aggregate**. The GPA for a course will be calculated only if the learner passes in that course.

2. A student who fails at UE in a course has to reappear only at UE as a backlog candidate and clear the HoP. Similarly, a student who fails in a course at IA has to reappear only at IA as a backlog candidate and clear the HoP. **A student who passes in aggregate in a course need not reappear even if he failed at IA if he/she obtains 25% at IA.**

3. The students of Semester I and II & III will be admitted to next Semester even if he/she gets backlog in any of the course. They can reappear in the next semester examination as a backlog candidate.

C: AWARD OF HONOURS:

1. A student who has completed the minimum credits specified for the program shall be declared to have passed in the program. The final result will be in terms of letter grade and CGPA only and is based on the CGPA of all courses studied and passed. The criteria for the award of honors are given in Table 3.

Table 3: Criteria for the award of honors at the end of the program

Range of CGPA	Final Grade	Performance Descriptor	Equivalent Range of Marks (%)
$9.50 \geq CGPA \leq 10.00$	O	Outstanding	$80 \geq Marks \leq 100$
$9.00 \geq CGPA \leq 9.49$	A+	Excellent	$70 \geq Marks \leq 80$
$8.00 \geq CGPA \leq 8.99$	A	Very Good	$60 \geq Marks \leq 70$
$7.00 \geq CGPA \leq 7.99$	B+	Good	$55 \geq Marks \leq 60$
$6.00 \geq CGPA \leq 6.99$	B	Average	$50 \geq Marks \leq 55$
$5.00 \geq CGPA \leq 5.99$	C	Satisfactory	$40 \geq Marks \leq 50$
CGPA Below 5.00	F	Fail	Marks below 40

THE FORMAT OF THE TRANSCRIPTS

The transcripts may be acquired by the students indicating his/her performance in every semester examination. The transcript shall show the performance indicators given in the following table, in addition to any other information.

;

<table>
<tr><th rowspan="2">Course Number</th><th rowspan="2">Course Description</th><th rowspan="2">Number of Credits</th><th colspan="2">University Examination</th><th colspan="2">IA/CA</th><th rowspan="2">Grade Point Average
(GPA)</th><th rowspan="2">Result</th></tr>
<tr><th>Grade</th><th>Grade Point</th><th>Grade</th><th>Grade Point</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">Total Cumulative Credits Completed</td><td>SGPA</td><td>CGPA</td><td>Equivalent Marks (%)</td><td colspan="4">Note: GPA is calculated by adding the UE marks out of 60 and IA marks out of 40. The total marks out of 100 are converted to Grade Point, which will be GPA</td></tr>
</table>

PATTERN FOR ASSESSMENT

A: Pattern of Evaluation for Internal Assessment of Theory Courses:

The weightage for Internal Assessment is 40%. Students for IA of every theory course will be assessed for total of 40 marks for 3 credit course and for 20 marks for 2 credit course which will be cumulative marks obtained in two separate assessments specified below.

1. Two internal written examinations of 20 marks each for 3 credit course and 10 marks each for 2 credit course. A total of the two tests will be considered.
2. An optional assignment/ oral/ open book examination may be undertaken if desired.

B. Pattern of Evaluation for Internal Assessment of Practical Courses:

The Internal Assessment for every practical course will be of 20 Marks for 2 credits and 40 marks for 4 credits practical courses. The students for IA will be assessed on the basis of;

1. Performance for every practical: 10 Marks/20 Marks for 2/4 Credits practical courses respectively. (Marks to be distributed depending on total number of practical)
2. Assignment/ Oral examination/Tour Report: 10/20 Marks for 2/4 Credits courses

C: Pattern of question paper at University Examination

University Examination for 3 credit and 2 credit theory course will be of 60 marks and 30 marks respectively. **For 3 credit course**, the question paper will comprise of 6 questions, 3 questions each in section I and section II. Each question will be of 10 marks. All questions will be compulsory. The pattern of question paper will be as given on next page.

PATTERN OF QUESTION PAPER FOR 3 CREDIT COURSE OF UNIVERSITY THEORY EXAMINATION OF M.Sc. BIOINFORMATICS 2019 CBCS COURSE (TOTAL MARKS: 60, TME: 3.00 HRS.)

Instructions to Paper Setter:

1. Question paper of each course will comprise of total 6 questions,
2. Section I will have 3 questions and Section II 3 questions.
3. All questions will be compulsory. Each question will carry an internal option of one extra sub-question.
4. Q. no 1 will be objective, comprising of 7 questions of 2 mark each. They will be based on entire portion of Section I. Students will have to attempt any 5 out of these.
5. Q no 4 will be objective, comprising of 7 questions of 2 mark each. They will be based on entire portion of Section II. Students will have to attempt any 5 out of these.
6. Questions 2 & 3 of **Section I** and 5 & 6 of **Section II** will be descriptive and contain 3 sub-questions of 5 marks each out of which students will attempt any two.

7. Q 2 and 3 will be based solely on Unit I and II whereas Q 5 and 6 will be based on Unit III and IV of the syllabus respectively.
8. Students will attempt answers to Section I and Section II in separate answer books

SECTION I

Q.1 Attempt **Any FIVE** of the following:(Define, Explain why, Fill in the blanks, Give examples, Answer in brief) (10)

a
b
c
d
e
f
g

Q.2 Attempt **Any Two** of the following: (Answer the following, Differentiate Between, Give neat labeled diagram of) (10)

a
b
c

Q.3 Write short notes on **Any Two** of the following (10)

a
b
c

SECTION II

Q.4 Attempt **Any five** of the following: :(Define, explain why, Fill in the blanks, Give examples, Answer in brief) (10)

a
b
c
d
e
f
g

Q.5 Attempt **Any Two** of the following: (Answer the following, Differentiate Between, Give neat labeled diagram of) (10)

a
b
c

Q.6 Write short notes on **Any Two** of the following (10)

a
b
c

QUESTION PAPER PATTERN FOR 2 CREDITS THEORY COURSE AT UNIVERSITY EXAMINATION

For 2 credit course, the question paper will comprise of 4 questions, 2 questions each in section I and section II. Q1 of section I and Q3 of section II will be of 7 marks each while Q2 and Q4 will be of 8 marks each .All questions will be compulsory. The pattern of question paper will be as given on next page.

Pattern of question paper for 2 credit course of university theory examination of M.Sc. BIOINFORMATICS 2019 CBCS COURSE (Total Marks: 30, Tme: 1.50 Hrs.)

Instructions to Paper Setter:

1. Question paper of each course will comprise of total 4 questions,
2. Section I will have 2questions and Section II 2 questions.
3. All questions will be compulsory. Each question will carry an internal option of one extra sub-question.
4. Questions 1 of section I and 3 of section II will be objective, and contain 7 questions of 1 mark each out of which students will attempt any 5. They will be based on entire portion of Section I and section II respectively.
5. Questions 2 of **Section I** and 4 of **Section II** will be descriptive and contain 3 sub-questions of 5 marks each out of which students will attempt any two.
6. Q 2 and 4 will be based solely on Unit I and II of the syllabus respectively.
7. Students will attempt answers to Section I and Section II in separate answer books

SECTION I

Q.1 Attempt **Any Five** of the following: :(Define, Explain why, Fill in the blanks, Give examples, Answer in brief) (05)

a
b
c
d
e
f
g

Q.2 Attempt **Any Two** of the following: (Answer the following, Differentiate Between, Give neat labeled diagram of, Write short notes on) (10)

a
b
c

SECTION II

Q.3 Attempt **Any Five** of the following: :(Define, Explain why, Fill in the blanks, Give examples, Answer in brief) (05)

a
b
c
d
e
f
g
h

Q.4 Attempt **Any Two** of the following: (Answer the following, Differentiate Between, Give neat labeled diagram of, Write short notes on) (10)

a
b
c

D: PATTERN FOR QUESTION PAPER OF UNIVERSITY PRACTICAL EXAMINATION OF M.Sc. BIOINFORMATICS 2019 CBCS COURSE

(Total Marks: 30/60 for 2/4 credit courses, Time: 3 .00/6.00 Hrs.)

Q. 1	Major Practical	(10/20)
Q. 2	Spotting/Minor Experiment	(10/20)
Q. 3	Viva	(05/10)
Q. 4	Journal	(05/10)

Course structure of M.Sc. Degree Course in Bioinformatics Under Choice Based Credit System 2019-20

SEMESTER I

Course No. & Description	Title	Credits	IA	Univ. Exam	Total Credits
MBI 101 Basic Course-Theory	Cell Biology **(C)**	**2**	20	30	**28**
MBI 102 Basic Course –Theory	Biochemistry **(C)**	**2**	20	30	
MBI 103 Basic Course –Theory	Biomathematics **(C)**	**2**	20	30	
MBI 104 Basic Course –Theory	Biostatistics **(C)**	**2**	20	30	
MBI 105 Basic Course –Theory	C Programming and Data structure **(C)**	**3**	40	60	
MBI 106 Basic Course – Theory	Biological Informatics **(C)**	**2**	20	30	
MBI 107 Basic Course – Theory	DBMS & MongoDB **(C)**	**3**	40	60	
MBI 108 Elective Course - Theory	PERL Programming / HTML Programming	**2**	20	30	
MBI 109 Basic Course –Practical	Cell Biology and Biochemistry Lab **(C)**	**2**	20	30	
MBI 110 Basic Course –Practical	C Programming and Data Structure Lab **(C)**	**2**	20	30	
MBI 111 Basic Course –Practical	Biological Informatics Lab **(C)**	**2**	20	30	
MBI 112 Basic Course –Practical	DBMS & MongoDB lab **(C)**	**2**	20	30	
MBI 113 Elective Course - Practical	PERL Programming Lab /HTML Programming Lab	**2**	20	30	

SEMESTER II

Course No. & Description	Title	Credits	IA	Univ. Exam	Total Credits
MBI 201 Core Course –Theory	Statistical Analysis System (SAS) **(C)**	**2**	20	30	**30**
MBI 202 Core Course –Theory	R and Data Analytics **(C)**	**3**	40	60	
MBI 203 Core Course –Theory	JAVA and BioJAVA Programming **(C)**	**3**	40	60	
MBI 204 Core Course –Theory	Science of Omics **(C)**	**3**	40	60	
MBI 205 Core Course - Theory	Proteomics **(C)**	**2**	20	30	
MBI 206 Core Course -Theory	Molecular Biology	**2**	20	30	
MBI 207 Core Course -Theory	Recombinant DNA Technology	**2**	20	30	
MBI 208 Core Course-Theory	Structural Biology and Molecular Modeling	**3**	40	60	
MBI 209 Core Course - Practical	SAS and Data Analytics lab **(C)**	**2**	20	30	
MBI 210 Core Course - Practical	JAVA and BioJAVA Programming lab **(C)**	**2**	20	30	
MBI 211 Core Course - Practical	Omics Analysis Lab **(C)**	**2**	20	30	
MBI 212 Core Course - Practical	Molecular Biology and Recombinant DNA Technology lab	**2**	20	30	
MBI 213 Core Course- Practical	Structural Biology and Molecular Modeling Lab	**2**	20	30	

SEMESTER III

Course No. & Description	Title	Credits	IA	Univ. Exam	Total Credits
MBI 301 Core Course –Theory	Scientific Writing Skills	**2**	20	30	**30**
MBI 302 Advance Course –Theory	Chemoinformatics and Drug Designing	**3**	40	60	
MBI 303 Advance Course- Theory	Machine Learning Techniques	**3**	40	60	
MBI 304 Advance Course- Theory	Current Bioinformatics	**2**	20	30	
MBI 305 Advance Course –Theory	Python Programming	**2**	20	30	
MBI 306 Elective Course- Theory	Introduction to Clinical Trials and Pharmacovigilance / Selenium	**2**	20	30	
MBI 307 Elective Course- Theory	Cancer Genomics/ Biodiversity Informatics & Molecular Phylogenetics	**2**	20	30	
MBI 308 Elective Course- Theory	System Biology/Artificial Intelligence	**2**	Continuous Assessment		
MBI 309 Advance Course- Practical	Chemoinformatics and Drug Designing Lab	**2**	20	30	
MBI 310 Advance Course- Practical	Machine Learning Techniques Lab	**2**	20	30	
MBI 311 Advance Course- Practical	Current Bioinformatics Lab	**2**	20	30	
MBI 312 Advance Course -Practical	Python Programming Lab	**2**	20	30	
MBI 313 Elective Course- Practical	Introduction to Clinical Trials and Pharmacovigilance Lab / Selenium lab	**2**	20	30	
MBI 314 Elective Course -Practical	Cancer Genomics Lab/ Biodiversity Informatics & Molecular Phylogenetics Lab	**2**	20	30	

SEMESTER IV

Course No. & Description	Title	Credits	IA	Univ. Exam	Total Credits
MBI 401 Core Course	Research Project	20	40	60	20

Total Credits Offered: 28 C, Sem I+ 30 C, Sem II +30 C, Sem III+ 20C, Sem IV = 108

SEMESTER I

MBI 101: Cell Biology (C) **Total**

Basic Course – Theory; 2 Credits **30L**

UNIT I

1.	**Cell as a basic unit of life.** Basic structure, organization and composition of prokaryotic and eukaryotic cell. Plasma membrane structure and functions, Membrane models. Components of blood & their functions (Plasma, RBC, WBC, Platelets).	**4**
2.	**Cytoskeleton:** Organization and functions cytoskeleton, Actin filaments, actin binding proteins, Intermediate filaments, Microtubules	**3**
3.	**Cell Junctions:** Gap junctions, Tight junctions, Cell adhesion-integrins, selectins, cadherins. desmosomes and hemidesmosomes, plasmodesmata	**2**
4.	**Membrane Transport:** Transport across membrane- passive diffusion, osmosis, active transport, Ion Channels, Na+ and K+ pump, Ca2+ ATPase pump, co-transport, symport, antiport, endocytosis and exocytosis.	**6**

UNIT II

6.	**Cell Cycle:** Molecular events of cell division and cell cycle, regulation of cell cycle events- Cyclins, Cyclin dependent kinases, inhibitors.	**5**
7.	**Cell Signaling**: General principles of cell signaling, signaling via G-protein coupled receptors, kinase receptors, role of secondary messengers.	**6**
8.	Ageing and apoptosis, abnormal development and teratogenesis in animals	**4**

References:

1. Alberts, B., Bray, D., Lewis, J., Raf, M., Roberts, K., Watson, J.D. (1994). Molecular Biology of the Cell
2. Cooper, G.M. (1997).The Cell: A molecular approach, ASM Press, USA.
3. Darnell, J., Lodish, H., Baltimore, D. (1990). Molecular Cell Biology. Scientific American Books Inc. NY.
4. Hallwell, B., Gutteridge, J.M.C. (2002). Free Radicals Biology and Medicine. Oxford Press.UK.
5. Karp, G. (1996). Cell and Molecular Biology concepts and experiments, John Wiley and Sons Inc. NY.
6. Lodish, H., Baltimore, D., Berk, A., Zipursky, B.L., Mastsydaira, P., Darnell, J. (2004). Molecular Cell Biology, Scientific American Books Inc. NY.

MBI 102: Biochemistry (C) **Total**

Basic Course – Theory; 2 Credits **30L**

UNIT I

1.	**Introduction to Biomolecules: Carboydrates, Lipids and Proteins – their biological roles and functions**	1
2.	**Carbohydrates:** Classification-on basis of carbon no & functional grs.; Structure with molecular formulae; Stereochemistry (structural and conformational isomers); isomers with one and more chiral centres, properties of stereoisomers; cyclization of sugars; conformations of cyclic forms (chair & boat forms).	3
3.	**Important monosaccharides:** 6 C sugars- glucose, fructose, mannose, galactose, 5 C sugars- ribose, deoxyribose, important modified monosaccharides (2-keto-3-deoxy-D-*manno*-octulosonic acid (KDO), 2-keto-3-deoxy-D-*glycero*-D-*galacto*-nonulosonic acid (KDN), sialic acids, aminosugars, sugar acids (their biological role/ occurrence)	1
4.	**Linkages in Sugars:** Structure and functions of important di-saccharides (sucrose, lactose, maltose, cellobiose) and poly-saccharides (homo and heteropolymers, storage and structural polymers – starch, glycogen, cellulose, chitin, peptidoglycan)	1
5.	**Physico-chemical properties of carbohydrates:**Reducing-non-reducing properties; Exploitation for detection and separation –Benedict's test; Introduction to conjugate sugars (Glycoproteins and glycolipids and their importance)	2
6.	**Significance of carbohydrates in microbial, plant and animal system:** ABO blood groups, heparin, lectins, carbohydrate vaccines	2
7.	**Lipids:** Occurrence/sources: plant derived oils, ghee	1
8.	**Structure :**Structure of fatty acids, triglycerides; Classification viz., simple and compound, structural and storage with examples; unsaturated-saturated, Oils (refined/unrefined significance); fats, waxes, rancidity/spoilage of fats (microbial, oxidative)	2
9	**Functions:**Functions in a living system: significance of hydrophobicity and water immiscibility for compartmentalisation, in vitamins, cofactors, for signalling, pigments	2
10.	**Lipid conjugates:**Important conjugates and their functions: phospholipids, glycerophospholipids, sphingolipids; structure of membranes, micellar structures, Cholesterol- Structure, function and significance	3

11.	**Liposomes:** Liposomes in drug delivery, soaps, bio-surfactants	**1**
	Physico-chemical properties of carbohydrates and lipids: Compare and contrast of properties arising due to differences in C:O ratio	

UNIT II

12.	**Amino acids, peptides and proteins:** Occurrence/sources, Naturally occurring amino acids; structures; abbreviated names; stereoisomerism; amphoteric nature of amino acids, Classification of amino acids (on basis of functional groups, essential/non-essential); Chemical reactivity due to functional group (amide, acid, amine); peptides & proteins peptide bond, bond properties, N-Terminal, C-Terminal, Complex proteins– Proteoglycans, Lipoproteins, metalloproteins, phosphoproteins, chromoproteins and their significance	**3**
13.	**Protein structure:** Primary, secondary (α helix and β conformation), tertiary and quaternary structure. Forces stabilizing molecular structure (covalent bond, ionic bond, hydrogen bond, salt linkage, van der Waal's forces) a brief over view of Ramchandran plot, Structure-Function relationship in proteins with examples of Fibrous proteins, Collagen, Hemoglobin; importance of sequence of amino acids for structure and function of protein (eg. Sickle cell anaemia),	**4**
14.	**Introduction to enzymes as biocatalysts:** Enzyme-substrate interaction, Rates of enzyme reaction, Concept of Km, Vmax	**4**

References:

1. Biochemistry, Berg, J.M., Tymoczko, J. L. and Stryer, L. W.H. Freeman and Co., New York, USA (2003).

2. Principles of Biochemistry by Lehninger, A., Nelson, D. L. and Cox, M.M. W.H. Freeman and Co., New York, USA (2008).
3. Biochemistry, Satyanarayan, U. Books and Allied (P) Ltd., Kolkata, India (2008).

MBI 103: Mathematics (C) — Total

Basic Course – Theory; 2 Credits — **30L**

UNIT I

1.	Calculus: Limits, Continuity, Differentiation (1D & Partial), Definite Integrals.	**8**
2.	Numerical Techniques: Vector algebra, Matrices, Finding Eigen values & Eigen vectors	**3**
3	2Dimensional & 3Dimensional Geometry: Cartesians and Polar Coordinates, Locus, Equation of a straight line, Pair of straight lines, Circle, Ellipse, Parabola, Hyperbola. 3D Coordinate System, Spherical and Cylindrical Coordinates.	**4**

UNIT II

4.	Ordinary & Partial Differential Equation: 1'st Order & 2'nd Order Ordinary Differential Equations. Nature of Partial Differential Equations, Method of Separation of Variables, Methods for Solving Equations, Solving ODE & PDE.	**5**
5.	Integral transform: Fourier Series, Fourier Transform, Laplace Transform	**7**
6.	Applications To Mathematical Biology: Enzyme kinetics, Immunology, Population genetics, Tumor modeling, Applications of ordinary & partial differential equations to Biology.	**3**

References:

1. Introduction to mathematical methods in bioinformatics by Isaev, Alexander Berlin Springer, 2004.
2. Mathematics in chemistry by K. V. Raman & Pal, Sourav, New Delhi, Vikas publishing house Pvt. ltd., 2005
3. Calculations in molecular biology and biotechnology: a guide to mathematics in the laboratory by Stephenson, F.H. Amsertdam, Academic Press, 2003.
4. Advanced mathematical methods for engineering & science students by Stephenson, G. & Radmore, P. M., 1990.
5. Mathematics and computer science in medical imaging by Viergever, Max A.& Todd-Pokropek, Andrew.,1988.
6. Mathematical Methods for Physicists by G. Arfken, Academic Press, New York,1970.
7. Some Modern Mathematics for Physicists and Other Outsiders by P. Roman, Pergamon, New York, 1975, Vol. 2, p. 660.
8. Mathematical Biology, by J. D. Murray Springer Verlag, 1989.
9. Mathematical Models in Molecular and Cellular Biology by Segal, L., ed. 1980. Cambridge: Cambridge University Press.
10. Numerical Methods by Balaguruswamy, TMH.

MBI 104: Biostatistics (C)

Basic Course – Theory; 2 Credits

Total 30L

UNIT I

1.	Probabilities, Random Variables, Multiple Random Variables, Distributions, Random Sampling, Maximum Likelihood Estimators, Bayes Estimators, Mean Squared Error. Hypothesis Testing., t-tests, Likelihood Ratio Tests , Interval Estimation	**7**
2.	Analysis of Variance ,One-Way Analysis of Variance , Two-Way Analysis of Variance , Regression Models, Simple Linear Regression, Logistic Regression	**8**

UNIT II

3.	Bayesian Method with Examples, basics of Markov chains, Dynamic Programming and Hidden Markov Model, Metropolis–Hastings Algorithm and Gibbs Sampling. Gene Expression and Microarray Analysis ,Unsupervised Learning , Supervised Learning Sequence Alignment, Pair-Wise Sequence Analysis, Multiple Sequence Alignment , Sequence Pattern Discovery	**7**
4.	Some Common Tools and Techniques : (Brief introduction and case studies only) Classification and clustering (Use WEKA) Artificial Neural Networks. Fuzzy Sets and Fuzzy Logic Genetic Algorithms	**8**

Note: Practice on XLminer analysis tool-pack or R (no hand computations). Theory to focus on basic concepts, applications and interpretations/inferencing with secondary data, no mathematical derivations

References:

1. Basics of Bioinformatics, Rui Jiang Xuegong Zhang Michael Q. Zhang Editors
2. “Basic Statistics for Bioinformatics,”(free open-source package called R. /XL Miner for solving problems / No hand computation, Use of XL-miner in practicals - 10 excercises, focus should be on estimating and interpreting outputs)
3. Probability statistics, and reliability for engineers by Boca Raton, Ayyub B. M. & McCuen, R H, CRC Press, 1997.
4. Statistics: concepts and applications by Frank, Harry & Althoen, S. C., Cambridge University Press, 1995.
5. Statistical methods in bioinformatics: an introduction by Ewens, W. J. & Grant, G. R., New York. Springer, 2001.

MBI 105: C Programming and Data Structure (C) Total

Basic Course – Theory; 3 Credits 45L

UNIT I

1. Introduction and First Program **3**

- Why Programming
- Types of Programming
- Introduction to C
- C programming features
- Benefits of C
- Some Facts about C
- Understanding First C Program

2. Variables and Data Types **3**

- Identifiers
- Keywords
- Data Types
- Variables
- Constants

3. Console IO Operations **2**

- printf function
- scanf function
- Unformatted Functions

4. Operators and Expressions **2**

- Expressions
- Types of Operators
- Type Casting

5. Control Flow Statements **2**

- Decision Making in C
- If Statement
- Switch Statement
- Unconditional Branching
- While Loop
- Do…While Loop
- For Loop
- Break and continue statements

UNIT II

6. Working with Functions **5**

- What is a Function
- Benefits of a Function
- Function Terminology
- Array of Structures
- How does Function Works
- Scope and Lifetime of Variables in function

- Storage Classes of Variables
- Call by value and call by reference
- Recursion

7. Working with Arrays and Strings **5**

- Understanding Arrays
- Arrays Declaration and Initialization
- Sample Programs
- Multidimensional Arrays.
- Arrays and Functions
- String operation

UNIT III

8. Pointers **4**

- Understanding Pointers
- Declaring and Initializing Pointers
- Function and Pointer Parameters
- Pointer Arithmetic
- Pointer and Arrays
- Two Dimensional Arrays and Pointers
- void Pointer
- Dynamic allocation of memory
- Difference between malloc and calloc

9. Structure and Unions **4**

- Overview of Structures
- Defining and Using a Structure
- Structures within a Structure
- typedef keyword
- Passing Structures to Functions
- Structure and Pointers
- Unions

10. File Handling **4**

- What is a Stream
- Opening and Closing of Files
- Writing and Reading in Text Format
- Writing and Reading in Binary Format

UNIT IV

11. Introduction to Data Structures **2**

Data Structures- Data structure, Abstract Data Types (ADT), Concept of linear and Non-linear, static and dynamic data structures, and relationship among data, data structure

10. Linked Lists **3**

Concept, Comparison of sequential and linked organizations, Primitive operations, Realization of Linked Lists, Linked list operations, Head

	pointer and header node, Types of linked list- Linear and circular linked lists, Doubly Linked List and operations, Circular Linked List, Singly circular linked list, Doubly circular linked list.	
11.	Stacks Stacks- concept, Primitive operations, stack operations, Applications of Stack- Expression, Evaluation and Conversion, Need for prefix and postfix expressions, Postfix expression evaluation, Recursion- concept, Backtracking algorithmic strategy, use of stack in backtracking.	**2**
12.	Queues Concept, Realization of Queues Using Arrays , Circular Queue, Advantages of using circular queues, operations on queue	**2**
13.	Sorting and Searching Searching- Search Techniques, Sequential search, Binary search. Sorting methods- Bubble sort, Insertion sort, Selection sort, Quick sort, Heap sort, Shell sort, Comparison of All Sorting Methods.	**2**

References:

1. C& Data structures - P. Padmanabham, B.S. Publications.
2. The C Programming Language, B.W. Kernighan, Dennis M.Ritchie, PHI/Pearson Education.
3. C Programming with problem solving, J.A. Jones & K. Harrow, Dreamtech Press.
4. Programming in C - Stephen G. Kochan, III Edition, Pearson Eductaion.
5. Data Structures and Program Design in C, R.Kruse, C.L. Tondo, BP Leung, Shashi M, Second Edition, Pearson Education.
6. Computer science, A structured programming approach using C, B.A. Forouzan and R.F. Gilberg, Third edition, Thomson.
7. DataStructures Using C - A.S.Tanenbaum, Y. Langsam, and M.J. Augenstein, PHI/Pearson education.
8. Let us C – Yashwant Kanetkar

MBI 106: Biological Informatics (C)	**Total**
Basic Course – Theory; 2 Credits	**30L**

UNIT I

1. Introduction to Bioinformatics	**5**

- Nature of biological data types
- Overview of available Bioinformatics resources on the web
- Primary Resource Institutes: NCBI, EMBL & DDBJ
- Hierarchy of Biological databases: Primary, Secondary & Derived

2. Biological Databases & Tools	**10**

- Database search engines: Entrez, SRS
- Nucleic acid databases: GenBank, ENA, Gententry
- Protein sequence databases: NCBI Protein, EMBL Protein, PIR-PSD , SwissProt/ UniProtKB/ TrEMBL, Expasy
- Structural Databases: PDB, SCOP, CATH, NDB, CCSD, CSD
- Molecular visualization tools: RasMol, Cn3D, SPDBV, Chime, Mol4D, etc
- Databases and search methods for chemical compounds: PubChem Compound, PubChem Substance, ChEBI, ChEMBL, PDBeChem, RESID, EuroCarbDB
- Sequence Submission Tools: Sequin, BankIt, ENA, IMGT/HLA, DGVa, SPIN, Metagenomics

UNIT II

3. Overview/concepts in Sequence Analysis	**7**

- Local & Global alignment, DotPlot, Gap Penalties
- Dynamic Programming, Heuristic Methods
- Pairwise Sequence Alignment algorithms: Needleman & Wunsch, Smith & Waterman
- Scoring matrices for Nucleic acids and proteins: PAM/MDM, BLOSUM, CSW

4. Database Similarity Searches	**8**

- BLAST & FASTA

Other Tools: LALIGN, Dotlet
- Multiple Sequence Alignment: ClustalW, ClustalX, PRAS
  Other Tools: DbClustal, Kalign, MAFFT, MUSCLE, MView, T-Coffee
- Motifs, Pattern & Profiles
- Derived Databases: PROSITE, BLOCK, ProDom, Pfam, PRINTS, SBASE

References:

1. Introduction to Bioinformatics by Attwood, T.K. & Parry-Smith, D.J., Delhi, Pearson Education (Singapore) Pte.Ltd., 2001.
2. Bioinformatics: Sequence and Genome Analysis by Mount, David, New York, Cold Spring Harbor Laboratory Press, 2004.
3. Current Protocols in Bioinformatics by Baxevanis, A.D., Davison, D.B., Page, R. D. M. & Petsko, G.A., New York, John Wiley & Sons Inc., 2004.
4. Claverie, J.M. and Notredame C. 2003 Bioinformatics for Dummies. Wiley Editor.
5. Letovsky, S.I. 1999 Bioinformatics. Kluwer Academic Publishers.
6. Baldi, P. and Brunak, S. 1998 Bioinformatics. The MIT Press.
7. Setubal, J. and Meidanis, J. 1996 Introduction to Computational Molecular Biology. PWS Publishing Co., Boston.
8. Lesk, A.M. 2002 Introduction to Bioinformatics. Oxford University Press.
9. Rastogi, S.C., Mendiratta, N. and Rastogi, P. 2004 Bioinformatics: Concepts, Skills & Applications. CBS Publishers & Distributors, New Delhi.
10. Vyas, S.P. and Kohli, D.V., Methods in Biotechnology and Bioengineering.
11. Genetic Library Construction and Screening: Advanced Techniques and Applications: Lab Manual
12. Mont, D.W., Bioinformatics: Sequence and Genome Analysis.
13. Pierre Baldi and Soren Brunak, Bioinformatics: The Machine Learning Approach.

MBI 107: DBMS & MongoDB (C) **Total**

Basic Course – Theory; 3 Credits **45L**

UNIT I

1. DBMS **10**

- Database designing, data capturing
- Data Abstraction
- Data Models
- Instances & Schemes
- E-R Model - Entity and entity sets
- Relations and relationship sets
- E-R diagrams

Reducing E-R Diagrams to tables

UNIT II

2. Basic concepts in Indexing and hashing **15**

- Types of Indexing
- Data warehousing
- Data mining
- Oracle Architecture
- Basic concepts in Oracle:
    - Table space
    - Data files
    - Blocks
    - Extents
    - Segments
    - Oracle Background Processes
    - Control files
    - Oracle Memory Management
    - Rollback
    - Redo logs etc.
- Oracle Report generation, Grants, Roles, Privileges
- Introduction to SQL

UNIT III

3. MongoDB **10**

- Introduction of mongoDB
- Uses and Advantages
- RDBMS/SQL vs. MongoDB
- Structure of MongoDB
- Database, Collection – operation
- Data type

UNIT IV

4. **10**

- CRVD commands
- Limit Records, Sorting Record
- Indexing, Aggregation
- Covered queries
- ObjectID
- Regular Expression

References:

1. Database System Concepts by Hanery Korth and Abraham Silberschatz, McGraw Hill publication.
2. An Introduction to Database Systems by C.J. Date, Addison-Wesley.
3. Database systems by Nilkamal Surve, Tech Max publications.
4. Data Mining: Concepts and Techniques by Jiawei Han and Micheline Kamber, Morgan Kaufmann Publishers.
5. Oracle PL/SQL Programming by Steven Feuerstein, Bill Pribyl, O-Reilly Media.
6. The Relational Database Dictionary by C.J. Date, O'Reilly Media
7. Fundamentals of Database Systems by Elmasri and Navathe, Pearson Education.
8. Database Design and Relational Theory by C.J. Date, O'Reilly Media.
9. Oracle Database 11g A Beginners Guide by Ian Abramson, Michael Abbey, Michael J. Corey and Michelle Malcher, McGraw Hill publication.
10. Developing Bioinformatics Computer Skills by Gibas, & Jambeck, O-Reilly.
11. Linux : The Complete Reference 6th Edition by Richard Petersen, Tata McGraw-Hill Education (2007)

MBI 108: PERL Programming **Total**
Elective Course – Theory; 2 Credits **30L**

UNIT I

1. **Introduction to Perl** **3**
   Perl Data types, Operators, Input/Output in Perl
2. String Functions, Array Functions, Hash Functions **3**
3. Control Statements and loops **4**

UNIT II

4. **Subroutine** **3**
   Defining subroutine, Calling subroutine, Passing Arguments to a Subroutine
   Returning Value from a Subroutine
5. Regular Expression, Pattern Matching, Referencing and Dereferencing **3**
6. **File handling** **4**
   File I/O, Opening and Closing Files, Reading and Writing Files

UNIT III

7. **Object Oriented Programming in Perl** **3**
   Objects, Methods, and Classes in Perl
8. Packages in Perl **3**
9. **Perl and Web** 4
   CGI programming, Web browsing

References:

1. Beginning Perl for Bioinformatics by James Tisdall, O-Reilly publication.
2. Perl: Complete Reference Perl by Martin C. Brown, McGraw-Hill publication.
3. Mastering Perl for Bioinformatics by James D. Tisdall, O-Reilly

Publication.

4. Teach Yourself Perl 5 in 21 days by David Till, Sams publishing.
5. Mastering Algorithms with Perl by Jon Orwant, Jarkko Hietaniemi and John Macdonald, O-Reilly Publication.
6. Professional Perl Programming by Peter Wainwrigh, Published by Wrox Press Ltd.
7. Beginning Web Development with Perl by Steve Suehring, Apress publication.

MBI 108: HTML Programming **Total**
Elective Course – Theory; 2 Credits **30L**

UNIT I

1.	**Introduction to HTML** HTML basic tags, Formatting Tags, Colour codings	**5**
2.	Lists, Tables, Headers	**5**
3.	Images, Frames	**5**

UNIT II

4.	**CSS,** Basis Syntax, Single style sheet, Multiple style sheets	**5**
5.	CSS text fonts, CSS backgrounds	**5**
6.	CSS lists, CSS links	**5**

References:

1. Learning Web Design- a beginner's guide to HTML, graphics, and beyond by Jennifer Niederst, Publisher: O'reilly digital studio.
2. HTML Black Book by Steven Holzner, Publisher: Dreamtech Press.
3. Head First HTML with CSS & XHTML, by Elisabeth Robson and Eric Freeman, Publisher: O'reilly Media.
4. XML Pocket Reference by Robert Eckstein and Michel Casabianca, Publisher: O'reilly.
5. XML Schema by Eric van der Vlist, Publisher: O'reilly.
6. Learning XML by Erik T. Ray, Publisher: O'reilly.

MBI 109:Cell Biology and Biochemistry Lab (C) Total
(2 Credits, 1 P, 4.00 Hr., Basic Course Practical) 20 P

Cell Biology Lab

1.	Chloroplast isolation from spinach leaves	**1**
2.	Staining of prokaryotic cell organelles: Staining of capsule, spore, cell wall and metachromatic granule	**2**
3.	Study of mitosis with onion root tip chromosomes	**2**
4.	Observation of permanent slides of meiosis	**1**
5.	Isolation of nuclei from rat liver	**2**
6.	To determine Erythrocyte (*RBC*) count of a blood sample	**1**
7.	To determine Leucocytes (*WBC*) count of a blood sample	**1**
8.	Temporary mounting of mitochondria by Janus green B	**1**

Biochemistry Lab

9.	Preparation of buffers-acetate buffer & Preparation of biochemical reagents (Benedict's reagent)	**2**
10.	Isolation of biomolecules a. Isolation of starch from corn (separation on the basis of density) b. Isolation of protein from a suitable source c. Extraction of triglycerides from oilseeds (separation on the basis of differential solubility)	**2**
11.	Quantitative estimation of Glucose by DNSA method	**1**
12.	Quantitative estimation of Protein by Biuret method and absorption at 280 nm	**2**
13.	Acid value or saponification value. Determination with reference to fatty acids.	**2**

References:

1. An Introduction to Practical Biochemistry, Plummer, D.T., Tata-McGraw-Hill Publishing Co., New Delhi (2005).

2. Principles and Techniques of Biochemistry and Molecular Biology, Wilson, K. and Walker, J., Cambridge University Press, New York (2005).

3. Laboratory Manual in Biochemistry, Jayraman J., New Age International (P) Ltd., New Delhi (2007).

4. Alberts B. and Jhonson A. 4th edition (2002) Molecular Biology of the cell, Garland science.

5. Berg J., Tymcoczko J, and Styrer L, 5th edition (2002) Biochemistry, W. H. Freeman and company, New York.

6. Cooper G.M., Hausman R. E. The cell: A molecular approach. 5th edition. ASM Press and Cinauer Associates Inc. 2009

7. Practical Microbiology: Principles and Techniques, (2005), 1st Edn., Kale, V and Bhusari, K. Himalaya Publishing House, New Delhi.

MBI 110: C Programming and Data Structure Lab (C) **(2 Credits, 1 P, 4.00 Hr., Basic Course Practical)**	**Total** **15 P**
1. Laboratory assignments based on the following topics in 'C' programming:	**5**

- Data types, operators and expressions, Hierarchy of operators,
- control statements including decision (if, if-else), loops (while, do-while, for), branching statements(switch, break, continue),
- Functions,
- Arrays (1D, 2D- all matrix operations including inverse of a matrix),
- Strings,
- File handling etc.

2. Writing C programs for Bioinformatics applications:	**10**

- Extract a protein or nucleic acid sequence from any of the databank files (GenBank entry, Swiss-Prot, EMBL entry etc.)
- Interconverting the sequence from one databank format to the other eg. GenBank format to FASTA format, FASTA to PIR format etc.
- Determining the base composition in a nucleic acid sequence and amino acid composition in a protein sequence.
- Generating the complimentary sequence of a DNA sequence o Pattern search algorithms o Search for a specific oligonucleotide pattern (eg. GAACATCC) in a given DNA sequence.
- Find the position where a specific sequence say "GGTCCCGAC" will hybridize a given DNA sequence.
- Find the restriction enzyme cleavage sites eg. where PVUZ, ECORI etc. will cut the DNA.
- Locate palindromic sequence stretches in a DNA sequence. Count the number of Open Reading frames (ORF's) in a DNA sequence.
- Calculate the codon usage in a nucleic acid sequence.
- Translate a DNA sequence into protein sequence in the forward and reverse frames.
- Implementation of the Needleman-Wunsch algorithm for pair wise

alignment and testing alignment score with randomized pairs of sequences also.

- Numerical Techniques (4 assignments)
- Basic Formalism, Methods for Solving Equations, Finding Eigen values & Eigenvectors (5 assignments)

References:

1. Let Us C by Yashavant Kanetkar, BPB Publications.
2. The C programming language by Kerighan and Richie, PHI Publication.
3. Programming in ANSI C by Balaguruswamy, Tata McGraw-Hill Education.
4. Sams Teach Yourself C in 21 Days Peter Aitken and Bradley L. Jones, Macmillan Computer publishing.
5. Schaum’s outline of programming with C by Byron Gottorfried.

MBI 111: Biological Informatics lab (C)
(2 Credits, 1 P, 4.00 Hr., Basic Course Practical)

		Total 15 P
1.	Exploring the integrated database system at NCBI server and querying the PUBMED and GenBank databases using the ENTREZ search engine.	**1**
2.	Exploring the integrated database system at EBI server and searching the EMBL Nucleotide database using the SRS search engine.	**1**
3.	Exploring & querying SWISSPROT & UniProtKB.	**1**
4.	Exploring and querying the PIR database.	**1**
5.	Exploring tools on ExPASY.	**1**
6.	Exploring utilities in EMBOSS packages.	**1**
7.	Explore Pair-wise global alignments.	**1**
8.	Explore Pair-wise local alignments.	**1**
9.	Database (homology) searches using different versions of BLAST and FASTA interpretation of the results to derive the biologically significant relationships of the query sequences (proteins/DNA) with the database sequences. Exploring other databases: LALIGN, Dotlet	**2**
10.	Multiple sequence alignments: CLUSTALW, Clustal Omega, DbClustal, Kalign, MAFFT, MUSCLE, MView, T-Coffee.	**2**
11.	Exploring Alignment Analysis tools: AMAS, CINEMA, MaxAlign, PhyloGibbs, SVA, PVS.	**1**
12.	Exploring and using the derived databases: PROSITE, PRINTS, BLOCKS, Pfam and Prodom for pattern searching, domain searches etc.	**1**
13.	Studying the format & content of structural databases & visualization of structure using Rasmol, Cn3D and other utilities.	**1**

References:

1. Bioinformatics: A Practical Guide to the analysis of Genes and Proteins (3rd Ed.) by Baxevanis, A.D. & Ouellettee, B., F. F., New York, John Wiley & Sons, Inc. Publications, 2004.
2. Practical Bioinformatics, Michael Agostino, 1stEdition,

ISBN:9780815344568 , September 26, 2012

3. Bioinformatics: A Practical Approach, Shui Qing Ye, CRC Press, 20-Aug-2007
4. Bioinformatics: Methods Express, Paul Dear (Editor) , Publication Date: September 18, 2007 | ISBN-10: 190484216X | ISBN-13: 978-1904842163 | Edition: 1st Edition
5. Current Protocols in Bioinformatics by Baxevanis, A.D., Davison, D.B., Page, R. D. M. & Petsko, G.A., New York, John Wiley & Sons Inc., 2004.
6. Bioinformatics: Sequence, Structure and Databanks: A Practical Approach [Paperback]

MBI 112: DBMS & MongoDB lab (C) **Total**

(2 Credits, 1 P, 4.00 Hr., Basic Course Practical) **20 P**

1. **Part I: ORACLE** **2**

   Assignment based on "Data Definition Language".
   - A set of SQL commands used to create table, modify table structure, drop table, rename table.
2. Assignment based on "Data Manipulation Language" **2**
   - A set of SQL commands used to change the data within the database.
   - Insertion of records in the tables, updating of all or specific set of records in tables.
   - Viewing the attributes of table's column.
3. Assignment based on "Data Query Language" **2**
   - Different forms of Select statement
   - Arithmetic and logical operators.
   - Range searching and pattern matching.
4. Assignment based on defining Constraints. **2**
   - Types: I/O constraints like Primary Key, Foreign key, Null and Unique
   - Constraints. Business constraints like check constraints.
5. Assignment based on using joins. **2**
   - Joining multiple tables, joining a table to it.
6.
   - Granting rights on user objects such as Tables, Views, and Sequences. 2
   - Revoking rights on user objects such as Tables, Views, and Sequences.
   - XAMPP for MySQL

7. **Part II: MongoDB based practical** **8**

References:

1. Database System Concepts by Hanery Korth and Abraham Silberschatz, McGraw Hill publication.

2. An Introduction to Database Systems by C.J. Date, Addison-Wesley.
3. Database systems by Nilkamal Surve, Tech Max publications.
4. Data Mining: Concepts and Techniques by Jiawei Han and Micheline Kamber, Morgan Kaufmann publishers.
5. Oracle PL/SQL Programming by Steven Feuerstein, Bill Pribyl, O-Reilly Media.
6. The Relational Database Dictionary by C.J. Date, O'Reilly Media
7. Fundamentals of Database Systems by Elmasri and Navathe, Pearson Education.
8. Oracle Database 11g A Beginners Guide by Ian Abramson, Michael Abbey, Michael J. Corey and Michelle Malcher, McGraw Hill publication.

MBI 113: PERL Programming Lab **Total**

(2 Credits, 1 P, 4.00 Hr., Elective Course Practical) **20P**

1 **Writing programs in Perl based on following assignment topics:** **10**

- Arithmetic and Logical operators
- Control statements and Loops
- Functions of Strings, Arrays, Hashes
- Input/output in Perl
- Subroutine
- Regular Expression
- Pattern Matching
- Working with files
- OOPs programming
- CGI programming

2 **Writing programs for following:** **10**

- Programs for Transcription. (DNA to RNA sequence)
- Programs for Translation. (RNA to Protein Sequences)
- Programs for finding Open Reading Frame.
- Programs for count bases in Sequence.
- Programs for finding Triplet Codon.
- Programs for multi FASTA files.

References:

1. Beginning Perl for Bioinformatics by James Tisdall, O-Reilly publication.
2. Perl: Complete Reference Perl by Martin C. Brown, McGraw-Hill publication.
3. Mastering Perl for Bioinformatics by James D. Tisdall, O-Reilly Publication.
4. Teach Yourself Perl 5 in 21 days by David Till, Sams publishing.
5. Mastering Algorithms with Perl by Jon Orwant, Jarkko Hietaniemi and John Macdonald, O-Reilly Publication.
6. Professional Perl Programming by Peter Wainwrigh, Published by Wrox Press Ltd.
7. Beginning Web Development with Perl by Steve Suehring, Apress publication.

MBI 113: HTML Programming lab **Total**

(2 Credits, 1 P, 4.00 Hr., Elective Course Practical) **20P**

1. **Writing programs in HTML based on following assignment topics:** **10**
   - HTML basic tags
   - Formatting Tags
   - Colour coding
   - Lists
   - Tables
   - Headers
   - Images
   - Frames

2. **Writing programs in CSS based on following assignment topics:** **10**
   - CSS tags
   - Single style sheet
   - Multiple style sheets
   - CSS text fonts
   - CSS backgrounds
   - CSS lists
   - CSS links

References:

1. Learning Web Design- a beginner's guide to HTML, graphics, and beyond by Jennifer Niederst, Publisher: O'reilly digital studio.
2. HTML Black Book by Steven Holzner, Publisher: Dreamtech Press.
3. Head First HTML with CSS & XHTML, by Elisabeth Robson and Eric Freeman, Publisher: O'reilly Media.
4. XML Pocket Reference by Robert Eckstein and Michel Casabianca, Publisher: O'reilly.
5. XML Schema by Eric van der Vlist, Publisher: O'reilly.
6. Learning XML by Erik T. Ray, Publisher: O'reilly.

SEMESTER II

MBI 201: Statistical Analysis System (SAS) (C) **Total**
Core Course – Theory; 2 Credits **30L**

UNIT I

1. Introduction to SAS **1**
- An overview of SAS foundation
- Introduction to SAS programs submitting a SAS program
- Working with SAS program syntax

2. Accessing Data **1**
- Examining SAS data sets
- Accessing SAS libraries- Proc contents

3. Reading and Creating SAS Data sets **2**
- Introduction to SAS Data sets
- Reading from existing SAS Data Sets
- Reading Spreadsheet and Database Data
- Reading Raw Data Files
    - Reading standard delimited data
    - Reading nonstandard delimited data
    - Handling missing data
    - Reading raw data files with formatted input
    - Controlling when a record loads

4. Formatting Data Values **1**
- Using SAS formats
- Creating user-defined formats

5. Manipulating Data **2**
- using SAS functions
- conditional processing

6. Combining SAS Data Sets **2**
- Concatening
- Merging - one-one, one-many, merging with non-matches

7. Processing Data in groups **1**

8.	**Processing Data Iteratively** ➢ DO loop processing ➢ conditional DO loop processing ➢ SAS array processing ➢ using SAS arrays	**2**
9.	**Restuction / Rotating SAS Data Set**	**1**
10.	**Creating Summary Reports** ➢ Proc Print ➢ Proc Freq ➢ Proc Report ➢ Proc Tabulate ➢ Report Enhancement	**2**

UNIT II

11.	**SAS Macros Language** ➢ Purpose of Macro Facility ➢ Program Flow	**1**
12.	**Macro Variables introduction to macro variables** ➢ Automatic macro variables ➢ Macro variable references ➢ User-defined macro variables ➢ Delimiting macro variable references	**2**
13.	**Macro Definitions defining and calling a macro** ➢ Macro parameters ➢ DATA Step and SQL Interfaces creating macro variables in the DATA step	**2**
14.	**DATA Step and SQL Interfaces creating macro variables in the**	**2**

DATA step

- Indirect references to macro variables
- Creating macro variables in SQL

15.	**Macro Programs conditional processing**	**2**

- Parameter validation
- Iterative processing
- Global and local symbol tables

16.	**SAS SQL – Language**	**1**

- Introduction to SAS- SQL
- Basic Queries
- Overview of the SQL procedure
- Specifying columns
- Specifying rows

17.	**Types of Joins**	**1**
18.	**Introduction to Subqueries**	**2**
19.	**Set Operators**	**2**

References:

1. Berger, J. O. (1985), Statistical Decision Theory and Bayesian Analysis, 2nd Edition, New York: Springer-Verlag.
2. Cameron, A. C. and Trivedi, P. K. (1998), Regression Analysis of Count Data, Cambridge: Cambridge University Press.
3. Gallant, A. R. (1987), Nonlinear Statistical Models, New York: John Wiley & Sons.
4. Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004) , Bayesian Data Analysis, 2nd Edition, London: Chapman & Hall.
5. Powers, D. A. and Xie, Y. (2000), Statistical Methods for Categorical Data Analysis, San Diego: Academic Press.
6. Schervish, M. J. (1995), Theory of Statistics, New York: Springer-Verlag.

MBI 202: R and Data Analytics (C) **Total**

Core Course – Theory; 3 Credits **45 L**

UNIT I

1. An Introduction to R and Basic Programming **5**

- Overview
- Environment set up (Note: Only for Lab Demonstration)
- Data Types
- Variables
- Operator
- Decision making

2. R Core Programming 5

- Loops
- Functions
- Strings
- Vector
- List
- Matrix

UNIT II

3. R Core Programming **5**

- Arrays
- Factors
- Data Frames
- Packages
- Data Shaping
- Library

4. R Charts and Graphs 5

- R-Pie Chart

- R – Bar Chart
- R- Box Plots
- R - Histogram
- R- Line Graph
- R- Scatter Plots

UNIT III

5.	R Data Interfaces 1	**5**

- R-CSV
- R-Excel
- R-Binary files
- R-XML files

6.	R Data Interfaces 1	**5**

- R-JSON files
- R-Web Data
- R-Database
- R- NoSQL

UNIT IV

7.	Analytics with R Statistics	**6**

- Mean, Median, Mode
- Normal Distribution
- Multiple regression
- Supervised Model
- Unsupervised Model

8.	Algorithm	**9**

- Logistic regression
- Linear Regression
- Decision tree
- Random Forest

- SVM model

References:

1. Data Analysis with R: A comprehensive guide to manipulating, analyzing, and visualizing data in R, Tony Fischetti, 2nd Edition Paperback – Import, 28 Mar 2018
2. Python Machine Learning, Uxi (Hayden) Liu By Example Paperback,– Import, 31 May 2017
3. Data Analytics for Beginners: Basic Guide to Master Data Analytics Paperback – November 3, 2016, by Paul Kinley
4. Data Analytics Made Accessible: 2019 edition Kindle Edition by Anil Maheshwari
5. Learn R in a Day, 2013, Kindle Edition, by Steven Murray
6. Beginning R: The Statistical Programming Language, 2013 by Mark Gardener, Paperback
7. Statistical Programming in R, 5 June 2017 by K.G. Srinivasa and G.M. Siddesh Paperback

MBI 203: JAVA and BioJAVA Programming (C) **Total**

Core Course – Theory; 3 Credits **45L**

UNIT I

1. An Introduction to Java **2**

A Short History of Java

- Features or buzzwords of Java
- Comparison of Java and C++
- Java Environment
- Simple java program
- Java Tools – jdb, javap, javadoc
- Java IDE – Eclipse/NetBeans (Note: Only for Lab Demonstration)

2. An Overview of Java **4**

- Types of Comments
- Data Types
- Final Variable
- Declaring 1D, 2D array
- Accepting input using Command line argument
- Accepting input from console (Using Buffered Reader class)

3. Objects and Classes **6**

- Defining Your Own Classes
- Access Specifiers (public, protected, private, default)
- Array of Objects
- Constructor, Overloading Constructors and use of 'this' Keyword
- Static block, static Fields and methods
- Predefined class – Object class methods (equals(), toString(), hashcode(),getClass())
- Creating, Accessing and using Packages
- Creating jar file and manifest file

- Wrapper Classes
- Garbage Collection (finalize() Method)

UNIT II

4. Inheritance and Interface 7

- Inheritance Basics (extends Keyword) and Types of Inheritance
- Superclass, Subclass and use of Super Keyword
- Method Overriding and runtime polymorphism
- Use of final keyword related to method and class
- Use of abstract class and abstract methods
- Defining and Implementing Interfaces
- Runtime polymorphism using interface
- Object Cloning

5. Exception Handling 4

Dealing Errors

- Exception class, Checked and Unchecked exception
- Catching exception and exception handling
- Creating user defined exception
- Assertions

UNIT III

6. Strings, Streams and Files 5

- String class and StringBuffer Class
- Formatting string data using format() method
- Using the File class
- Stream classes
- Byte Stream classes
- Character Stream Classes
- Creation of files

- Reading/Writing characters and bytes
- Handling primitive data types
- Random Access files

7. Collection **6**

Introduction to the Collection framework

- List – ArrayList, LinkedList and Vector,Stack,Queue
- Set - HashSet, TreeSet, and LinkedHashSet
- Map – HashMap, LinkedHashMap, Hashtable and TreeMap
- Interfaces such as Comparator, Iterator, ListIterator, Enumeration

UNIT IV

8. Database Programming **6**

- The design of jdbc, jdbc configuration
- Types of drivers
- Create and Execute sql statements, query execution
- The Result Set Object

9. Multithreading **3**

- What are threads?
- Life cycle of thread
- Running and starting thread using Thread class
- Thread priorities
- Running multiple threads
- The Runnable interface
- Synchronization and interthread communication

10. Networking **2**

Networking basics – Protocol, Addressing, DNS, URL, Socket, Port

- The java.net package – InetAddress, URL, URLConnection class
- SocketServer and Socket class
- Creating a Socket to a remote host on a port (creating TCP

client and server)

- Simple Socket Program Example

References:

1. Introduction to Bioinformatics by Attwood, T.K. & Parry-Smith, D.J., Delhi, Complete reference Java by Herbert Schildt(5th edition)
2. Java 2 programming black books, Steven Horlzner
3. Programming with Java , A primer ,Forth edition , By E. Balagurusamy
4. Core Java Volume-I-Fundamentals, Eighth Edition, Cay S. Horstmann, Gary Cornell,
   Prentice Hall, Sun Microsystems Press
5. Core Java Volume-II-Advanced Features, Eighth Edition, Cay S. Horstmann, Gary
6. Cornell, Prentice Hall, Sun Microsystems Press

MBI 204: Science of Omics (C) **Total**

Core Course – Theory; 3 Credits **45L**

UNIT I

1. ➢ **Introduction to Genomics:** **5**

- ➢ Genome sequencing:
  - ○ Strategies & Approaches
  - ○ Conventional DNA sequencing methodologies
  - ○ NGS(Next generation sequencing)
    - ✓ Introduction, Next-generation sequencing methods,
    - ✓ NGS File formats (Recognizing different file formats related to genome sequencing data),
    - ✓ NGS data quality check & cleaning,
  - ○ Third generation sequencing

2. ➢ Genome Assemblies: **6**

- ○ Reference assembly, Assembly statistics & visualization,
- ○ De-novo assemblies & assemblers for genome makeup(MIRA, RAYMETA etc)

- ➢ Basic Aspects of Genome Annotation
- ➢ Genome mapping techniques: Genetic Mapping and Physical mapping
- ➢ Structural Genomics and Functional Genomics

UNIT II

3. **Genomics and Comparative genomics Databases:** **6**

- ➢ Genome Databases: Genome Sequence DataBase (GSDB), Genome Database (GDB), 1000 Genomes Project , AceDB, FlyBase, UCSC Genome Browser, Wormbase, MaizeGDB, MGI
- ➢ Comparative Genomics Databases: COG, VirGen, CORG, HOBACGEN, Homophila, XREFdb, Gramene
- ➢ Genetic Disorders Databases: OMIM, OMIA, Genetic Association

Database, Genetic Disorder Guide, IGDD, DisGenet, Genetic Disorder UK

4. Genomics and Comparative genomics Tools: 6

- Genome Alignments tools:BLAST2, MUMmer, PipMaker, VISTA
- Comparison of Gene Order :GeneOrder, Gene Synteny
- Prediction of genes: ORFs, Prediction of Signal sequences (Promoters, Primers, splice sites, UTRs etc.), Operons
- Primer Designing Tools: Primer, Primer3, NetPrimer, Primerfinder
- DNA/RNA Sequence Analysis: CENSOR, Gene Finder, GENEID, GenHunt, GENIE, GRAIL, ORD ID, ORF Finder, ORFGene, Pol3Scan, tRNAscan
- Translation Tools: Translation Tool, The Protein Machine, 6 Frame Translation Tool, Reverse Translation Tools
- Restriction Analysis Tools: WEB Cutter, ENZFINDER, TACG

UNIT III

5. Transcriptomics: 6

- Biology of Transcription
- Search for transcription factor binding sites
- RNA Sequencing & Techniques of RNA analysis RNA-Seq, Microarrays, Regulatory RNAs: small or large, Computational prediction of miRNA target genes, RNA Darkmatter
- Generating Transcriptome expression Data

6. Metagenomics: 6

- Introduction & Study of metagenomics
- Metagenomics Samples
- Qualitative and quantitative analysis of metagenome: DGGE, T-RFLP, RT-PCR, NGS
- Sequence to identification of biodiversity
- Metagenomics Algorithms (metagenomics data analysis): RAST, IMG/M,

MEGAN

UNIT IV

7. Metabolomics: **5**

- Fundamental concept,
- Tools of metabolomics- Capillary electrophoresis, Gas chromatography, Electrochemical detectors
- Case studies.

8. Lipidomics: **2**

Basic concepts and tools Case studies

9. Degradomics: **3**

- Techniques and concepts
- Approaches to identify the protease and protease-substrate repertoires, or
- 'degradomes', on an organism-wide scale
- Uncover new roles for proteases in vivo.
- Identification of new pharmaceutical targets to treat disease (Emerging degradomic)

References:

1. She Has Her Mother's Laugh: The Powers, Perversions, and Potential of Heredity Hardcover – May 29, 2018, Dutton; 1st Edition by Carl Zimmer
2. The Gene: An Intimate History Paperback – May 2, 2017, **by** Siddhartha Mukherjee
3. Introduction to genomics by Arthur M. Lesk, 2007, Published by OUP Oxford
4. Bioinformatics and Functional Genomics, Textbook by Jonathan Pevsner, 2003, Wiley publication
5. Metabolomics- Methods and Protocols by Wolfram Weckwerth, Humana Press.
6. Lipidomics- Technologies and Applications by Kim Ekroos, Wiley-VCH.
7. Web/Journal Resources.
8. Transcriptomics: Expression Pattern Analysis, Virendra Gomase, Somnath Tagore; VDM Publishing, 2009 – Science

MBI 205: Proteomics (C) **Total**

Core Course – Theory; 2 Credits **30L**

UNIT I

1. Overview of Proteomics: **7**

- Introduction and scope of proteomics
- Protein separation techniques:
    - ion-exchange
    - size-exclusion and
    - affinity chromatography techniques
    - Polyacrylamide gel electrophoresis
    - Two dimensional PAGE for proteome analysis; Image analysis of 2D gels

2. **1**

- Determination of Amino acid composition
    - Hydrolysis
    - Separation
    - Quantitative analysis

3. **1**

- Protein sequencing Methods
    - Sanger's method
    - Edman's method

4. **3**

- Protein structure determination methods:
    - X ray crystallography,
    - Mass spectrometry
    - NMR

5. **Proteomics tools:** **3**

- Protein Databases
- Structural databases: PDB, MMDB, SCOP, CATH.
- 3D structure visualization tools: Rasmol, Pymol, SPDBV, Cn3D (already covered in MBI 106)
- Secondary structure prediction algorithms: Chou-Fasman, Jpred, Psipred, GOR methods; PHD
- 3D structure validation databases: PROSA, Ramchandran Plot,

Procheck
- Proteomics tools on Expasy

UNIT II

6.
- Protein modifications: Post transcriptional and post translational **9**
- Applications of proteome analysis to drug; Protein-protein interaction (Two hybrid interaction screening)
- Protein engineering, Protein chips and functional proteomics
- Clinical and biomedical application of proteomics
- Proteome databases
- Proteomics industries

7. Protein-protein interaction Databases : **4**
- BIND - Biomolecular Interaction Network Database,
- STRING
- DIP (Database of Interacting Proteins)
- PPI Server
- BIND - Biomolecular Interaction Network Database
- PIM -Hybrigenics
- PathCalling Yeast Interaction Database
- MINT - a Molecular Interactions Database
- GRID - The General Repository for Interaction Datasets
- InterPreTS - protein interaction prediction through tertiary structure

References:

1. Fundamentals of Data Mining in Genomics and Proteomics, By Werner Dubitzky, Martin Granzow, Daniel P. Berrar, 2007, Springer Science + Bussiness Media, LLC.
2. Protein Arrays, Biochips and Proteomics: The Next Phase of Genomic Discovery edited by Joanna S. Albala, Ian Humphery-Smith, ISBN-0-8247-4212-1, 2003, Marcel Dekker.
3. Proteomics: Methods Express (Methods Express Series) Paperback – Import,

2007, by C. David O'Connor (Editor), B. David Hames (Editor)

4. Proteomics: A Comprehensive Study of Proteins Hardcover – Import, 30 Jun 2017 by Tanner Perry (Editor)

5. Principles of Proteomics, English, Paperback, Twyman, 2004

6. Introduction to Proteomics -Tools for the New Biology by Daniel C. Liebler, Humana Press

7. Mass Spectrometry for Biotechnology by Gary Siuzdak, Academic Press.

8. Proteomics for Biological Discovery by Timothy Veenstra and John Yates, Wiley.

MBI 206: Molecular Biology **Total**
Core Course – Theory; 2 Credits **30L**

UNIT I

1.	Organization of prokaryotic genome, Structure of nucleosome and organization of chromatin, structure of chromosome, chromatin remodeling, Repetitive sequences, clusters and repeats, Concept of gene, Interrupted and uninterrupted genes, intron, exon and their relationships centromere and telomere, Genome sizes of different organisms, C Value	**2**
2.	Types of mutations, effect of mutations, mutation related inherited human disorders, Physical and chemical mutagenic agents, frequency of mutations, measuring mutation rate, somatic &germline mutations, role of mutations in evolution	**4**
3.	Salient features of genetic code, codon – anticodon recognition, Wobble hypothesis, exceptions to the universal genetic code	**3**
4.	DNA replication & repair DNA polymerases, mechanism of replication in prokaryotes and eukaryotes, DNA damage, Mechanisms of DNA repair in prokaryotes and eukaryotes, Excision repair, mismatch repair, recombination repair, error prone repair, SOS response, repair system in eukaryotic cells.	**6**

UNIT II

5.	Transcription and posttranscriptional mechanisms RNA polymerase and mechanism of prokaryotic transcription, Eukaryotic RNA polymerases and their promoters, activating transcription, role of enhancers, gene silencers, CpG Islands, post transcriptional modifications, RNA splicing reactions, catalytic RNA	**7**
6.	Regulatory RNA, Micro RNAs & RNA Interference	**4**
7.	Translation Mechanism of translation in prokaryotes and eukaryotes, post translational	**4**

modifications, transport of proteins, role of chaperons

References:

1. Watson J. and Stephen (2004) Molecular Biology of the Gene, Dorling Kindersley(India) pvt ltd,New delhi. Taylor and Francis group, NewYork.
2. Cooper G.M. and Hausman R.E. (2004) The Cell: A molecular approach, Sinauer Associates, Inc., ASM Press, Washington DC.
3. Lewin (2007) Genes IX: Pearson Prentice Hall, Pearson Education, Inc. Upper Saddle River, NJ 07458
4. Strickberger 1985, Genetics. Macmillan.
5. Russell 2002, Genetics. Benjamin.
6. Cornell, Prentice Hall, Sun Microsystems Press

MBI 207: Recombinant DNA Technology **Core Course –Theory; 2 Credits**	**Total** **30L**

UNIT I

1.	Scope of Genetic engineering	**1**
2.	Tools and techniques: DNA Purification, Hybridization, PCR and Types of PCR, Electrophoresis	**7**
3.	Enzymes used in gene manipulation: Restriction endonucleases, Methylases, DNA polymerases, Ligases, Kinases, Phosphatases, Topoisomearses, DNA labelling methods	**7**

UNIT II

4.	Cloning vectors of plasmids (pBR, pUC18) and λ phages, Cosmids, BACs, PACs.	**6**
5.	Introduction of DNA into living cells, Identification of recombinants,	**2**
6.	DNA libraries: construction and of Genomic library and cDNA library. Library screening	**3**
7.	Mapping and Sequencing genome: RFLP, SNP, AFLP, Physical mapping, Radiation Hybrid mapping	**4**

References:

1. Brown T. A., 7th edition (2016), Gene cloning and DNA analysis, Blackwell publishing, UK
2. Primrose S., Twyman R. M. , 8th edition (2016), Principles of Gene Manipulation and Genomics, Blackwell Publishing, UK

MBI 208: Structural Biology & Molecular Modeling

Core Course –Theory; 3 Credits

Total 45L

UNIT I

1.	Secondary and tertiary structures of macromolecules: protein, DNA and RNA	**3**
2.	Secondary structure prediction using Chou Fasman, GOR methods; analysis of results and measuring the accuracy of predictions using Q3, Segment overlap, Mathew's correlation coefficient.	**2**
3.	Fundamentals of the methods for 3D structure prediction (sequence similarity/identity of target proteins of known structure, fundamental principles of protein folding etc.)	**5**

UNIT II

4.	Homology Modeling, fold recognition, and ab-initio structure prediction methods - protocols/algorithms.	**5**
5.	3-D structure comparison and concepts: FSSP, CE, VAST and DALI, Fold Classes.	**6**
6.	Databases of structure-based classification: CATH and SCOP.	**2**

UNIT III

7.	General features- bond stretching, angle bending, improper torsions, cross terms, non-bonded interactions, point charges, calculation of atomic charges, polarization, Van Der Waals interactions, hydrogen bond interactions, Water models.	**3**
8.	Types of force field - all atoms force field, united atom force field, etc.	**4**
9.	Molecular Energy minimization methods: steepest descent, conjugate gradient– derivatives, First order steepest decent and conjugate gradients.	**3**
10.	Second order derivatives Newton-Raphson, Minima, maxima saddle points and convergence criteria.-non derivatives minimization methods, the simplex.	**3**

UNIT I V

11.	Molecular Dynamics Simulation: Newtonian dynamics, Periodic boundary conditions and minimum image convention, Potential truncation and shifted-force potentials, Neighbor list, Force calculations.	**3**
12.	Classical Monte Carlo: Random numbers, Evaluating integrals using random numbers, Importance sampling, Metropolis algorithm.	**3**
13.	Analysis of simulated trajectories: Radial distribution functions, Self diffusion coefficient, Time correlation functions.	**3**

References:

1. Molecular Modeling Principles and Applications (2nd Ed.) by Andrew R. Leach, 2001, Prentice Hall, USA.
2. Principles of Protein Structure by G. E. Schulz, 2009, Springer
3. Textbook of Structural Biology by Anders Liljas, Lars Liljas, Jure Piskur and GÃran Lindblom, World Scientific Publishic Co. Pvt. Ltd.
4. Structural Bioinformatics, Edited by Jenny Gu, Philip E. Bourne, Wiely Blackwell
5. Computational Structural Biology: Methods and Applications by orsten Schwede, Torsten Schwede, Manuel C. Peitsch, 2008, World Scientific Publishing Company.
6. Molecular Modeling Of Nucleic Acids, Leontis N.B. and SantaLucia J. Jr. (eds), 1998, American Chemical Society.
7. Molecular Modeling of Proteins by Andreas Kukol, 2008, Humana Press.
8. Molecular Modeling and Simulation: An Interdisciplinary Guide by Schlick T., 2012, Springer.
9. Biological Modeling and Simulation: A Survey of Practical Models, Algorithms, and Numerical Methods (Computational Molecular Biology) by Russell Schwartz, The MIT Press, Cambridge, Massachusetts, London, England

MBI 209: SAS and Data Analytics lab (C)	**Total**
(2 Credits, 1 P, 4.00 Hr., Elective Course Practical)	**15P**

1.	**SAS based practical**	**05**
2.	**Data Analysis lab**	**15**

- Install R from CRAN and setup R studio
- Write a Program to print list of Student, Marks and Percentage considering
    - Different data types variables in R using logical, numeric
    - integer, complex, character and raw
    - Valid/invalid variables declaration
    - Variable assignment
    - Getting variables in environment
    - Class of variables
    - Deleting variables
    - Vectors declaration
    - Different operators demonstration using Arithmetic Operators, Relational Operators, Logical Operators, Assignment Operators, Miscellaneous Operators
    - if-else demonstration for negative, positive and zero number test
    - Switch case for different operations
- Write program to classify odd and even number to demonstrate For, While and Repeat loops with break and next keywords
- Create a function to calculate Area and circumference of circle/rectangle
    - calling a function with/without arguments
    - calling by argument position and names
- Write R program to demonstrate string functions in R , String manipulations in R
    - Vector declaration
    - Accessing vector elements
    - Vector manipulations

- Create a list and modify
    - Name elements
    - Access elements
    - Manipulate the list
    - Merging list
    - convert list to vector
- Create a numeric matrix
    - Access elements
    - Matrix computation
    - Demonstration
- Create array name columns, rows, matrix, Access array elements matrix computations
- Create a categorical vector and convert it to
    - Factor
    - Factor validations
    - Generating levels
    - Order change
- Create a data frame
    - Get structure and summary
    - Extract data from DF
    - Expand the data frame add column and rows
    - Check packages
    - Install packages
    - Load package
    - Join cols/rows in DF
    - Merge DF
    - Melting and casting
- Read/write files and access from
    - CSV files
    - Binary file
    - Xls file
- Read access and manipulate JSON data

    - Download and load data from website
    - Connect to mysql/oracle DB and access data from table
    - Working with MongoDB/Redis
- Create data records and create
    - Plot Pie chart for Profit sharing among individuals. Percentage plotting , pie chart
    - Plot a bar graph for monthly revenue. Group bar graph/stacked chart
    - Plot a boxplot for match data mpg vs cylinder
    - Create a vector and plot a histogram , Ranged histogram
    - Create a vector and plot line, point both graph
    - Plot scatterplot for mtcars$wt and mtcars$mpg
- Create data from Hospital, Patient, and disease use case
    - Calculate mean, median, mode of a distribution for data
    - Demonstration of dnorm, pnorm, qnorm and rnorm
    - Demonstration of multiple regression with mtcar dataset
- Get records and perform below
    - Use case of admission into university
    - Use case for weight calculation
    - Use case for Titanic survival prediction
    - Use case for Car acceptability
    - Use case for Heart disease recognition with caret package

References:

1. Data Analysis with R: A comprehensive guide to manipulating, analyzing, and visualizing data in R, Tony Fischetti, 2nd Edition Paperback – Import, 28 Mar 2018
2. Python Machine Learning, Uxi (Hayden) Liu By Example Paperback,– Import, 31 May 2017
3. Data Analytics for Beginners: Basic Guide to Master Data Analytics Paperback – November 3, 2016, by Paul Kinley
4. Data Analytics Made Accessible: 2019 edition Kindle Edition by Anil Maheshwari
5. Learn R in a Day, 2013, Kindle Edition, by Steven Murray
6. Beginning R: The Statistical Programming Language, 2013 by Mark Gardener, Paperback
7. Statistical Programming in R, 5 June 2017 by K.G. Srinivasa and G.M. Siddesh Paperback

MBI 210: JAVA and BioJAVA programming Lab (C) **(2 Credits, 1 P, 4.00 Hr., Elective Course Practical)**	**Total** **20 P**
1. ➢ Class and Method Implementation by – ○ Method overloading ○ Constructor Overloading ○ Static members and methods	**2**
2. ➢ Implementation of Multiple Inheritance using Interface.	**1**
3. ➢ Implementation of Inheritance by ○ Method overriding ○ super constructor and super keyword ○ abstract class	**2**
4. ➢ Implementation of Package	**1**
5. ➢ Program to read basic data types from keyboard using Scanner/BufferedReader and check the entered values' data type for its appropriateness.	**1**
6. ➢ Exception Handling for – ○ Divide by zero error ○ Null values	**2**
7. ➢ Program to read the data from user and save it to two different files, display the contents and exchange the contents of those two files using IO package.	**2**
8. ➢ Synchronization of threads. ➢ Develop an animation program using Multithreading viz. Bouncing Ball. ➢ Implementation of Client / Server mechanism using Socket classes. ➢ Design Database program for Employee details and implement INSERT, ➢ SELECT, DELETE, UPDATE queries. ➢ Design concurrent server that will handle multiple clients using ➢ Multithreading.	**9**

- Develop a simple client-server application using RMI.
- Programs using IO streams.
- Programs using files.

References:

1. Introduction to Bioinformatics by Attwood, T.K. & Parry-Smith, D.J., Delhi, Complete reference Java by Herbert Schildt(5th edition)
2. Java 2 programming black books, Steven Horlzner
3. Programming with Java , A primer ,Forth edition , By E. Balagurusamy
4. Core Java Volume-I-Fundamentals, Eighth Edition, Cay S. Horstmann, Gary Cornell,
   Prentice Hall, Sun Microsystems Press
5. Core Java Volume-II-Advanced Features, Eighth Edition, Cay S. Horstmann, Gary
6. Cornell, Prentice Hall, Sun Microsystems Press

MBI 211: Omics Analysis Lab (C) **Total**

(2 Credits, 1 P, 4.00 Hr., Elective Course Practical) **20P**

1.	Browsing & viewing genome data Ensembl@EBI ➢ MapViewer@NCBI	**1**
2.	Explore comparative genomics resources and NCBI and EBI	**1**
3.	Viewing regions exhibiting Synteny ➢ Genome assembly ➢ Genome Annotation: Using integrated genome annotation servers such as the server developed at IMTech, Chandigadh (http://imtech.res.in/raghava/gp.html)	**1**
4.	Gene Prediction and Gene Modeling ➢ Prediction of genes and gene structures (gene modeling) using online (web) servers of different methods tailored for prokaryotic and eukaryotic organisms such as GLIMMER, GeneMark, Grail, GENSCAN etc. Interpretation of results and comparison with known gene models (where available). Evaluation of accuracy of the methods. ➢ Prediction of promoters using methods such as Neural Network Promoter Prediction (NNPP) at Berkeley Drosophila Genome Project server, Genome inspector for combined analysis of multiple signals in genomes etc. Using Promoter databases. ➢ Prediction of alternate splice sites using methods such as Splice Site Prediction by Neural Network (at Berkeley Drosophila Genome Project server), GenScan, NetGene2 GeneSplicer etc. Prediction of PCR primers using Primer 3, ePCR etc.	**2**

5.	Comparison of full / partial genomic sequences using following methods to identify conserved genes and map/compare the annotations of the two sequences ➢ BLAST2 ➢ MegaBLAST and Discontiguous MegaBLAST ➢ MUMmer ➢ PipMaker ➢ VISTA ➢ Artemis	**2**
6.	Compare gene order of given genomic sequences using the GeneOrder tool	**1**
7.	Explore and query the comparative genomics databases: COG, VirGen, CORG, HOBACGEN, Homophila, XREFdb, Gramene etc.	**1**
8.	Explore and query SNP and SNP-related databases	**1**
9.	RNA seq data Analysis	**1**
10.	Transcriptomics data analysis	**1**
11.	Metagenomics Data Analysis	**1**
12.	Explore and query the protein-protein interaction databases: DIP, PPI Server, BIND, PIM, PathCalling, MINT, GRID, InterPreTS	**2**
13.	Gene annotation of unknown sequence	**1**
14.	Functional Genomics Using primary databases (such as UniProt) and derived databases such as InterPro, PRINTS, BLOCKS, PRODOM, Pfam etc. along with advanced sequence analysis tools such as profiles searches, pattern searches for function annotation of genomic sequences. Validation and verification of results for known case studies ➢ Using sequence-based and structure-based Function Annotation Servers such as: ProKnow (http://www.doe-mbi.ucla.edu/Services/ProKnow/) ➢ Joined Assembly of Function Annotations (JAFA) at http://jafa.burnham.org/learnMore.html etc. which are integrated	**4**

services for function annotation

- ProFunc (http://www.ebi.ac.uk/thornton-srv/databases/ProFunc)

References:

1. Fundamentals of Data Mining in Genomics and Proteomics, By Werner Dubitzky, Martin Granzow, Daniel P. Berrar, 2007, Springer Science + Bussiness Media, LLC.

2. Protein Arrays, Biochips and Proteomics: The Next Phase of Genomic Discovery edited by Joanna S. Albala, Ian Humphery-Smith, ISBN-0-8247-4212-1, 2003, Marcel Dekker.

3. Proteomics: Methods Express (Methods Express Series) Paperback – Import, 2007, by C. David O'Connor (Editor), B. David Hames (Editor)

4. Principles of Proteomics, English, Paperback, Twyman, 2004.

5. Dale Jeremy, Schantz Malcolm Von. From genes to genomes: concepts and applications of DNA technology. Publisher: John Wiley & Sons Ltd., 2007. ISBN:9780470017340.

9. Brown, T.A., Genomes, 3rd edition Garland Science publishers, 2006. ISBN:
1859960294.

10. Primrose Sandy B., Twyman Richard. Principles of Gene Manipulation and Genomics. Publisher: Wiley-Blackwell. 7th edition 2006. ISBN: 1405135441.

MBI 212: Molecular Biology and Recombinant DNA Technology lab (2 Credits, 1 P, 4.00 Hr., Core Course Practical)

Total 15P

Molecular Biology Lab:

1.	Isolation of genomic DNA from bacterial cells/cell lines/blood /plant	**3**
2.	Analysis of DNA and RNA preparations by agarose gel electrophoresis	**2**
3.	Quantitation of DNA and RNA preparations and determination of purity by UV spectrophotometry	**1**

Recombinant DNA Technology lab:

4.	Competent cells preparation and GFP cloning in E.Coli	**2**
5.	Plasmid isolation from *E. coli*	**3**
6.	Restriction digestion	**2**
7.	Restriction mapping	**2**

References:

1. Molecular cloning: A laboratory Manual, Sambrook and Russell, Wiley New York, Volume I – III.
2. Watson J. and Stephen (2004) Molecular Biology of the Gene, Dorling Kindersley(India) pvt ltd,New delhi. Taylor and Francis group, NewYork.
3. Cooper G.M. and Hausman R.E. (2004) The Cell: A molecular approach, Sinauer Associates, Inc., ASM Press, Washington DC.
4. Lewin (2007) Genes IX: Pearson Prentice Hall, Pearson Education, Inc. Upper Saddle River, NJ 07458
5. Strickberger 1985, Genetics. Macmillan.
6. Russell 2002, Genetics. Benjamin.

MBI 213: Structural Biology and Molecular Modeling Lab **Total**
(2 Credits, 1 P, 4.00 Hr., Elective Course Practical) **15P**

Structural Biology:

1.	**Visualization Tools:** RasMol, Cn3D, SPDBV, Chime, PyMOL	**3**
2.	**Secondary Structure Prediction Tools:** PHYRE, PSI-PRED, PREDIC PROTEIN,APSSP2, YASPIN SECONDARY STRUCTURE PREDICTION TOOLS, JPRED3	**4**
3.	**Structural Databases:** FSSP, VAST, DALI, CATH, SCOP	**3**

Molecular Modeling:

4. ➢ Conformational Analysis **5**
➢ BABEL, MOPAC
➢ Molecular dynamics simulations
➢ Molecular Dynamics Visualization
➢ Homology Modeling

References:

1. Molecular Modeling Principles and Applications (2nd Ed.) by Andrew R. Leach, 2001, Prentice Hall, USA.
2. Principles of Protein Structure by G. E. Schulz, 2009, Springer
3. Textbook of Structural Biology by Anders Liljas, Lars Liljas, Jure Piskur and GÃran Lindblom, World Scientific Publishic Co. Pvt. Ltd.
4. Structural Bioinformatics, Edited by Jenny Gu, Philip E. Bourne, Wiely Blackwell
5. Computational Structural Biology: Methods and Applications by orsten Schwede, Torsten Schwede, Manuel C. Peitsch, 2008, World Scientific Publishing Company.
6. Molecular Modeling Of Nucleic Acids, Leontis N.B. and SantaLucia J.

Jr. (eds), 1998, American Chemical Society.

7. Molecular Modeling of Proteins by Andreas Kukol, 2008, Humana Press.
8. Molecular Modeling and Simulation: An Interdisciplinary Guide by Schlick T., 2012, Springer.
9. Biological Modeling and Simulation: A Survey of Practical Models, Algorithms, and Numerical Methods (Computational Molecular Biology) by Russell Schwartz, The MIT Press, Cambridge, Massachusetts, London, England.

SEMESTER III

MBI 301: Scientific Writing Skills **Total 30L**

Core Course – Theory; 2 Credits

Course Statement/Objectives:

This course aims to demystify the writing process and teach the fundamentals of effective scientific writing. Instruction will focus primarily on the process of writing and publishing scientific manuscripts but grant writing will also be addressed. The course will be presented in two segments:
Part (1) teaches students how to write effectively, concisely, and clearly.
part (2) takes them through the preparation of an actual scientific manuscript or grant.

The course may be taken for 2 units or 3 units. Students taking the class for 2 units will be asked to attend a weekly lecture and to complete short writing and editing exercises, including writing a real "Letter to the Editor" piece. Students taking the class for 3 units will additionally develop a manuscript of their choice and will meet periodically with the instructor for individual editing sessions.

1. Demystifying the Writing Process **15**

- What makes good writing? Are there "good writers" and "bad writers"?
- Words, word choice, the basic elements of sentences and sentence structure. Writing in the active voice. Introduction to Bioinformatics
- Dissecting the news article: News-writing is the art of maximizing information and minimizing words; it's the barest-bones form of writing. The fundamentals of good writing can be learned by dissecting news articles.
- In-Class Exercise: Sorting through news articles.
- Writing Basics I: Punctuation and Parallelism. Tricks for clarity, brevity, and finesse.
- In-Class Exercise: Peer interviews and write-up mini-profiles
- Writing Basics II: Paragraphs, logic, and organization. Organizational strategies.
- In-Class Exercise: Peer interviews and write-up mini-profiles (swap)
- Writing Basics III: Putting it all together…
- In Class exercise: group rewrites of hard-to-read scientific

snippets

2. Good Writing Applied: The Scientific Manuscript 10

- Methods and Results Sections: How to present data effectively. How to write prose that complements a table or figure.
- In-Class Exercise: Discuss a variety of journal articles that present data in different ways; rewrite a results paragraph.
- The Abstract, Introduction, and Discussion: Getting to the main point and summarizing effectively. How to conduct literature reviews. Writing an effective discussion.
- Wrap-up scientific manuscripts plus Overview of grant writing:
    - Submission and authorship for scientific manuscripts.
    - Overview of the NIH grants writing process.
    -

3. Communicating effectively with the media and lay public and peer review 5

- How to write articles for the lay public. How to deal with the media. Ethics. Also: how to write a peer review.
- In-Class Exercise: Practice writing a peer review
- SUBMIT completed (or near-completed) manuscript

References:

1. Sin and Syntax, Constance Hale
2. Successful Scientific Writing: A step-by-step guide for biomedical scientists, Matthews and Bowen

MBI 302: Chemoinformatics and Drug Designing **Total 45L**

Advance Course – Theory; 3 Credits

UNIT I

1. Introduction to Cheminformatics 12

- Introduction to cheminformatics: aims, scope, role of cheminformatics in pharmaceutical/chemical research.
- Chemical Structure representation file format: 1D, 2D and 3D structures, molecular file formats (SMILES, PDB, SDF, MOL).
- Molecular Descriptors: 1D, 2D, 3D, topological (Molecular connectivity, electrotoplogical and refractotopological) and shape indices descriptors.
- Introduction to graph theory, vertex partitioning algorithms, Molecular Similarity and Molecular Diversity Analysis.

UNIT II

2. Introduction to Molecular Properties 11

- Similarity metrics: Tanimoto Coefficient, Euclidean distance and Tversky Index.
- Chemical Databases: Design, storage and retrieval methods, molecular database screening, Lipinski rule, Drug/Lead like molecules, chemical structure based search techniques
- Introduction to molecular properties: Hydrophobicity, molecular refractivity, electronic charges (Huckel charge, Wang-ford charge), principal moment of inertia.
- Pharmacokinetics study: *In-silico* ADMET Studies

UNIT III

3. Drug discovery and QSAR 11

- Drug discovery process, role of Bioinformatics in drug design.
- Target identification and validation, lead optimization and validation.
- Structure-based drug design and ligand based drug design.

- Concept of *de novo* design for lead identification.
- Structure Activity Relationship: QSARs and QSPRs, QSAR Methodology, Training data, test data and external validation data, applicability domain in QSAR
- 3D QSAR techniques: CoMFA and CoMSIA.

UNIT IV

4. Pharmacophore Modeling 11

- Pharmacophore features, Pharmacophore model, Receptor-based and ligand-based pharmacophore modeling.
- Virtual screening based on pharmacophore model.
- Receptor site, molecular docking study, flexible docking, rigid docking, molecular interactions.
- Scoring functions, correlation between ligand-based and receptor-based studies.

References:

1. Chemoinformatics -A Textbook by Johann Gasteiger and Thomas Engel, Wiley-VCH 2003.
2. Handbook of Chemoinformatics by Johann Gasteiger, Wiley-VCH 2003.
3. An Introduction to Chemoinformatics by Andrew R. Leach, Valerie J. Gillet, Springer 2007.
4. Burger’s Medicinal Chemistry edited by Donald J. Abraham, Volume – I.
5. Computer-Aided Molecular Design: Theory and Applications by Jean-Pierre Doucet, J. P. Doucet, Jacques Weber, Elsevier Science & Technology Books.
6. Receptor-based Drug Design edited by Paul Leff, Marcel Dekker Inc., New York.
7. Advanced Drug Design and Development: A Medicinal Chemistry Approach by P. N. Kourounakis, 1994, Taylor & Francis.
8. Biopharmaceutical Drug Design and Development by Susanna Wu-Pong, Yon Rojanasakul, 2008, Humana Press.
9. Combinatorial Library Design and Evaluation: Principles, Software, Tools, and Applications in Drug Discovery by Arup Ghose, Vellerkad Viswanadhan, 2001.
10. Computer-Aided Drug Design and Delivery Systems by Ahindra Nag, Baishakhi De, 2010, McGraw-Hill Professional.

MBI 303: Machine Learning Techniques **Total**

Advance Course – Theory; 3 Credits **45L**

UNIT I

1. Introduction to Machine Learning (ML) **2**

- .Introduction to Artificial Intelligence & Machine Learning,
- Elements of ML,
- Life Cycle of ML,
- Applications of Machine Learning.

2. Introduction to Data Mining **3**

- Data Science and Visualization
- Important Python Libraries
- pandas
- numpy
- scikit-learn
- matplotlib
- seaborn.

3. Biological Data and its Application in Machine Learning **1**

4. Genetic Algorithm **1**

- What is genetic algorithm?
- Types of genetic algorithm
- Introduction to feature selection

5.
- Introduction to Linear Regression **3**
- loss functions
- over fitting
- gradient descent

UNIT II

6. Classification Algorithm **6**

- Logistic Regression,
- Support Vector Machine
- Naïve Bayes Classifier.

7. Non Parametric Learning: **4**

- K- Nearest Neighbors
- Decision Trees
- Random Forest Classifier

8. Introduction to Cross-Validation, Hyper parameter tuning and ensemble Models **2**

UNIT III

7. Introduction & Types of Unsupervised Learning **1**

8. Association Rule Mining : (Market Basket Analysis) 2

9. Clustering : **3**

- K-means Clustering
- Hierarchical Clustering

10.

- Principal Component Analysis (PCA) **3**
- Singular Value Decomposition (SVD)
- SOM(self-organizing maps)

UNIT IV

8. Neural Computation **2**

- Introduction to Neural computation
- Neural network model

9. Evolutionary computation 4

- Introduction to evolutionary Processes

- Genetic Operators
- Evolutionary Optimization Algorithms

10.	Fuzzy Computation	**3**

- Introduction
- Fuzzy sets & systems
- Ant colony Optimization (ACO)

11.	Introduction to Human Language Technology & Machine Learning	**3**

- Stochastic Grammar & linguistics,Language models
- Statistical language Model,
- Markov model
- Hidden Markov Model

12.	Optimization Techniques	**2**

- Conjugate Gradient,
- Newton Raphson,
- Steepest descent,
- Simulated annealing

References:

1. Data Mining: Concepts and Techniques by Han and Kamber, Morgan Kaufmann.
2. Machine Learning by Tom Mitchell, McGraw Hill.
3. Data Mining: Practical Machine Learning Tools and Techniques by Witten and Frank, Elsevier.
4. Biological Sequence Analysis: probabilistic models of proteins and nucleic acids by Durbin, R., Eddy, S., Krogh, A. & Mitchison, G.Cambridge Univ. Press, 1998.
5. Optimization Theory and Application by Rao, S.S., 1984.
6. Discrete optimization by Parker, R. G. & Rardin, R. L., 1988.
7. Stochastic simulation by Repley, Brian D, Wiley series, 1987.
8. Methods of microarray data analysis III by Johnson, K.F. & Lin, S.M. Boston.

Kluwer academic publishers, 2003.

9. Exploration and analysis of DNA microarray and protein array data by Amaratunga, D. & Cabrera, J.New Jersey. John Wiley & Sons Inc., 90 BVDU-RGITBT-M.Sc. BIOINFORMATICS- SEM III 2004.
10. Ant colony optimization by Dorigo, Marco & Stutzle, Thomas New Delhi, Prentice-Hall of India Pvt Ltd, 2004.
11. Data mining: introductory and advanced topics by Dunham, M.H.:New Delhi, Pearson Education, 2003.
12. An introduction to bioinformatics algorithms by Jones, Neil.C. & Pevzner, Pavel A. New Delhi, Anne Books, 2005.
13. Fuzzy sets and fuzzy logic: theory and applications by Klir, G.J. & Yuan Bo, New Delhi. Printice Hall of India, 2002. 81-203-1136-1.
14. Fuzzy and neuro-fuzzy systems in medicine by Teodorrescu, HoriaNicolai, Kandel, Abraham. & Jain, Lakhmi, C New York, CRS Press

MBI 304: Current Bioinformatics	**Total**
Advance Course – Theory; 2 Credits	**30L**

UNIT I

1.	**Linux:** Linux Essentials, Hands-on Exercises Querying Biological Databases with SQL	**3**
2.	Statistics and graphing software: GraphPad Prism Creating Phylogenetic Trees with MEGA	**3**
3.	**NGS:** ➢ QC'ing Reads ➢ Mapping Next Generation Sequence Reads ➢ Analysis of ChIP-seq data in Galaxy ➢ Analysis of RNA-seq data in Galaxy ➢ Analysis of differential gene expression ➢ Finding and annotating indels in Human genome ➢ Bedtool: Toolset for genome arithmetic	**4**
5.	**Structural Variant Analysis** ➢ Variant detection: VarScan and SVDetect ➢ Variant annotation: TASSEL, GATK	**5**

UNIT II

6.	**Advanced Genomics** Visualizing Genomes: Browsers Introduction to ENSEMBL Integrative Genomics Viewer (IGV) Juggling Genome Coordinates Gene list enrichment analysis Visualizing Genomes: Circos Plots	**5**
7.	**Introduction to Microarray** DNA and Protein Array Microarray data Analysis pipeline	**6**

Microarray Techniques

Data analysis with R/Bioconductor

Clustering and Displaying Microarray Data

Expression Atlas

Cool BaRC Web Tools

8. Sequence analysis using DotPlot **4**

GeneGo: Bioinformatics Technology for Systems Biology

Visualizing Networks: Cytoscape

References:

1. Next-Generation DNA Sequencing Informatics,Stuart M. Brown, New York University School of Medicine *2013 ISBN 978-1-936113-87-3*

2. Chip Technology, Volume 77,Jörg D. Hoheisel, Alvis Brazma, illustratedPublisherSpringer, 2002, ISBN3540432159, 9783540432159

3. Microarray biochip technology, Mark Schena , illustratedPublisherEaton Pub., 2000, Original fromthe University of California, ISBN1881299376, 9781881299370

4. Discovering Statistics Using R, Andy Field, Jeremy Miles, Zoë Field, SAGE, 2012, ISBN1446258467, 9781446258460

5. The R Book, Michael J. Crawley, John Wiley & Sons, 2012ISBN1118448944, 9781118448946

6. A Guide to MATLAB: For Beginners And Experienced Users, Brian R. Hunt, Ronald L. Lipsman, Jonathan M. Rosenberg, Edition2, illustratedPublisherCambridge University Press, 2006, ISBN1139452533, 9781139452533

7. MATLAB Guide, Desmond J. Higham, Nicholas J. Higham, Edition2, illustratedPublisherSIAM, 2005ISBN0898715784, 9780898715781

8. MATLAB Primer, , Kermit Sigmon, Timothy A. Davis, Edition7, CRC Press, 2004ISBN1420034952, 9781420034950

9. Foundations of Comparative Genomics, Arcady R. Mushegian, Academic Press, 2010ISBN0080546099, 9780080546094

10. Essentials of Genomic and Personalized Medicine, Geoffrey S. Ginsburg, Huntington F Willard, Academic Press, 2009ISBN0080958117

MBI 305: Python Programming	**Total**
Advance Course – Theory; 2 Credits	**30L**

UNIT I

1.	**Introduction and Overview of Python**	4

- Installation and understanding Python Environment
- Introduction to Python variables
- basic Operators, keywords, python blocks

2.	**Data types and Program Flow Structures**	**6**

- Numeric data types: int, float, complex, string data type and string operations,
- Control Structure: if, else and elif , nested control structures
- Python Loop : for loops in python using ranges, string, list and dictionaries, while loops in python, Controlling loop using pass, continue, break and else

3.	**String , Lists, Tuples & Dictionaries**	**5**

- String: String operators, String manipulation using in built methods
- Lists: list operations, slicing, list methods, - list comprehension.
- Tuples: definition, assignment, tuple operations.
- Dictionaries: definition, operations and methods

UNIT II

4.	**Functions , Modules & Packages**	**4**

- Function: Definition, Function calling, lambda function
- Modules: Introduction to inbuilt Modules, User defined Modules

- Package: Understanding in built package, building own package, importing package

5. File Handling & Exceptional Handling 6

- File Handling: Opening a file, Read and Write operations, File and Directory related methods
- Exceptional Handling: try-except-else, Assert, try-finally, Raise, Exception Names & descriptions, User defined exceptions.

6. Python Database Interaction 5

- Python Database interaction, Insertion and retrieval information in database.
- Reading and storing information on database.

References:

1. Python: The Complete Reference Paperback – 20 Mar 2018, Martin C. Brown
2. Python Programming: A modular approach by Pearson Paperback – 26 Sep 2017, **by** Taneja Sheetal , Kumar Naveen
3. Python Machine Learning By Example Paperback – Import, 31 May 2017
4. by Yuxi (Hayden) Liu
5. Artificial Intelligence with Python Paperback – Import, 27 Jan 2017, **by** Prateek Joshi
6. Python Deep Learning Paperback – Import, 28 Apr 2017, **by** Valentino Zocca , Gianmario Spacagna , Daniel Slater, Peter Roelants

MBI 306: Introduction to Clinical Trials and Pharmacovigilance **Total**
Elective Course – Theory; 2 Credits **30L**

UNIT I

1. **Clinical Research:** **3**
   Introduction to clinical Pharmacology, basic components and scope. The concept of safety: measuring risk, safety in practice, risk–benefit balance, lack of benefit causation: causality assessment in individual cases, assessing causality from clinical trial data, using the Bradford hill criteria to assess causality.
2. Introduction to basic terms: tolerance, addiction, idiosyncrasy, allergy, hypersensitivity, antagonism, synergism, potentiation, tachyphylaxis, adverse drug reactions and its monitoring. **2**
3. Clinical evaluation of new drugs, organizations, types of clinical research **2**
4. Phases of clinical research: pre-clinical studies, human volunteer studies (Phase I), clinical trials (Phase II and III studies), post-marketing surveillance (Phase IV studies), prescription-event monitoring and registries. **2**
5. **Introduction to clinical trials:** **2**
   History, terminologies, type of clinical trials, role of clinical trials in new drug developments.
6. Regulatory affairs in clinical trials: **2**
   IND, NDA, ANDA-parts and contents, safety monitoring boards, FDA in various countries including India.
7. Clinical trial designs: **2**
   Designs used in clinical trials, inclusion and exclusion criteria, randomization blinding and controls.

UNIT II

8.	**Concept of Pharmacovigilance** Standard terms and terminologies in Pharmacovigilance.	2
9.	Adverse Drug Reactions: nature and mechanisms of ADRs, predisposing factors for ADRs.	2
10.	The process of Pharmacovigilance: signal detection, evaluation and investigation and communication.	2
11.	Regulatory aspects of Pharmacovigilance: legislation and guidelines, key elements of European legislation, regulatory Pharmacovigilance systems, obligations of pharmaceutical companies.	2
12.	International collaboration: international regulatory collaboration, World Health Organization, Council for International Organizations of Medical Sciences.	2
13.	Clinical aspects of adverse drug reactions: clinical burden of ADRs, important ADRs and minimizing risk, improving patient safety in clinical practice.	2
14.	Ethical and societal considerations: ethical principles, informed consent, privacy and confidentiality, ethical issues for the pharmaceutical industry, conflicts of interest.	2
15.	Reporting Database, managements and Risk assessment: PSUR (periodic safety update reports), risk assessments and managements. Regulatory guidelines and laws n PV.	1

References:

1. An Introduction to Pharmacovigilance (2nd edition) by Patrick Waller, Mira Harrison-Woolrych, Wiley-Blackwell, 2017
2. Pharmacovigilance- An Industry Perspective by By Deepa Arora, Pharmapublisher, 2012.
3. Pharmacovigilance: A Practical Approach edited by Thao Doan, Fabio Lievano, Mondira Bhattacharya, Linda Scarazzini, Cheryl Renz, Elsevier, 2018.

4. Dictionary of Pharmacovigilance by Amer Alghabban, Pharmaceutical Press, 2004.
5. Textbook of Pharmacovigilance by Gupta Sk, Jaypee Brothers, Medical Publishers, 2011.
6. Fundamentals of Clinical Trials, Textbook by David L. DeMets and Lawrence M. Friedman, 1981, Springer.
7. Design and analysis of clinical trials, Book by Shein-Chung Chow, Wiley, 1995.
8. Clinical Trials: A Practical Approach, Book by Stuart J. Pocock, 1983, Wiley.
9. The Sourcebook for Clinical Research: A Practical Guide for Study Conduct

   Book by Jeff Nelligan and Natasha Martien, 2018, Academic Press.

MBI 306: Selenium **Total**
Elective Course – Theory; 2 Credits **30L**

UNIT I

1. **Selenium Overview:** **6**
   - Selenium Overview & Selenium IDE basics
   - Introduction to Selenium
   - Overview of Selenium tool suite-IDE, Selenium RC, Web Driver, grid.
2. **Selenium IDE basics:** Selenium IDE basics **4**
   - Installation
   - IDE Features
   - Recording, running, Pausing and Debugging script
   - Verifying Page Elements
   - Assertion and Verification
   - Locating Elements- use of firebug, Firepath
     - What is firebug?
     - Installing Firebug
     - Firepath
3. **Selenium Web Driver I:** 5
   - Architecture of selenium web driver
   - web driver basis – Installation
   - web driver basis – Download/Setup/Configuration
   - web driver basis – writing the first test case
   - working with multiple browser- IE, Chrome, Firefox
   - working with Firefox profile – create Firefox profile and use in Web Driver script
4. **Introduction to TestNG :** 5
   - Introduction to TestNG’s role in testing
   - Why TestNG?
   - Testing with TestNG – Advantages

- Testing with TestNG – Disadvantages
- Features of TestNG
- Downloading TestNG
- Installing TestNG onto Eclipse
- Different basic types of annotations in TestNG

  @Test

  @Before method

  @After method

  @ Before class

  @After class

  Generation of test execution report using TestNG.

UNIT II

5. Managing UI Objects : 5

- Locating the element use of firebug, Firepath
- Inspect HTML with Firebug
- How to use Firepath?
- Different types of web elements.
- Identifying Web-elements using id, name, linkText, class xpath, tagname, cssselector
- Handling input box/buttons, list/selection boxes, radio buttons, check boxes, Grids/Web Tables
- Handling dynamic objects using regular expressions
- Extracting links and other Web-Elements.

6. Understanding XPath and CSS **2**

- Locating objects using regular expression
- Extracting data from web-Elements

7. Selenium Web Driver II : 3

- Switching between the frames, window
- Tabbed browsing with Selenium
- Managing Alerts/pop-ups

- Capturing screenshots

References:

1. Selenium Testing Tools Cookbook, Book by Unmesh Gundecha, November 2012, Unmesh Gundecha, Packt Publishing Ltd.
2. Instant Selenium Testing Tools Starter Book by Unmesh Gundecha, 2013, Unmesh Gundecha, Packt Publishing Ltd.
3. Front End Testing and Continuous Integration, Book by Daniel Ellis, Daniel Ellis, O’Relly publishing
4. Selenium 1.0 Testing Tools Beginner's Guide Book by David Burns, 2010, Packt Publishing Ltd.

MBI 307: Cancer Genomics **Total**
Elective Course – Theory; 2 Credits **30L**

UNIT I

1.
- HUMAN TUMORS -- epidemiology, classification, conventional treatment **7**
- Clonal selection in human tumors. Metastasis as a product of cell evolution.
- Tumor initiation. Chemical, Viral and other cancerogenic routes

2.
- Tumor suppressor genes and oncogenes. **8**
- Mutational process in tumors. Instability of tumor genome
- A review of the cancer gene cloning strategies in pre-genomic and post-genomic eras
- Genomic analysis of familial cancers
- Review of cancer and genomics

UNIT II

3.
- Signaling pathways damaged or short-circuited in human tumors: RB genes network, cyclines, CDKs, CDKIs **15**
- Signaling pathways damaged or short-circuited in human tumors: Receptors and RTKs, SMADs, RAS-cascade, PTEN, NF1 etc
- Extracellular matrix signaling, hypoxia, angiogenesis-related pathways and VHL
- Differentiation-related pathways in leukemia and lymphomas
- Cancer checkpoints
- P53 as guardian of genome
- Programmed cell death.
- Cancer- associated polymorphisms, Epigenetics
- Breakthrough Research and Therapeutics

References:

1. Stewart, B. W. K. P., and Christopher P. Wild. "World cancer report 2014." (2014).
2. Neidle, Stephen, ed. Cancer drug design and discovery. Elsevier, 2011.
3. Dellaire, Graham, Jason N. Berman, and Robert J. Arceci, eds. *Cancer* genomics: from bench to personalized medicine. Academic Press, 2013.
4. Cancer Genomics and Proteomics: Methods and Protocols (Methods in Molecular Biology) **by** Paul B. Fisher , Humana Press; 2007 edition
5. Cancer Genomics, Molecular Classification, Prognosis and Response Prediction, Pfeffer**,** Ulrich, Springer 2013
6. The Biology of Cancer, Robert Weinberg, WW Norton & Co, 2014

MBI 307: Biodiversity Informatics & Molecular Phylogenetics **Total 30L**

Elective Course – Theory; 2 Credits

UNIT I

1. Molecular Evolution 6

- Overview & protocols
- Nature of data used
- Basic Principles of Taxonomy & Phylogeny
- Modern Taxonomical Methods
- Molecular Systematics: Molecular data types, generation and analysis
- Diversity of Life: Genetic, Species, Community and Landscape Diversity
- Species Diversity: Measurement methodologies, Indices
- Trends of Species Diversity: Gradients, Biodiversity Hot Spots

2. Ecological / Ecosystem diversity 2

- Urban Biodiversity
- National, Regional and Global Diversity Information Systems and Networks

3. Biodiversity Informatics Databases 7

- Species 2000
- Tree of Life
- ATCC
- GBIF
- ICTV
- Species Analyst Collaboration
- Animal Virus information System
- Online Biodiversity and Ecosystem based Databases

UNIT II

4.	Probabilistic models and associated algorithms	**4**
	➢ Probabilistic models of evolution ➢ Maximum likelihood algorithm	
5.	Phylogenetic analysis algorithms	**6**
	➢ Distance-based: UPGMA, Transformed Distance, Neighbors-Relation, Neighbor-Joining Character optimization; delayed and accelerated transformation ➢ Maximum Parsimony Reliability of trees: Bootstrap, jackknife, decay, randomization tests.	
6.	Phylogenetic trees and their comparison	**5**
	➢ Definition and description, various types of trees; ➢ Consensus (strict, semi-strict, Adams, majority rule, Nelson). ➢ Data partitioning and combination. ➢ Tree to tree distances, similarity	

References:

1. Recknagel, F. 2002 Ecological Informatics: Understanding Ecology by Biologically- Inspired Computation. Springer, New York.
2. Phillipson, J. 1972 Ecological Energetics, Edward Arnold.
3. Odum, E.P. 1983 Basic Ecology. Saunders International Edition, Japan.
4. Atkinson, P.M. and Tate, N.J.(Eds.) 1999 Advances in remote sensing and GIS analysis., Wiley, New York.
5. Gunther,O. 1998 Environmental Information Systems. Berlin, New York, Springer.
6. Pankhurst, R.J. 1981 Practical taxonomic computing. Cambridge University Press, Cambridge, U.K.
7. Judd, W.S., Campbell, C.S., Kellogg, E.A.,Stevens, P.F. and Donoghue, M.J.(2002) Plant Systematics: A Phylogenetic Approach, 2nd Ed., Sinauer Associates, Inc. Publishers, Sunderland, Masssachusetts, U.S.A.

8. Nordenstam, B., El Gazaly, G. and Kassas, M. 2000 Plant Systematics for 21st Century. Portland Press Ltd.,London.

9. Bromham Lindell. Reading the Story in DNA: A Beginner's Guide to Molecular Evolution. Publisher: USA, Oxford University Press. 2008. ISBN: 9780199290918.

10. Bernardi Giorgio. Structural and Evolutionary Genomics, Volume 37: Natural Selection in Genome Evolution (New Comprehensive Biochemistry). Publisher: Netherlands, Elsevier Science.2005. ISBN: 9780444521361.

11. Marco Salemi, Anne-Mieke Vandamme. The phylogenetic handbook: a practical approach to DNA and protein phylogeny. Publisher: Cambridge University Press, 2003. ISBN: 052180390X.

12. Patthy Laszlo. Protein Evolution. Publisher: London, Blackwell Science Ltd. 1999. ISBN: 0632047747.

13. Takahata Naoyuki, Clark Andrew G.(Editor). Mechanisms of Molecular Evolution: Introduction to Molecular Paleopopulation Biology. Publisher: Japan, Japan Scientific Socities Press and Sinauer Associates, Inc. 1993. ISBN: 476226718X.

MBI 308: System Biology **Total 30L**

Elective Course – Theory; 2 Credits

UNIT I

1. Biological Systems – their analysis and modeling **8**

- Systems and modeling, modeling types, model constraints, process of modeling, constraints and analysis categories
- Qualitative modeling, principles, types, Forrester diagrams, formulations, simplification, Qualitative modeling using dynamic editors
- Quantitative modeling, categories, formulation, physical process modeling

2. Models of mathematics in Systems Biology **7**

- Difference equation based modeling, linear models, non-linear models, discrete models, continuous models, logistic models
- Interaction modeling, predator-prey models, stability, linearization, equilibria

UNIT II

3. Numerical techniques in analyzing bio-models **10**

- Numerical integration
- Numerical differentiation, ODEs, PDEs
- Adaptive modeling using ODE tools
- Numerical methods implementation

4. Validation techniques **5**

- Model validation, techniques, and discrimination models
- Dynamic validations
- System-level validation

References:

1. Elizabeth S. Allman, John A. Rhodes, Mathematical models in Biology: An Introduction, Cambridge University Press.
2. Bernhard O. Palsson, Systems Biology: Properties of Reconstructed Networks, Cambridge University Press.
3. James W. Haefner, Modeling Biological Systems: Principles and Applications Springer.
4. Zoltan Szallasi, Jörg Stelling, Vipul Periwal, System Modeling in Cell Biology: From Concepts to Nuts and Bolts, MIT Press.

MBI 308: Artificial Intelligence **Total**

Elective Course – Theory; 2 Credits **30L**

UNIT I

1. **Intelligent Agents:** **4**
    - PEAS Representation for an Agent,
    - Agent Environments, Concept of Rational Agent, Structure of Intelligent agents, Types of Agents
2. **Problem Solving**: **4**
    - Solving problems by searching, - DFS, BFS, A* searches, Performance Evaluation.
    - Use of graphs in bioinformatics
3. **Knowledge and Reasoning**: **4**
    - A knowledge Based Agent,
    - Introduction To Logic,
    - First Order Logic: Syntax and Semantics,
    - Inference in First Order Logic, Unification, Forward and backward chaining, Resolution
4. **Planning:** **3**
    - Planning problem,
    - Planning with State Space Search
    - Partial Order Planning
    - Hierarchical Planning
    - Conditional Planning

UNIT II

5. **Genetic Algorithms:** **2**
    - Single-objective genetic algorithms,
    - Multi-objective genetic algorithms,

- GA – bioinformatics applications genetic programming

6. Learning: **3**

- Introduction to neural networks
- Perceptrons
- Multilayer feed forward network
- Application of ANN
- Reinforcement learning: Passive & Active Reinforcement learning

7. Introduction to Robotics: **2**

- Robot Classification, Robot Specification.

8. Direct and Inverse Kinematics: **4**

- Coordinate Frames, Rotations, Homogeneous Coordinates,
- Arm Equation of four axis SCARA Robot,
- TCV
- Direct Kinematics of Four Axis SCARA Robot
- Inverse Kinematics of Four Axis SCARA Robot

9. Application: **4**

- DNA sequencing, hybridization
- RNA folding prediction using GA
- Genomics and proteomics
- MAGE (MicroArray *and* Gene Expression*)*
- Advances in tele surgery and surgical robotics
-

References:

1. Stuart Russell and Peter Norvig Artificial Intelligence A Modern Approach Edition, Pearson Publication.
2. Robert J. Schilling, Fundamentals of Robotics Analysis and Control, PHI Publication.
3. Intelligent Bioinformatics: The Application of Artificial Intelligence

Techniques to Bioinformatics Problems Edward Keedwell, Ajit Narayanan, John Wiley & Sons Ltd,

4. A Beginner's Guide to Microarrays: Eric M. Blalock, Springer
5. Artificial Intelligence and Bioinformatics group(AIBIG), Afnizanfaizal Abdullah Springer
6. Surgical Robotics Systems Applications and Visions : Rosen, Jacob; Hannaford, Blake; Satava, Richard M. (Eds.) Springer
7. Robotic Surgery: Farid Gharagozloo, Farzad Najam, McGraw-Hill Professional Publishing
8. Robin R Murphy Introduction to AI Robotics ISBN-81-203-2458-7 PHI Publication.

MBI 309: Chemoinformatics and Drug Designing Lab **Total**

(2 Credits, 1 P, 4.00 Hr., Advance Course Practical) **20 labs**

10

1. Cheminformatics:

- Practical session for chemical structure representation and storage in special file formats.
- Importance of 3D structures and method of generation from 1D & 2D representations.
- A brief introduction to building molecular databases with special emphasis on retrieval using structure input.
- Substructure/Exact/similar structure based searching.
- Development of descriptors (2D descriptors, Refracto-topological and Electro-topological state indices, 3D descriptors).

2. Drug Design: **10**

- Development of 2D QSAR models.
  - Classical QSAR
  - kNN QSAR
  - GFA QSAR
- Development of 3D QSAR models.
  - CoMFA
  - CoMSIA
  - Topomar CoMFA
  - HQSAR
- Web-based pharmacophore modeling.
  - HypoGen
  - HipHop
  - DiscoTech
- Molecular docking study.
  - Rigid Docking
  - Flexible Docking

  - Self Docking
- Virtual Screening.
- Combinatorial library design.

References:

1. Chemoinformatics -A Textbook by Johann Gasteiger and Thomas Engel, Wiley-VCH 2003.
2. Handbook of Chemoinformatics by Johann Gasteiger, Wiley-VCH 2003.
3. An Introduction to Chemoinformatics by Andrew R. Leach, Valerie J. Gillet, Springer 2007.
4. Burger's Medicinal Chemistry edited by Donald J. Abraham, Volume – I.
5. Computer-Aided Molecular Design: Theory and Applications by Jean-Pierre Doucet, J. P. Doucet, Jacques Weber, Elsevier Science & Technology Books.
6. Receptor-based Drug Design edited by Paul Leff, Marcel Dekker Inc., New York.
7. Advanced Drug Design and Development: A Medicinal Chemistry Approach by P. N. Kourounakis, 1994, Taylor & Francis.
8. Biopharmaceutical Drug Design and Development by Susanna Wu-Pong, Yon Rojanasakul, 2008, Humana Press.
9. Combinatorial Library Design and Evaluation: Principles, Software, Tools, and Applications in Drug Discovery by Arup Ghose, Vellerkad Viswanadhan, 2001.
10. Computer-Aided Drug Design and Delivery Systems by Ahindra Nag, Baishakhi De, 2010, McGraw-Hill Professional.

MBI 310: Machine Learning Technique Lab **Total**

(2 Credits, 1 P, 4.00 Hr., Advance Course Practical) **20P**

Implementation of following Algorithms using Python Programming

20

1. Logistic Regression,
2. Support Vector Machine,
3. Naïve Bayes Classifier. :
4. K- Nearest Neighbors,
5. Random Forest Classifier.
6. K-means Clustering
7. Association Rule Mining
8. Ant colony optimization
9. Simulated annealing algorithm for energy minimization
10. Estimator of transition probabilities for markov models based¬ on various sample sizes

References:

1. Data Mining: Concepts and Techniques by Han and Kamber, Morgan Kaufmann.
2. Machine Learning by Tom Mitchell, McGraw Hill.
3. Data Mining: Practical Machine Learning Tools and Techniques by Witten and Frank, Elsevier.
4. Biological Sequence Analysis: probabilistic models of proteins and nucleic acids by Durbin, R., Eddy, S., Krogh, A. & Mitchison, G.Cambridge Univ. Press, 1998.
5. Optimization Theory and Application by Rao, S.S., 1984.
6. Discrete optimization by Parker, R. G. & Rardin, R. L., 1988.
7. Stochastic simulation by Repley, Brian D, Wiley series, 1987.
8. Methods of microarray data analysis III by Johnson, K.F. & Lin, S.M. Boston. Kluwer academic publishers, 2003.
9. Exploration and analysis of DNA microarray and protein array data by Amaratunga, D. & Cabrera, J.New Jersey. John Wiley & Sons Inc., 90 BVDU-RGITBT-M.Sc. BIOINFORMATICS- SEM III 2004.
10. Ant colony optimization by Dorigo, Marco & Stutzle, Thomas New Delhi, Prentice-Hall of India Pvt Ltd, 2004.
11. Data mining: introductory and advanced topics by Dunham, M.H.:New Delhi, Pearson Education, 2003.
12. An introduction to bioinformatics algorithms by Jones, Neil.C. & Pevzner, Pavel A. New Delhi, Anne Books, 2005.
13. Fuzzy sets and fuzzy logic: theory and applications by Klir, G.J. & Yuan Bo, New Delhi. Printice Hall of India, 2002. 81-203-1136-1.

MBI 311: Current Bioinformatics Lab **(2 Credits, 1 P, 4.00 Hr., Advance Course Practical)**	**Total 20P**

1.	**Linux:** Linux Essentials, Hands-on Exercises Querying Biological Databases with SQL	**2**
2.	Statistics and graphing software: GraphPad Prism Creating Phylogenetic Trees with MEGA	**2**
3.	**NGS:** ➢ QC'ing Reads ➢ Mapping Next Generation Sequence Reads ➢ Analysis of ChIP-seq data in Galaxy ➢ Analysis of RNA-seq data in Galaxy ➢ Analysis of differential gene expression ➢ Finding and annotating indels in Human genome ➢ Bedtool: Toolset for genome arithmetic	**4**
5.	**Structural Variant Analysis** ➢ Variant detection: VarScan and SVDetect ➢ Variant annotation: TASSEL, GATK	**2**
6.	**Advanced Genomics** Visualizing Genomes: Browsers Introduction to ENSEMBL Integrative Genomics Viewer (IGV) Juggling Genome Coordinates Gene list enrichment analysis Visualizing Genomes: Circos Plots	**4**
7.	**Introduction to Microarray** DNA and Protein Array Microarray data Analysis pipeline Microarray Techniques Data analysis with R/Bioconductor Clustering and Displaying Microarray Data	**4**

	Expression Atlas	
	Cool BaRC Web Tools	
8.	Sequence analysis using DotPlot	**2**
	GeneGo: Bioinformatics Technology for Systems Biology	
	Visualizing Networks: Cytoscape	

References:

1. Next-Generation DNA Sequencing Informatics,Stuart M. Brown, New York University School of Medicine *2013 ISBN 978-1-936113-87-3*
2. Chip Technology, Volume 77,Jörg D. Hoheisel, Alvis Brazma, illustratedPublisherSpringer, 2002, ISBN3540432159, 9783540432159
3. Microarray biochip technology, Mark Schena , illustratedPublisherEaton Pub., 2000, Original fromthe University of California, ISBN1881299376, 9781881299370
4. Discovering Statistics Using R, Andy Field, Jeremy Miles, Zoë Field, SAGE, 2012, ISBN1446258467, 9781446258460
5. The R Book, Michael J. Crawley, John Wiley & Sons, 2012ISBN1118448944, 9781118448946
6. A Guide to MATLAB: For Beginners And Experienced Users, Brian R. Hunt, Ronald L. Lipsman, Jonathan M. Rosenberg, Edition2, illustratedPublisherCambridge University Press, 2006, ISBN1139452533, 9781139452533
7. MATLAB Guide, Desmond J. Higham, Nicholas J. Higham, Edition2, illustratedPublisherSIAM, 2005ISBN0898715784, 9780898715781
8. MATLAB Primer, , Kermit Sigmon, Timothy A. Davis, Edition7, CRC Press, 2004ISBN1420034952, 9781420034950
9. Foundations of Comparative Genomics, Arcady R. Mushegian, Academic Press, 2010ISBN0080546099, 9780080546094
10. Essentials of Genomic and Personalized Medicine, Geoffrey S. Ginsburg, Huntington F Willard, Academic Press, 2009ISBN0080958117, 9780080958118

MBI 312: Python Programming lab **Total**
(2 Credits, 1 P, 4.00 Hr., Core Course Practical) **20 labs**

1. Basic Python Programming **10**

- Python Basic
- Python Basic (Part -II)
- Python Data Types - String
- Python Data Types - List
- Python Data Types - Dictionary
- Python Data Types - Tuple
- Python Data Types - Sets
- Python Array
- Python Conditional statements and loops
- Python functions

2. Python Data Structures and Algorithms **3**

- Data Structure
- Search and Sorting
- Recursion

3. **7**

- Python Date Time
- Python Class
- Python Math
- Python File Input Output
- Python Regular Expression

References:

1. https://www.w3resource.com/python-exercises/
2. Python: The Complete Reference Paperback – 20 Mar 2018, Martin C. Brown
3. Python Programming: A modular approach by Pearson Paperback – 26 Sep 2017, **by** Taneja Sheetal , Kumar Naveen
4. Python Machine Learning By Example Paperback – Import, 31 May 2017
5. by Yuxi (Hayden) Liu
6. Artificial Intelligence with Python Paperback – Import, 27 Jan 2017, **by** Prateek Joshi
7. Python Deep Learning Paperback – Import, 28 Apr 2017, **by** Valentino Zocca , Gianmario Spacagna , Daniel Slater, Peter Roelants

MBI 313: Introduction to Clinical Trials and Pharmacovigilance Lab

(2 Credits, 1 P, 4.00 Hr., Elective Course Practical)

		Total 20 labs
1.	➢ Study of case papers to study drug interactions. ➢ Studying various Bioinformatics tools to study drug interactions. ➢ Reporting Database	**10**
2.	➢ Clinical trial phase reports	**10**

References:

1. An Introduction to Pharmacovigilance (2nd edition) by Patrick Waller, Mira Harrison-Woolrych, Wiley-Blackwell, 2017
2. Pharmacovigilance- An Industry Perspective by By Deepa Arora, Pharmapublisher, 2012.
3. Pharmacovigilance: A Practical Approach edited by Thao Doan, Fabio Lievano, Mondira Bhattacharya, Linda Scarazzini, Cheryl Renz, Elsevier, 2018.
4. Dictionary of Pharmacovigilance by Amer Alghabban, Pharmaceutical Press, 2004.
5. Textbook of Pharmacovigilance by Gupta Sk, Jaypee Brothers, Medical Publishers, 2011.
6. Fundamentals of Clinical Trials, Textbook by David L. DeMets and Lawrence M. Friedman, 1981, Springer.
7. Design and analysis of clinical trials, Book by Shein-Chung Chow, Wiley, 1995.
8. Clinical Trials: A Practical Approach, Book by Stuart J. Pocock, 1983, Wiley.
9. The Source book for Clinical Research: A Practical Guide for Study Conduct Book by Jeff Nelligan and Natasha Martien, 2018, Academic Press.

MBI 313: Selenium Lab **(2 Credits, 1 P, 4.00 Hr., Elective Course Practical)**		**Total** **17 labs**
1.	Selenium Overview	**2**
2.	Selenium IDE basics	**3**
3.	Selenium Web Driver I	**3**
4.	Introduction to Test NG	**3**
5.	Managing UI Objects	**3**
6.	Selenium Web Driver II	**3**

References:

1. An Introduction to Pharmacovigilance (2nd edition) by Patrick Waller, Mira Harrison-Woolrych, Wiley-Blackwell, 2017
2. Pharmacovigilance- An Industry Perspective by By Deepa Arora, Pharmapublisher, 2012.
3. Pharmacovigilance: A Practical Approach edited by Thao Doan, Fabio Lievano, Mondira Bhattacharya, Linda Scarazzini, Cheryl Renz, Elsevier, 2018.
4. Dictionary of Pharmacovigilance by Amer Alghabban, Pharmaceutical Press, 2004.
5. Textbook of Pharmacovigilance by Gupta Sk, Jaypee Brothers, Medical Publishers, 2011.
6. Fundamentals of Clinical Trials, Textbook by David L. DeMets and Lawrence M. Friedman, 1981, Springer.
7. Design and analysis of clinical trials, Book by Shein-Chung Chow, Wiley, 1995.
8. Clinical Trials: A Practical Approach, Book by Stuart J. Pocock, 1983, Wiley.
9. The Source book for Clinical Research: A Practical Guide for Study Conduct Book by Jeff Nelligan and Natasha Martien, 2018, Academic Press.

MBI 314: Cancer Genomics Lab **(2 Credits, 1 P, 4.00 Hr., Elective Course Practical)**	**Total** **20 labs**
1. TCGA data portal --> cBioPortal (use this as a database)	**3**
2. Data mining using SRA or GEO databases for rna seq data	**3**
3. Data clustering and analysis	**2**
4. Single cell transcriptomics	**3**
5. Data visualization using the following plots: tSNE, MDS, heatmap	**3**
6. Single sample gene set enrichment analysis to find gene panels	**3**
7. Survival analysis using kaplan meier plots	3

References:

1. Stewart, B. W. K. P., and Christopher P. Wild. "World cancer report 2014." (2014).
2. Neidle, Stephen, ed. Cancer drug design and discovery. Elsevier, 2011.
3. Dellaire, Graham, Jason N. Berman, and Robert J. Arceci, eds. Cancer genomics: from bench to personalized medicine. Academic Press, 2013.
4. Cancer Genomics and Proteomics: Methods and Protocols (Methods in Molecular Biology) **by** Paul B. Fisher , Humana Press; 2007 edition
5. Cancer Genomics, Molecular Classification, Prognosis and Response Prediction, Pfeffer**,** Ulrich, Springer 2013
6. The Biology of Cancer, Robert Weinberg, WW Norton & Co, 2014

MBI 314: Biodiversity Informatics & Molecular Phylogenetics Lab **Total**
(2 Credits, 1 P, 4.00 Hr., Elective Course Practical) **20 labs**

1. Biodiversity Informatics Lab **3**

Biodiversity Informatics Databases:

- Species 2000
- Tree of Life
- ATCC
- GBIF
- ICTV

2. **4**

- Online Biodiversity and Ecosystem based Databases
- Software for identification
- Accessing existing databases on the WWW
- Software for identification of species
- Probabilistic and deterministic identification
- Delta MicroIS, AVIS, ICTV

3. Molecular Phylogenetics Lab **1**

Compilation & curation of dataset, format conversion

4. Survey of software programs available for phylogenetic analysis **1**

- Installation of at least 2 public domain packages for both Windows & Unix environment: Phylip, PAUP, MEGA

5. MSA using ClustalW: writing files for phylogenetic analysis **2**

- Informative and variable sites; Singleton sites etc.
- Difference between dendogram & phylogenetic tree
- Converting sequence data into distance data

6. Reconstruction of phylogenetic tress using molecular data **2**

- Distance-based methods: UPGMA, Neighbor-joining, Neighbor-relations & Transformed distance
- Maximum Parsimony

	➢ Maximum likelihood	
7.	Using bootstrapping tool to generate multiple datasets from the original input data & generation of consensus tree	**1**
8.	Plotting, visualizing & printing phylogenetic trees: TreeView and other tools ➢ Various rendering ➢ Formatting & labeling ➢ Interpretation of trees	**2**
9.	Comparison of trees drawn using ➢ RNA, Nucleotide & protein data ➢ Gene tress & species tree	**1**
10.	Un-rooted & rooted tree ➢ Rooting un-rooted tree using an out group	**1**
11.	Reconstruction of phylogenetic trees using whole genome data of viruses	**1**
12.	Numerical taxonomy & Bacterial identification using matrices	**1**

References:

1. Recknagel, F. 2002 Ecological Informatics: Understanding Ecology by Biologically- Inspired Computation. Springer, New York.
2. Phillipson, J. 1972 Ecological Energetics, Edward Arnold.
3. Odum, E.P. 1983 Basic Ecology. Saunders International Edition, Japan.
4. Atkinson, P.M. and Tate, N.J.(Eds.) 1999 Advances in remote sensing and GIS analysis., Wiley, New York.
5. Gunther,O. 1998 Environmental Information Systems. Berlin, New York, Springer.
6. Pankhurst, R.J. 1981 Practical taxonomic computing. Cambridge University Press, Cambridge, U.K.
7. Judd, W.S., Campbell, C.S., Kellogg, E.A.,Stevens, P.F. and

Donoghue, M.J.(2002) Plant Systematics: A Phylogenetic Approach, 2nd Ed., Sinauer Associates, Inc. Publishers, Sunderland, Masssachusetts, U.S.A.

8. Nordenstam, B., El Gazaly, G. and Kassas, M. 2000 Plant Systematics for 21st Century. Portland Press Ltd.,London.
9. Bromham Lindell. Reading the Story in DNA: A Beginner's Guide to Molecular Evolution. Publisher: USA, Oxford University Press. 2008. ISBN: 9780199290918.
10. Bernardi Giorgio. Structural and Evolutionary Genomics, Volume 37: Natural Selection in Genome Evolution (New Comprehensive Biochemistry). Publisher: Netherlands, Elsevier Science.2005. ISBN: 9780444521361.
11. Marco Salemi, Anne-Mieke Vandamme. The phylogenetic handbook: a practical approach to DNA and protein phylogeny. Publisher: Cambridge University Press, 2003. ISBN: 052180390X.
12. Patthy Laszlo. Protein Evolution. Publisher: London, Blackwell Science Ltd. 1999. ISBN: 0632047747.
13. Takahata Naoyuki, Clark Andrew G.(Editor). Mechanisms of Molecular Evolution: Introduction to Molecular Paleopopulation Biology. Publisher: Japan, Japan Scientific Socities Press and Sinauer Associates, Inc. 1993. ISBN: 476226718X.

SEMESTER IV

MBI 401: RESEARCH PROJECT

Core Course – 20 Credits

Guidelines for Research Project Submission during Sem IV for Master of Science (M. Sc.) in Bioinformatics

Eligibility:

- If student fails to pay any of the pending dues before the beginning of the project date, he/ she will not be considered to be eligible to undertake research project.

Project Duration, Dissertation writing and Submission:

- The project duration will be from beginning till the end of Semester IV
- Students shall submit dissertation title, name of research guide, name of co-guide (for off-campus only), name & place of research work within 10 days after the start date of semester IV. (As per the format enclosed).
- Student must have to submit Two hard copies (copy of Library/ Co-Guide & Student's copy) and one soft copy (Guide) of Dissertation only in the prescribed format (read below), duly approved by Research Guide(s) on or before April 15th of every year. The dissertation will not be accepted for evaluation for those who submit after March 15th, their presentation will be taken along with the next year batch.
- Students must acknowledge all the figures, maps, tables, methods, texts, etc., that are used, taken from other sources for writing the dissertation, except for original work that they have carried out. Dissertation having more than 10 % of plagiarism found will not be considered for evaluation.
- Dissertation must be written in specified format only as mentioned below:
  a. The paper used for printing shall be of A4 size;

b. Printing shall be in a standardized form (word size of 12, font in Times New Roman) on one side of the paper and in 1.5 line spacing;

c. A margin of 1.5 inches shall be on the left hand side, top, bottom and right hand margin shall be of 1 inch.

d. The card for cover shall not be more than 330 GSM.

e. The title of the dissertation, name of the candidate, degree, name of the guide, co-guide, place of research and the date, month and year of submission shall be printed on the title page and on the front cover.

f. The hard- bound thesis cover shall be of black color. Spine of the binding [side cover] should mention 'M Sc. Bioinformatics dissertation on the top, name of the candidate and date, month and year.

- Student must follow following chapter scheme for Dissertation submission:

Chapter Scheme of Dissertation :

- ✓ Introduction
- ✓ Review of Literature
- ✓ Aims and Objectives
- ✓ Materials & Methods
- ✓ Observations and Results
- ✓ Discussion
- ✓ Summary
- ✓ Conclusions
- ✓ Bibliography –Reference etc.

Place of Research Project:

- Student may prefer to undertake his/ her research project in-house or off-campus. Students those preferring to pursue research at off-campus will have to undertake research work only in any of the Department of Scientific and Industrial research (DSIR), Government of India recognized laboratory (Government, State-Government, Private).

Research Guide(s):

- Students who are opting for off-campus will have one Major Research Guide from the host organization (Preferably a Ph D. qualified scientist), however, a Co-Guide from RGIITBT will be appointed.

Dissertation Evaluation:

- Students will have to submit **Two progress reports (45 days of Intervals) (Format Enclosed) and One evaluation report from Research Guide having 40 % weight age (Format Enclosed)** at the time of dissertation submission on the execution of research project duly signed by Guide / Co-Guide. The progress report will include, attendance percentage, review collection, research progress, sincerity, topic understanding, and systematic execution of research project, data collection and management. The evaluation report will include attendance, review work, project execution, critical thinking, originality of work, presentation of result, understanding of research, dissertation write-up, presentation of tables, figures, maps, references, etc.
- Student will have to give 20 min presentation on the work done in the presence of expert committee (between **April 21 to May 5 of every year**). (Note only working dates will be considered). The power point presentation format shall contain project title, name of candidate, place of research work, name of Guide/ Co-Guide, introduction, review, objectives, significance of the work, methodology, results & discussion, conclusion, references and acknowledgement. The presentation shall be of 12-15 min with 5-8 minutes of discussion.
- The presentation will carry **60 % of weight age** based on the following consideration. The overall understanding of the research project, objectives, methodology. The outcome of research work, data analysis and statistics, clarity in presentation and question – answer session (**Format Enclosed**).
- Student will be assigned a grade as per the Rules mentioned.

RAJIV GANDHI INSTITUTE OF IT AND BIOTECHNOLOGY

"Write here **approved title** of the Dissertation in all upper-case
(capital letters) with a 'centre' alignment. Place
this title on the upper central part of the cover
with sufficient margin from top and both sides.
Use font size suitable to length of the title"

A DISSERTATION SUBMITTED TO

RAJIV GANDHI INSTITUTE OF I.T. AND BIOTECHNOLOGY,
BHARATI VIDYAPEETH DEEMED UNIVERSITY, PUNE

FOR AWARD OF DEGREE OF
MASTER OF SCIENCE in BIOTECHNOLOGY

SUBMITTED BY
………………………………..

UNDER THE GUIDANCE OF
……………………………………..
…………………………………

Name of Co-Guide Name of Guide

RESEARCH CENTRE
……………………………………
…………………………………..
…………………………………..

WRITE HERE DATE, MONTH & YEAR OF SUBMISSION

CERTIFICATE

This is to certify that the work incorporated in the dissertation entitled "…………………………………………………………………………………..." for the degree of 'Master of Science' in the subject of Biotechnology under the faculty of Interdisciplinary Science has been carried out by Mr/ Mrs……………… Rajiv Gandhi Institute of I.T and Biotechnology, BharatiVidyapeeth Deemed University, Dhankawadi, Pune (OR NAME OF THE LABORATORY, PLACE OF THE WORK) during the period from ………… to…………under the guidance of Dr…………….

Place: Pune (Signature of Head of the Institute with seal)
Date : Principal / Director
Seal

CERTIFICATION OF GUIDE

This is to certify that the work incorporated in the dissertation entitled "……………………………………………………………………………………..…… ……"

Submitted by……………. for the degree of 'Master of Science' in the subject of 'Biotechnology' under the faculty of Interdisciplinary Science has been carried out in the Department (laboratory) of…………….., RGIITBT, BVDU (**Institute/ Private Lab, Govt Lab etc**) ……………, Pune (**OR Place**) during the period from …….to……, under my direct supervision/ guidance.

Place : (Signature of Research Guide)

Date : (Name & Designation)

Place : Pune (Signature of Research Co-Guide)

Date : (Name & Designation)

DECLARATION BY THE CANDIDATE

I hereby declare that the dissertation entitled "
" ____ submitted by me to the Bharati Vidyapeeth (Deemed to be University), Pune for the degree of Master of Science (M.Sc.) in Bioinformatics under the Faculty of Interdisciplinary Sciences
has been carried out by me under the supervision of **(Name of Guide)** and **(Name of Co-guide (if any).**

I further declare that it has not been submitted to this or any other university or Institution for the award of any Degree or Diploma.

I also confirm that all the material which I have borrowed from other sources and incorporated in this dissertation is duly acknowledged. If any material is not duly acknowledged and found incorporated in this dissertation, it is entirely my responsibility. I am fully aware of the implications of any such act which might have been committed by me advertently or inadvertently.

Place :
Date : / /

Name & signature of
Research Student

BHARATI VIDYAPEETH DEEMED UNIVERSITY
(Re-accredited with **A grade** by NAAC in 2011, Accredited with **A$^+$ Grade** by NAAC in 2017)

Rajiv Gandhi Institute of Information Technology and Biotechnology
Pune-Satara Road, Katraj, Pune – 411 046

Proforma of Progress Report – I

(To be submitted to Principal, RGIITBT, BVDU, Pune)

Name of the Student :
Registration Number of the Student :
Degree Program :
Project Title :
Name of the Research Guide :
Name of Internal Guide (Co-Guide) :
(Only in case of off-campus student)
Period under report : Dec 1 – Jan 15

Name & Place of Research Work :
Objectives of Research Work : 1)
2)
3)

(**Tick mark, wherever applicable**)

	Very Good	**Good**	**Poor**	**Special Remark, if any**
Percent Attendance				
Getting well acquainted with colleague and laboratory procedures, sincerity				
Technical Aspects: Understanding research topic, review collection, systematic execution of research project, research progress, data collection and management				
Overall Performance				

Name of Research Guide with signature and seal:

Place & Date:

BHARATI VIDYAPEETH DEEMED UNIVERSITY

(Re-accredited with **A grade** by NAAC in 2011, Accredited with **A$^+$ Grade** by NAAC in 2017)

Rajiv Gandhi Institute of Information Technology and Biotechnology
Pune-Satara Rd, Katraj, Pune – 411 046

Proforma of Progress Report – II

(To be submitted to Principal, RGIITBT, BVDU, Pune)

Name of the Student :
Registration Number of the Student :
Degree Program :
Project Title :
Name of the Research Guide :
Name of Internal Guide (Co-Guide) :
(Only in case of off-campus student)
Period under report : Jan 16 – Feb 28

Name & Place of Research Work :
Objectives of Research Work:
1)
2)
3)

(**Tick mark, wherever applicable**)

	Very Good	**Good**	**Poor**	**Special Remark, if any**
Percent Attendance				
Getting well acquainted with colleague and laboratory procedures, sincerity				
Technical Aspects: Understanding research topic, review collection, systematic execution of research project, research progress, data collection and management				
Overall Performance				

Name of Research Guide with signature and seal:

Place & Date:

BHARATI VIDYAPEETH DEEMED UNIVERSITY
(Re-accredited with **A grade** by NAAC in 2011, Accredited with **A$^+$ Grade** by NAAC in 2017)

Rajiv Gandhi Institute of Information Technology and Biotechnology
Pune-Satara Rd, Katraj, Pune – 411 046

Proforma of Internal Evaluation Report

(To be submitted to Principal, RGIITBT, BVDU, Pune)

Name of the Student :
Registration Number of the Student :
Degree Program :
Project Title :
Name of the Research Guide :
Name of Internal Guide (Co-Guide) :
(Only in case of off-campus student)

Name & Place of Research Work :
Completion of Research Objectives : (Yes / No)

(**of 40 % weight age**)

	Out of	**Marks obtained**
Percent Attendance	10	
Getting well acquainted with colleague and laboratory procedures, sincerity	5	
Technical Aspects: Understanding research topic, review collection, systematic execution of research project, research progress, data collection and management	20	
Overall Performance	5	
Total		

1. Name of Research Guide with signature and seal:
2. Name of Research Guide:

Place & Date:

BHARATI VIDYAPEETH DEEMED UNIVERSITY

(Re-accredited with A grade by NAAC in 2011, Accredited with A$^+$ Grade by NAAC in 2017)

Rajiv Gandhi Institute of Information Technology and Biotechnology
Pune-Satara Rd, Katraj, Pune – 411 046

Proforma of Evaluation Report on Presentation

(To be submitted to Principal, RGIITBT, BVDU, Pune)

Name of the Student :
Registration Number of the Student :
Degree Program :
Project Title :
Name of the Research Guide :
Name of Internal Guide (Co-Guide) :
(Only in case of off-campus student)

Name & Place of Research Work :
Completion of Research Objectives : (Yes / No)

(**of 60 % weight age**)

	Out of	**Marks obtained**
Overall understanding of the research project - Research Objectives	10	
Significance of Research / Review	10	
Results – Data presentation, statistical analysis, Softwares used, Result Interpretation	15	
Presentation – Clarity, power point slides, communication skills, question – answer session.	15	
Significant outcome – Technical abstract, Seminar, etc.	10	
Total		

Name & Signature of Expert Pane - 1. (External):

2. (Internal):

3. (Internal):

Signature of Principal (RGIITBT):

BHARATI VIDYAPEETH DEEMED UNIVERSITY

(Re-accredited with A grade by NAAC in 2011, Accredited with A^{+} Grade by NAAC in 2017)

Rajiv Gandhi Institute of Information Technology and Biotechnology
Pune-Satara Rd, Katraj, Pune – 411 046

Plan of Research Outline

(To be submitted to Principal, RGIITBT, BVDU, Pune)

Name of the Student :
Registration Number of the Student :
Degree Program :
Proposed Project Title :
Name of the Research Guide :
Name of Internal Guide (Co-Guide) :
(Only in case of off-campus student)

Name & Place of Research Work :
Proposed Research Objectives :

Start Date of Research project :

Likely Date of Project Completion :

Significance of Research Project :

Name & Sign of Student

Approved by

1. Name of Research Guide with signature and seal:

Place & Date:

2. Name of Research Guide (Co-Guide) with signature and seal:

Place & Date:

BHARATI VIDYAPEETH
(DEEMED TO BE UNIVERSITY), PUNE

FACULTY OF INTERDISCIPLINARY STUDIES
M.Sc. - BIOINFORMATICS
Old Syllabus

BHARATI VIDYAPEETH DEEMED UNIVERSITY, PUNE

Rajiv Gandhi Institute of IT & Biotechnology, Katraj, Pune- 411046

FACULTY OF INTERDISCIPLINARY STUDIES

MASTER OF SCIENCE IN BIOINFORMATICS PROGRAMME

STRUCTURE AND DETAILED SYLLABUS OF SEMESTERS I to IV
(UNDER CHOICE BASED CREDIT SYSTEM)
TO BE EFFECTIVE FROM 2013 -14

BHARATI VIDYAPEETH DEEMED UNIVERSITY, PUNE

Rajiv Gandhi Institute of IT & Biotechnology, Katraj, Pune- 46

FACULTY OF INTERDISCIPLINARY STUDIES

Course Structure of Master of Science in Bioinformatics Program

(UNDER CHOICE BASED CREDIT SYSTEM)
TO BE EFFECTIVE FROM 2013 -14 at Semester I

INTRODUCTION

The Master of Science (M.Sc.) program in Bioinformatics is a full time **100 credit program** offered by Bharati Vidyapeeth Deemed University (BVDU) in its constitutive unit Rajiv Gandhi Institute of IT and Biotechnology, Pune. The Institute has excellent infrastructure, faculty, state-of-art laboratories, and library to provide appropriate learning facilities and environment. The Institute is approved by UGC to conduct graduate and post graduate courses in Biotechnology & Bioinformatics and the University itself is **reaccredited by NAAC** and **accredited by Ministry of MHRD**, Government of India with **'A'** grade.

The expectations of the Bioinformatics industry and research are envisaged while designing the M.Sc. Program. The feedback of students, faculty, employers and parents has contributed in giving an industrial accord to this curriculum.

M. SC. BIOINFORMATICS (MBI) PROGRAM:

❖ 21st century biology has been transformed in to integrative biology due to paradigm shift from reductionist to holistic approach. The data explosion is one of the obvious causes of the transformation to holistic or systems biology mode. In order to have deeper insights in to biological systems, integration of the data from genome to phenome levels to generate dynamic models has surfaced as key area of bioinformatics.

- Integration of biological data from all levels requires expertise in tools and techniques of mathematics, computer science and information technology. This post graduate program imparts training in core techniques and formalisms of integrative bioinformatics and also provides required flexibility in selecting the project topic of choice at the each student level.

- The technical side of the M. Sc. Bioinformatics Program gives emphasis on imparting training in programming, algorithm design, data storage, mining protocols, advanced tools of integrating databases, systems biology and kinetic modeling with respect to developing skill set to efficiently address the needs of data analyses of integrative biology.

VISION STATEMENT

To generate professionally competent human resource in Bioinformatics to serve the academia, research, agriculture, health and IT industry.

MISSION STATEMENT

Social Transformation through Dynamic Education.

OBJECTIVES

Objectives of this course (M.Sc. Bioinformatics) essentially focus on the development of skills of students for a successful career in industry or research.

- The course structure emphasizes enough effort on theory as well as laboratory practical work so as to have fundamentals clear and strengthened through practical.
- It is essential for the students to read research papers and deliver seminars that would help them to know the recent advances in the subject and also develop the communication skills.

- The course includes project work that would develop the scientific approach and research attitude to plan and implement experiments, record observations, interpret them on the basis of existing knowledge and communicate through clear and concise English.
- The course emphasizes on the delivery of the state of the art technologies in Genomics, Proteomics and Drug discovery lab trainings to know how biological data is generated.
- The course has been designed to fulfill needs of the industry for the manpower with the specific skills sets in the area of bioinformatics.

ELIGIBILITY FOR ADMISSION TO THE COURSE

Admission to the course is open to any graduate of any recognized university satisfying the following conditions

- **i.** Graduate in any discipline of Science and Technology. (e.g. B.Sc. /BCS/B. Pharm. /MBBS /BDS / BHMS/BE etc.) with minimum **45% (40% for Reserve Category)** marks in aggregate.
- **ii.** Post graduate in any discipline of Science and Technology with minimum **45% (40% for Reserve Category)**.
- **iii.** Subject to the above conditions, the final admission is based solely on the merit at the all India entrance test (MBI) conducted by Bharati Vidyapeeth Deemed University.

DURATION OF THE COURSE

The duration of this course is two years (four semesters). The medium of instruction and examination will be only **English**.

RULES AND REGULATIONS

1. The entire Course is of 100 credits and will be executed in 4 semesters.
2. One credit for theory course is equivalent to 15 hrs of interaction of student with facilitator while one credit for practical course is equivalent to 45 hrs. of practical work.

3. A total credit for respective courses indicates number of hours of classroom teaching for respective theory courses per week. e.g. For every theory course of 4 credits, there will be 3L +1T per week.
4. Credits for practical courses indicate number of laboratory sessions per week (each session is of 3 hours). For every practical course of 2 credits, there will be 2 labs per week.
5. The number of credits offered in Semesters I, II, III and IV are 30, 30, 28 and 12 respectively.
6. The distribution of the credits as per level is 28 for Foundation, 24 for Core and 20 for Advanced level.
7. Out of 100 credits, 16 are Elective credits acquired by the students out of 32 total elective credits.
8. The student will undertake dissertation at the start of Semester IV for which 12 credits are allotted. 600 marks are allotted for the assessment of this dissertation work. The distribution of these marks is as follows;
    i. 200 Marks for updated literature survey and framing of Aims & Objectives of the selected dissertation topic.
    ii. 200 Marks for progress of work (based on day – to – day work)
    iii. 100 Marks for fulfillment of objectives
    iv. 100 Marks for presentation/ Viva (External evaluation)
9. The field/ industrial/ academic institute visits will be organized for demonstration of advance experimental techniques. The students are expected to submit a report on the basis of the observations carried out during the visit.
10. Student must have minimum 70% attendance. A student who does not have the required attendance may not be allowed to appear for the university term.

RULES FOR EXAMINATION

1. Examination for the theory and practical courses will be conducted at the end of each respective semester.
2. The weightage allotted for internal assessment is 40%. Thus 40 marks out of 100 will be allotted for internal assessment which will be conducted throughout the semester in the form of tests/tutorials/seminars/oral presentations etc.

3. The university shall conduct an examination for 60 marks.
4. The elective courses like Elective I, II, III, V and General Course MBI 301 will be evaluated by the Institute on the basis of continuous assessment. There will be no semester End Examination conducted by the University. But Elective IV, Elective Lab I and II will be evaluated by the University.

STANDARD OF PASSING

1. The 10 –point scale would be used to convert marks out of 100 to grades and grade points according to the following table.
2.

Marks as percentage	Grade	Grade point
[75, 100]	O	10.0
[70, 74.9]	A$^+$	9.0
[65, 69.9]	A	8.0
[60, 64.9]	B+	7.0
[55, 59.9]	B	6.0
[50, 54.9]	C+	5.5
[45, 49.9]	C	5.0
[40, 44.9]	D	4.5
[00, 39.9]	F	0.0

3. For courses which have University Examination (UE) and Internal Assessment (IA)

   In order to pass at the University Examination, a student must obtain at least a grade point of 4.5 at the university examination and minimum CGPI of 4.0 in aggregate of UE and IA. There will be a separate grade assigned for performance in IA. There is no minimum grade for passing in IA. The **Grade Point Index (GPI)** will be calculated for a course as and when the student passes in the UE by securing at least D grade. In such a case, the GPI is calculated by combining grade points at UE and IA with respective weights 0.6 and 0.4. A student passes in the course, must clear the course by re-appearing at UE as a backlog candidate.

Postgraduate CBCS Program in Bioinformatics (M.Sc. Bioinformatics)

4. For courses which have no University Examination (UE):

For some General courses and Elective courses, there is no UE. The institute conducts '**Continuous Assessment**' and report the grade. For such courses the corresponding grade point would be the GPI. In order to pass in such courses the student must obtain a minimum GPI of 4.0.

5. At the end of each semester, a **Cumulative Grade Point Index (CGPI)** will be calculated as a weighted average of the GPI of all courses in which the student have passed till that semester.

6. A student who passes in all the courses will be declared to have passed the entire M.Sc. program with the following honors.

CGPI in [0.00, 3.99] -- Fail
CGPI in [4.00, 4.99] -- Pass Class
CGPI in [5.00, 5.49] -- Second Class
CGPI in [5.50, 5.99] – Higher Second Class
CGPI in [6.0, 7.99] -- First Class
CGPI in [8.00, 10.00] -- Distinction

6. The percent marks equivalent a CGPA/SGPA is calculated by the formula:

$$\% \text{ Marks} = \left| \begin{array}{ll} 10 * \text{CGPA} & \text{if CGPA is in } [4.00,\ 6.00], \\ 5 * \text{CGPA} + 30 & \text{if CGPA is in } [6.00,\ 9.00], \\ 25 * \text{CGPA} - 150 & \text{if CGPA is in } [9.00,\ 10.00] \end{array} \right.$$

Semester/ Duration	Subject Code	Title of the Subject	Credits	L/P + T / week	Weightage for UE/IA	Exam Conduction
Semester - I **Foundation Level** (16-17 Weeks)	MBI 101	Basic Biosciences	**4**	**3L+1T**	0.6/0.4	University
	MBI 102	C programming and Data structure	**4**	**3L+1T**	0.6/0.4	University
	MBI 103	Bio-computing and DBMS	**4**	**3L+1T**	0.6/0.4	University
	MBI 104	Essential Biomathematics	**2**	**1L+1T**	0.6/0.4	University
	MBI 105	Biostatistics	**2**	**1L+1T**	0.6/0.4	University
	MBI 106	Biological Informatics	**4**	**3L+1T**	0.6/0.4	University
	MBI 107 **Elective I**	Bio-physics/ Metabolomics	**2**	**1L+1T**	Continuous Assessment	Institute
	MBI 108	Basic Biosciences Lab I	**2**	**2P**	0.6/0.4	University
	MBI 109	C programming and Data structure Lab II	**2**	**2P**	0.6/0.4	University
	MBI 110	Bio-computing, DBMS and Biostatistics Lab III	**2**	**2P**	0.6/0.4	University
	MBI 111	Biological Informatics Lab IV	**2**	**2P**	0.6/0.4	University
			Total = 30 credits (28 Foundation + 2Elective) Hours per day = 7.66 hrs.			**Total = 750**
Semester - II **Core Level** (16-17 Weeks)	MBI 201	Java and Biojava Programming	**4**	**3L+1T**	0.6/0.4	University
	MBI 202	Structural Biology & Molecular Modeling	**4**	**3L+1T**	0.6/0.4	University
	MBI 203	Genomics & Proteomics	**4**	**3L+1T**	0.6/0.4	University
	MBI 204	Perl and Bioperl Programming	**4**	**3L+1T**	0.6/0.4	University
	MBI 205 **Elective II**	Bio-ethics & IPR/ Immuno-informatics	**2**	**1L+1T**	Continuous Assessment	Institute
	MBI 206 **Elective III**	Software Engineering / Emerging Trends in Information Technology	**2**	**1L+1T**	Continuous Assessment	Institute
	MBI 207	Java and Biojava Programming Lab V	**2**	**2P**	0.6/0.4	University
	MBI 208	Structural Biology & Molecular Modeling Lab VI	**2**	**2P**	0.6/0.4	University
	MBI 209	Genomics & Proteomics Lab VII	**2**	**2P**	0.6/0.4	University

	MBI 210	Perl and Bioperl Programming Lab VIII	**2**	**2P**	0.6/0.4	University
	MBI 211 **Elective Lab I**	HTML and XML/ Introduction to Python Lab IX	**2**	**2P**	0.6/0.4	University
			Total = 30 Credits (24 Core + 6 Elective) Hours per day = 8.33 hrs.			**Total = 750**
Semester - III **Advanced Level** (16-17 Weeks)	MBI 301	Scientific Writing Skills	**2**	**1L+1T**	Continuous Assessment	Institute
	MBI 302	Cheminformatics & Drug Design	**4**	**3L+1T**	0.6/0.4	University
	MBI 303	Biological Data Mining	**4**	**3L+1T**	0.6/0.4	University
	MBI 304	Systems Biology	**4**	**3L+1T**	0.6/0.4	University
	MBI 305 **Elective IV**	Recent trends in Bioinformatics/ Commercial Bioinformatics	**4**	**3L+1T**	0.6/0.4	University
	MBI 306 **Elective V**	Bio-diversity Informatics/ Artificial Bio-intelligence	**2**	**1L+1T**	Continuous Assessment	Institute
	MBI 307	Cheminformatics & Drug Design Lab X	**2**	**2P**	0.6/0.4	University
	MBI 308	Biological Data Mining Lab XI	**2**	**2P**	0.6/0.4	University
	MBI 309	Systems Biology Lab XII	**2**	**2P**	0.6/0.4	University
	MBI 310 **Elective Lab II**	Recent trends in Bioinformatics/ Commercial Bioinformatics Lab XIII	**2**	**2P**	0.6/0.4	University
			Total = 28 Credits (20 Core +8 Elective) Hours per day = 7.33 hrs.			**Total = 700**
Semester – IV (16-17 Weeks)	MBI 401	Project	**12**	**12P**	-	University
			Total = 12 Credits Hours per day = 6.00 hrs.			**Total = 600**
M.Sc. Bioinformatics (Sem I to Sem IV) Total: 100 Credits						**Total Marks= 2800**

Semester I

MBI 101: Basic Biosciences

Foundation Course **(4 Credits, 3L+1T, 60L)**

Cell & Molecular Biology

Unit 1: **10L**

1.1. Prokaryotic & Eukaryotic cell concepts, structural organization of plant and animal cells, cell membrane and cell wall
1.2. Cell cycle: Cell division phases, regulation & inhibitors, Apoptosis
1.3. Basic concepts of signal transduction.

Unit 2: **10L**

2.1 Membrane transport**:** Passive, Facilitated and active transport, Na+/K+ ATPase structure and function, secondary active transport, Ion channels, leak channels and gated channels.
2.2 Cytoskeleton: Actin filaments, Intermediate filaments, Microtubules, Organization and functions, Cell - cell interactions, Tight junctions, gap junctions, Desmosomes and hemidesmosomes, plasmodesmata.

Unit 3: **10L**

3.1 Genetic code: A brief account
3.2 Nucleic acids: A brief historical review of major discoveries in the last century, Structure, properties of different forms of DNA and RNA and their respective functions.
3.3 Nature of Genome: Genome sizes of different organisms, C value paradox, Repetitive sequences, clusters and repeats, Concept of gene, Interrupted and uninterrupted genes, intron, exon and their relationships, overlapping gene, Introduction to promoters, Regulatory sequences, Transposons.

Genetics

Unit 4: **10L**

4.1 Mendel's laws of inheritance and their chromosomal basis, extra chromosomal inheritance
4.2 Organization of prokaryotic and eukaryotic genome: Organization of bacterial genome, Nucleoid structure, Nature of supercoiling, Twisting number, Writhing number, Chromatin, Euchromatin and Heterochromatin, Organization of chromatin, Chromosome structure and role, Chromosome banding, Telomere and replication of telomere sequences, Histones, Nucleosome structure, Extra-Chromosomal Genomes like Mitochondrial, Chloroplast & Bacterial Plasmids.

Unit 5: **10L**

5.1 Molecular mechanism of general recombination, homologous and site-specific, Holiday model of recombination, gene conversion.
DNA Replication: Concept of replicon, DNA polymerases, Priming reactions, Semiconservative replication of DNA, Replication fork, Synthesis of leading and lagging strands, Termination in Prokaryotes and Eukaryotes, Comparative account of phage, Prokaryotic and eukaryotic replication, regulation of initiation, licensing factor.
5.2 Types of mutation, mutagens, molecular mechanisms of mutation, site-directed mutagenesis, detection and isolation of mutants, transposons in mutation, repair mechanisms
5.3 DNA repair: DNA damage and response, Excision repair, mismatch repair, recombination repair, error prone repair, SOS response, repair system in eukaryotic cells.

Unit 6: **10L**

5.1 Prokaryotic Transcription: RNA polymerase structure, role, promoter recognition, Sigma factors
5.2 Eukaryotic transcription: Eukaryotic RNA polymerase and their promoters, TBP, Promoters and enhancers, Activating transcription, Post transcriptional modifications, RNA splicing reactions, catalytic RNA.

References:

- Alberts B. and Jhonson A.4th edition (2002) Molecular Biology of the cell, Garland science.
- Berg J., Tymcoczko J, and Styrer L, 5th edition(2002) Biochemistry, W. H. Freeman and company, New York.
- Watson J. and Stephen (2004) Molecular Biology of the Gene, Dorling Kindersley(India) pvt ltd,New delhi. Taylor and Francis group, NewYork.
- Cooper G.M. and Hausman R.E. (2004) The Cell: A molecular approach, Sinauer Associates, Inc., ASM Press, Washington DC.
- Lewin (2007) Genes IX: Pearson Prentice Hall, Pearson Education, Inc. Upper Saddle River, NJ 07458
- Human Molecular Genetics, 2nd edition, Tom Strachan and Andrew P Read, New York: Wiley-Liss; 1999., ISBN-10: 1-85996-202-5
- Genomes, 2nd edition,Terence A Brown,Oxford: Wiley-Liss; 2002, ISBN-10: 0-471-25046-5
- Developmental Biology, 6th edition, Scott F Gilbert,2000, ISBN-10: 0-87893-243-7
- Modern Genetic Analysis, Anthony JF Griffiths, William M Gelbart, Jeffrey H Miller, and Richard C Lewontin, New York: W. H. Freeman, 1999,ISBN-10: 0-7167-3118-5
- Lodish et. al. 2004, Molecular Cell Biology. Freeman.
- Karp 2002, Cell and Molecular Biology. John Wiley.
- Pollard & Earnshaw 2002, Cell Biology. Saunders.
- Tobin & Morcel 1997, Asking about Cells. Saunders.
- Atherly et. al. 1999, The Science of Genetics. Saunders.
- Hartl & Jones 1998, Genetics - Principles & Analysis. Jones & Bartlett.
- Snustad et. al. 1998, Principles of Genetics. Wiley & Sons.
- Strickberger 1985, Genetics. Macmillan.
- Russell 2002, Genetics. Benjamin.

Semester I

MBI 102: C Programming and Data Structure

Foundation Course **(4 Credits, 3L+1T, 60L)**

Unit 1: **10L**

1.1. Introduction to C Programming
1.2. Structure of C Program
1.3. Tokens in C
1.4. Identifiers
1.5. Operators
1.6. Expressions
1.7. Hierarchy of operators
1.8. Constants

Unit 2: **10L**

2.1. Data types
2.2. Type conversions
2.3. Type casting
2.4. C Instructions
2.5. Concept of flow chart

Unit 3: **10L**

3.1. Decision Making and Looping
3.2. Arrays -Array declaration and initialization

Unit 4: **10L**

4.1. Functions
4.2. Strings and string functions

Unit 5: **10L**

5.1. Pointers
5.2. Structures

Unit 6: **10L**

6.1. File Input/ Output
6.2. Data structures
6.3. Introduction to data structures
6.4. Types of Data structures

References:

- Let Us C by Yashavant Kanetkar, BPB Publications.
- The C programming language by Kerighan and Richie, PHI Publication.
- Programming in ANSI C by Balaguruswamy, Tata McGraw-Hill Education.
- Sams Teach Yourself C in 21 Days Peter Aitken and Bradley L. Jones, Macmillan Computer publishing.
- Schaum’s outline of programming with C by Byron Gottorfried.
- Algorithms in bioinformatics by Guigo R. Ed. & Gusfield D., Ed.: Berlin. Springer-Verlag, 2002.

Semester I

MBI 103: Bio-Computing and DBMS

Foundation Course **(4 Credits, 3L+1T, 60L)**

Unit 1: **20L**

1.1. Fundamentals of computing
1.2. Introduction to operating systems
1.3. WINDOWS, UNIX / LINUX
1.4. UNIX: Unix architecture, Unix file system, Unix commands, Pipes and filters (grep, cut etc.), file permissions (read, write, execute) for user, group, others

1.5. Computer viruses- An Introduction
1.6. Use of Internet and world wide web, searches on Medline, bibliographic databases, PDB
1.7. Networking-LAN, WAN, MODEM

Unit 2: **10L**

2.1. Database designing, data capturing
2.2. Data Abstraction
2.3. Data Models
2.4. Instances & Schemes
2.5. E-R Model - Entity and entity sets
2.6. Relations and relationship sets
2.7. E-R diagrams
2.8. Reducing E-R Diagrams to tables

Unit 3: **20L**

3.1. Basic concepts in Indexing and hashing
3.2. Types of Indexing
3.3. Data warehousing
3.4. Data mining
3.5. Oracle Architecture
3.6. Basic concepts in Oracle:
    - Tables
    - Views
    - Indexes
    - Synonyms
    - Table space
    - Data files
    - Blocks
    - Extents

    - Segments
    - Oracle Background Processes
    - Control files
    - Oracle Memory Management
    - Rollback
    - Redo logs etc.

Unit 4:

Part 1 **10L**

4.1. Oracle Report generation
4.2. Grants
4.3. Roles
4.4. Privileges

Part 2

4.5. SQL Statements:
    - Data Definition Statements
    - Data Manipulation Statements
    - Data Control Statements

4.6. Different forms of Select Statements
4.7. Create Table , View etc
4.8. Group functions

References:

- Database System Concepts by Hanery Korth and Abraham Silberschatz, McGraw Hill publication.
- An Introduction to Database Systems by C.J. Date, Addison-Wesley.
- Database systems by Nilkamal Surve, Tech Max publications.
- Data Mining: Concepts and Techniques by Jiawei Han and Micheline Kamber, Morgan Kaufmann Publishers.
- Oracle PL/SQL Programming by Steven Feuerstein, Bill Pribyl, O-Reilly Media.
- The Relational Database Dictionary by C.J. Date, O'Reilly Media
- Fundamentals of Database Systems by Elmasri and Navathe, Pearson Education.
- Database Design and Relational Theory by C.J. Date, O'Reilly Media.
- Oracle Database 11g A Beginners Guide by Ian Abramson, Michael Abbey, Michael J. Corey and Michelle Malcher, McGraw Hill publication.
- Developing Bioinformatics Computer Skills by Gibas, & Jambeck, O-Reilly.
- Linux : The Complete Reference 6th Edition by Richard Petersen, Tata McGraw-Hill Education (2007)

Semester I

MBI 104: Essential Biomathematics

Foundation Course **(2 Credits, 1L+1T, 30L)**

Unit 1: **10L**

1.1. Calculus: Limits, Continuity, Analysis, Differentiation (1D & Partial), Reimann Integration, Definite Integrals, Recursion Theorems.
1.2. Ordinary & Partial Differential Equation: 1'st Order & 2'nd Order Ordinary Differential Equations. Self Adjoint Equations, Special Functions, Nature of Partial Differential Equations, Method of Separation of Variables.
1.3. 2Dimensional & 3Dimensional Geometry: Cartesians and Polar Coordinates, Locus, Equation of a straight line, pair of straight lines, circle, ellipse, parabola, hyperbola. Asymptotic Analysis, Equation of sphere, cone. 3D Coordinate System, Spherical and Cylindrical Coordinates.

Unit 2: **10L**

2.1 Integral transform: Fourier Series, Fourier Transform, Laplace Transform
2.2 Series And Sequences: Sequences, Infinite Series, Infinite Products
2.3 Numerical Techniques: Vector Algebra, Basic Computations, Matrices, Basic Formalism, Methods for Solving Equations, Finding Eigenvalues & Eigenvectors, Solving ODE & PDE, Differentiation

Unit 3: **10L**

3.1 Applications To Mathematical Biology: Enzyme kinetics, Immunology, Population genetics, Tumor modeling, Applications of ordinary & partial differential equations to Biology.

References:

- Introduction to mathematical methods in bioinformatics by Isaev, Alexander Berlin Springer, 2004.
- Mathematics in chemistry by K. V. Raman & Pal, Sourav, New Delhi, Vikas publishing house Pvt. ltd., 2005
- Calculations in molecular biology and biotechnology: a guide to mathematics in the laboratory by Stephenson, F.H. Amsertdam, Academic Press, 2003.
- Advanced mathematical methods for engineering & science students by Stephenson, G. & Radmore, P. M., 1990.
- Mathematics and computer science in medical imaging by Viergever, Max A.& Todd-Pokropek, Andrew.,1988.
- Mathematical Methods for Physicists by G. Arfken, Academic Press, New York,1970.
- Some Modern Mathematics for Physicists and Other Outsiders by P. Roman, Pergamon, New York, 1975, Vol. 2, p. 660.
- Mathematical Biology, by J. D. Murray Springer Verlag, 1989.
- Mathematical Models in Molecular and Cellular Biology by Segal, L., ed. 1980. Cambridge: Cambridge University Press.
- Numerical Methods by Balaguruswamy, TMH.

Semester I

MBI 105: Biostatistics

Foundation Course **(2 Credits, 1L+1T, 30L)**

Unit 1: **10L**

1.1. Introduction to principles of statistical sampling from a population Calculus: Limits, Continuity, Analysis, Differentiation (1D & Partial), Reimann Integration, Definite Integrals, Recursion Theorems.
1.2. Probability Distributions Introductory Concepts, Conditional Probability, Bayesian Rules, Frequency Distributions and Statistical Measures, Random variable, Integration of random variables, Vector valued Random Variables

Unit 2: **10L**

2.1 Correlation and regression analysis.
2.2 Multivariate analysis
2.3 Hypothesis testing
2.4 Markov Models, Markov Chains & Hidden Markov Models
2.5 Cluster Analysis:
   - Nearest neighbour search
   - Phylogenetic Analysis Tools: UPGMA, Maximum Likelihood, Parsimony methods, Distance methods, Model Comparison, Neighbor joining Methods

Unit 3: **10L**

3.1 Neural Networks:
   - Neural Networks
   - Neural Networks Tools- Brain Box, MATLAB etc.
   - Use in Protein Structure Prediction

3.2 Probabilistic Modelling and Inference Ghgfh
3.3 Graph theory in bioinformatics
3.4 Fuzzy Logic and Application

References:

- Probability statistics, and reliability for engineers by Boca Raton, Ayyub B. M. & McCuen, R H, CRC Press, 1997.
- Statistics: concepts and applications by Frank, Harry & Althoen, S. C., Cambridge University Press, 1995.
- Statistical methods in bioinformatics: an introduction by Ewens, W. J. & Grant, G. R., New York. Springer, 2001.
- Handbook of computational statistics: concepts and methods by Gentle, J.E., Hardle, W. & Mori, Y., Berlin, Springer-Verlag, 2004.
- Statistical design and analysis of industrial experiments by Ghosh, Subir, Ed., 1990.
- Scan Statistics by Glaz, J., Naus, J. & Wallenstein S, New York, Springer, 2001.
- Statistical design for research by Kish, L., Wiley series in probability and mathematical statistics, New York, John Wiley $ Sons, 1987.
- Introduction to probability and statistics by Lipschutz, S. & Schiller, J. J., New York. McGraw-Hill, 1999.
- Schaum's outline of theory and problems of statistics by Spiegel, M. R. & Stephens, L. J., Ed. 3, New Delhi, Tata McGraw-Hill Publishing Co. Ltd., 2001.
- Mathematical and statistical methods for genetic analysis by Lange, K., 2nd Ed., New York. Springer-Verlag, 2002.

Semester I

MBI 106: Biological Informatics

Foundation Course **(4 Credits, 3L+1T, 60L)**

Unit 1: **10L**

1.2. Nature of biological data
1.3. Introduction to Bioinformatics
1.4. Overview of available Bioinformatics resources on the web
1.5. Primary Resource Institutes: NCBI, EMBL & DDBJ
1.6. Hierarchy of Biological databases: Primary, Secondary & Derived

Unit 2: **10L**

2.1 Primary Nucleic acid databases: GenBank, EMBL Nucleotide, DDBJ Nucleotide
2.2 Primary Protein sequence databases: NCBI Protein, EMBL Protein, PIR-PSD , SwissProt/ UniProtKB/ TrEMBL
2.3 Database search engines: Entrez, SRS
2.4 Database Filters: SEG, XNU,DUST
2.5 Sequence Submission Tools: Sequin, BankIt, ENA, IMGT/HLA,DGVa, SPIN, Metagenomics

Unit 3: **10L**

3.1 Overview/concepts in sequence analysis: Local & Global alignment, DotPlot, Gap Penalties
3.2 Dynamic Programming, Heuristic Methods
3.3 Pairwise Sequence Alignment algorithms: Needleman & Wunsch, Smith & Waterman
3.4 Scoring matrices for Nucleic acids and proteins: PAM/MDM, BLOSUM, CSW

Unit 4: **10L**

4.1 Database Similarity Searches: BLAST & FASTA
    Other Tools: LALIGN, Dotlet
4.2 Multiple Sequence Alignment: ClustalW, ClustalX, PRAS
    Other Tools: DbClustal, Kalign, MAFFT, MUSCLE, MView, T-Coffee
4.3 Motifs, pattern & Profiles
4.4 Alignment Analysis tools: AMAS, CINEMA, MaxAlign, PhyloGibbs, SVA, PVS
4.5 Phylogenetic Analysis Tools: BIONJ, PHYLIP, PHYML, POWER, BlastO, Evolutionary Trace Server

Unit 5: **10L**

Biological Databases:

5.1 Genome Databases: Human, model organisms, microbes & viral
Genetic Disorders Databases: OMIM, OMIA, Genetic Association Database, Genetic Disorder Guide, IGDD, DisGenet, Genetic Disorder UK

5.2 Structural Databases: PDB, SCOP, CATH, NDB, CCSD, CSD

5.3 Derived Databases: PROSITE, BLOCK, ProDom, Pfam, PRINTS, SBASE,

Unit 6: **10L**

6.1 Primer Designing Tools: Primer, Primer3, NetPrimer, Primerfinder

6.2 DNA/RNA Sequence Analysis: CENSOR, Gene Finder, GENEID, GenHunt, GENIE, GRAIL, ORD ID, ORF Finder, ORFGene, Pol3Scan, tRNAscan

6.3 Translation Tools: Translation Tool, The Protein Machine, 6 Frame Translation Tool, Reverse Translation Tools

6.4 Restriction Analysis Tools: WEB Cutter, ENZFINDER, TACG

6.5 Molecular visualization tools: RasMol, Cn3D, SPDBV, Chime, Mol4D, etc

6.6 Databases and search methods for chemical compounds: PubChem Compound, PubChem Substance, ChEBI, ChEMBL, PDBeChem, RESID, EuroCarbDB

References:

- Introduction to Bioinformatics by Attwood, T.K. & Parry-Smith, D.J., Delhi, Pearson Education (Singapore) Pte.Ltd., 2001.
- Bioinformatics: Sequence and Genome Analysis by Mount, David, New York, Cold Spring Harbor Laboratory Press, 2004.
- Current Protocols in Bioinformatics by Baxevanis, A.D., Davison, D.B., Page, R. D. M. & Petsko, G.A., New York, John Wiley & Sons Inc., 2004.
- Claverie, J.M. and Notredame C. 2003 Bioinformatics for Dummies. Wiley Editor.
- Letovsky, S.I. 1999 Bioinformatics. Kluwer Academic Publishers.
- Baldi, P. and Brunak, S. 1998 Bioinformatics. The MIT Press.
- Setubal, J. and Meidanis, J. 1996 Introduction to Computational Molecular Biology. PWS Publishing Co., Boston.
- Lesk, A.M. 2002 Introduction to Bioinformatics. Oxford University Press.
- Rastogi, S.C., Mendiratta, N. and Rastogi, P. 2004 Bioinformatics: Concepts, Skills & Applications. CBS Publishers & Distributors, New Delhi.
- Vyas, S.P. and Kohli, D.V., Methods in Biotechnology and Bioengineering.
- Genetic Library Construction and Screening: Advanced Techniques and Applications: Lab Manual
- Mont, D.W., Bioinformatics: Sequence and Genome Analysis.
- Pierre Baldi and Soren Brunak, Bioinformatics: The Machine Learning Approach.

Semester I

MBI 107: Bio-physics

Elective Course I **(2 Credits, 1L+1T, 30L)**

Unit 1: **10L**

1.1 Introduction to biophysics: historical overview, connections with physics, biology and medicine.
1.2 Causation as the main acting principle in the physical reality.
1.3 Mathematics as the exact instrument of the quantitative methods.
1.4 Principle of measurement. Physical units of measurement, their systems. The SI system. Measurement uncertainty, ways to express it.
1.5 Equalities and equations as the language of formulae used in natural sciences.
1.6 Functions: notion, properties, forms of presentation, applications in physics. Proportionality and linear function. Linear and non-linear systems. Principle of superposition. Derivative and integral of a function, their applications.

Unit 2: **10L**

2.1 X-rays, their properties, sources of X-radiation, Application of X-rays to biology and medicine.
2.2 Basic principles of X-ray diffraction studies.
2.3 Electron density map and its interpretation.
2.4 Protein structure determination by X-rays and electron diffraction.

Unit 3: **10L**

3.1 NMR techniques for structure determination, elementary idea of UV/visible, IR, Raman techniques in biological context.
3.2 Mass spectrometry
3.3 Microscopy

References:

- Jan Drenth Principles of Protein X-ray Crystallography (Springer Advanced Texts in Chemistry), Springer-Verlag Telos.
- Lipson, H. and Steeple, H., Interpretation of X-Ray Powder Diffraction Patterns. St. Martin's Press.
- Bovey, F,A., Mirau, P.A. and Gutowsky, H.S., Nuclear Magnetic Resonance Spectroscopy (2nd Edition). Academic Press.
- Hallet, F.R., Stinson, R.H., Speight, P.A. and Graham, W.G., 2004 Physics for the Biological Science, Toronto: Nelson Can.
- Yadav, L.D.S., Organic Spectroscopy (1st Edition). Springer.
- Hollas, J.M., High Resolution Spectroscopy (2nd Edition). John Wiley & Sons.

Semester I

MBI 107: Metabolomics

Elective Course I **(2 Credits, 1L+1T, 30L)**

Unit 1: Metabolomics – Data Collection and Analysis **10L**

1.1 Introduction to Metabolomics, datasets, collection strategies

1.2 Repostories for metabolomics datasets, KEGG, Panther, Brenda, BioCyc

Unit 2: Metabolic Control Analysis **10L**

2.1 Reductionist theory in MCA, Control and regulation studies

2.2 Steady state analysis, control coefficients

2.3 Elasticity and connectionist theorem

Unit 3: Flux-Balance Analysis **10L**

3.1 Systems – their definition and analysis

3.2 Mass balance analysis, measurement of fluxes

3.3 Optimization strategies

References:

- Niek Yoan, Flux Balance Analysis, Miss Press.
- Luonan Chen, Rui-Sheng Wang, Xiang-Sun Zhang, Biomolecular Networks: Methods and Applications in Systems Biology (Wiley Series in Bioinformatics), Wiley.
- Sang Yup Lee, E. Terry Papoutsakis, Metabolic Engineering (Biotechnology and Bioprocessing), CRC Press.

Semester I

MBI 108: Basic Biosciences Lab I

Foundation Course: Practical Course in Basic Biosciences **(2 Credits, 2P=6 Hrs.)**

Cell and Molecular Biology

1. Isolation, purification and estimation of DNA from viruses/ bacteria /Yeasts / plant cells/ animal cells. **2P**
2. Demonstration of Metagenomics DNA extraction **1D**
3. Isolation, purification and estimation of RNA & mRNA from various cell types **2P**
4. Estimation of DNA & proteins using spectrophotometric methods. **2P**
5. Size fractionation of DNA by gel electrophoresis **2P**
6. Determination of Tm and GC% of DNA. **2P**
7. Demonstration of DNA amplification by PCR, Real Time PCR. **3P**
8. Maintenance and sub-culturing of viral/bacterial/plant/animal cell cultures **1D**

Genetics

9. Cloning of DNA fragment using bacterial host vector system. **2P**
10. Transformation of host cells by reporter genes and expression analysis **2P**
11. Plasmid isolation **2P**
12. Restriction mapping **2P**
13. In Vitro transcription **2P**
14. Phage DNA Isolation **1P**
15. Restriction digestion and size fractionation of DNA **2P**
16. Ligation **1P**
17. Industrial Visit **1D**

References:

- Sombrook J. and Russell D. (2001) Molecular cloning, A Laboratory manual, Cold spring harbor laboratory press, New York.
- Wilson K. and Walker J. (2005) Principles and Techniques of Biochemistry and Molecular Biology, Cambridge University press, New York.
- Tablot N. (2005) Molecular and cellular Biology of Filamentous Fungi, Practical Approach, Indian Edn., Oxford University Press, New Delhi (ISBN: 0-19-567943-1).
- Joseph sambrook & david W. Russell, Cold Spring Harbor Laboratory, New York.

Semester I

MBI 109: C-Programming and Data Structure Lab II

Foundation Course: Practical Course in C-Programming and Data Structure

(2 Credits, 2P=6 Hrs.)

1. Laboratory assignments based on the following topics in 'C' programming: 15P

- Data types, operators and expressions, Hierarchy of operators,
- control statements including decision (if, if-else), loops (while, do-while, for), branching statements(switch, break, continue),
- Functions,
- Arrays (1D, 2D- all matrix operations including inverse of a matrix),
- Strings,
- File handling etc.

2. Writing C programs for Bioinformatics applications: 15P

- Determining the base composition in a nucleic acid sequence and amino acid composition in a protein sequence.
- Generating the complimentary sequence of a DNA sequence
- Pattern search algorithms
- Search for a specific oligonucleotide pattern (eg. GAACATCC) in a given
- DNA sequence.
- Locate palindrome sequence stretches in a DNA sequence.

References:

- Let Us C by Yashavant Kanetkar, BPB Publications.
- The C programming language by Kerighan and Richie, PHI Publication.
- Programming in ANSI C by Balaguruswamy, Tata McGraw-Hill Education.
- Sams Teach Yourself C in 21 Days Peter Aitken and Bradley L. Jones, Macmillan Computer publishing.
- Schaum's outline of programming with C by Byron Gottorfried.
- Algorithms in bioinformatics by Guigo R. Ed. & Gusfield D., Ed.: Berlin. Springer-Verlag, 2002.

Semester I

MBI 110: Bio-computing, DBMS and Biostatistics Lab III

Foundation Course: Practical Course in Bio-computing, DBMS and Biostatistics

(2 Credits, 2P=6 Hrs.)

Part I: ORACLE 10P

1.1. Assignment based on "Data Definition Language".
  - A set of SQL commands used to create table, modify table structure, drop table, rename table.

1.2. Assignment based on "Data Manipulation Language"
  - A set of SQL commands used to change the data within the database.
  - Insertion of records in the tables, updating of all or specific set of records in tables.
  - Viewing the attributes of table's column.

1.3. Assignment based on "Data Query Language"
  - Different forms of Select statement
  - Arithmetic and logical operators.
  - Range searching and pattern matching.

1.4. Assignment based on defining Constraints.
  - Types: I/O constraints like Primary Key, Foreign key, Null and Unique
  - Constraints. Business constraints like check constraints.

1.5. Assignment based on using joins.
  - Joining multiple tables, joining a table to itself.

1.6. Granting rights on user objects such as Tables, Views, and Sequences.

1.7. Revoking rights on user objects such as Tables, Views, and Sequences.

Part 2: Basic Computer Commands 05P

2.1. Familiarization with DOS, Windows, UNIX operating systems

2.2. UNIX commands, Pipes and Filters

Part 3: R and Matlab 15P

3.1.Basic Commands in R and Matlab

References:

- Database System Concepts by Hanery Korth and Abraham Silberschatz, McGraw Hill publication.
- An Introduction to Database Systems by C.J. Date, Addison-Wesley.
- Database systems by Nilkamal Surve, Tech Max publications.
- Data Mining: Concepts and Techniques by Jiawei Han and Micheline Kamber, Morgan Kaufmann Publishers.
- Oracle PL/SQL Programming by Steven Feuerstein, Bill Pribyl, O-Reilly Media.
- The Relational Database Dictionary by C.J. Date, O'Reilly Media
- Fundamentals of Database Systems by Elmasri and Navathe, Pearson Education.
- Database Design and Relational Theory by C.J. Date, O'Reilly Media.
- Oracle Database 11g A Beginners Guide by Ian Abramson, Michael Abbey, Michael J. Corey and Michelle Malcher, McGraw Hill publication.
- Developing Bioinformatics Computer Skills by Gibas, & Jambeck, O-Reilly.
- Linux : The Complete Reference 6th Edition by Richard Petersen, Tata McGraw-Hill Education (2007)
- Probability statistics, and reliability for engineers by Boca Raton, Ayyub B. M. & McCuen, R H, CRC Press, 1997.
- Statistics: concepts and applications by Frank, Harry & Althoen, S. C., Cambridge University Press, 1995.
- Statistical methods in bioinformatics: an introduction by Ewens, W. J. & Grant, G. R., New York. Springer, 2001.
- Handbook of computational statistics: concepts and methods by Gentle, J.E., Hardle, W. & Mori, Y., Berlin, Springer-Verlag, 2004.
- Statistical design and analysis of industrial experiments by Ghosh, Subir, Ed., 1990.
- Scan Statistics by Glaz, J., Naus, J. & Wallenstein S, New York, Springer, 2001.
- Statistical design for research by Kish, L., Wiley series in probability and mathematical statistics, New York, John Wiley $ Sons, 1987.
- Introduction to probability and statistics by Lipschutz, S. & Schiller, J. J., New York. McGraw-Hill, 1999.
- Schaum's outline of theory and problems of statistics by Spiegel, M. R. & Stephens, L. J., Ed. 3, New Delhi, Tata McGraw-Hill Publishing Co. Ltd., 2001.
- Mathematical and statistical methods for genetic analysis by Lange, K., 2nd Ed., New York. Springer-Verlag, 2002.
- Jae K. Lee, Statistical Bioinformatics, John Wiley & Sons Inc.
- Patterson, B.K., Techniques in Quantification and Localization of Gene Expression.
- Evens, W.J. and Grant, G.R., Statistical Methods in Bioinformatics: An Introduction.
- Liu, B.H., Statistical Genomics: Linkage Mapping and QTL Analysis

Semester I

MBI 111: Biological Informatics Lab IV

Foundation Course: Practical Course in Biological Informatics (2 Credits, 2P=6 Hrs.)

1.	Exploring the integrated database system at NCBI server and querying the PUBMED and GenBank databases using the ENTREZ search engine	**1P**
2.	Exploring the integrated database system at EBI server and searching the EMBL Nucleotide database using the SRS search engine	**1P**
3.	Exploring & querying SWISSPROT & UniProtKB	**1P**
4.	Exploring and querying the PIR database	**1P**
5.	Pair-wise global alignments of protein and DNA sequences using Needleman Wunsch algorithm & interpretation of results to deduce homology between the sequences, use of scoring matrices	**1P**
6.	Pair-wise local alignments of protein and DNA sequences using Smith-Waterman algorithm and interpretation of results.	**1P**
7.	Database (homology) searches using different versions of BLAST and interpretation of the results to derive the biologically significant relationships of the query sequences (proteins/DNA) with the database sequences	**1P**
8.	Exploring other databases: LALIGN, Dotlet	**1P**
9.	Database (homology) searches using different versions of FASTA & interpretation of the results to derive the biologically significant relationships of the query sequences (proteins/DNA) with the database sequences	**1P**
10.	Multiple sequence alignments of sets of sequences using web-based and stand-alone version of CLUSTAL. Interpretation of results to identify conserved and variable regions and correlate them with physico-chemical & structural properties.	**2P**
11.	Exploring other databases: DbClustal, Kalign, MAFFT, MUSCLE, MView, T-Coffee	**2P**
12.	Exploring Alignment Analysis tools: AMAS, CINEMA, MaxAlign, PhyloGibbs, SVA, PVS	**2P**

13. Exploring Phylogenetic Analysis Tools: BIONJ, PHYLIP, PHYML, POWER, BlastO, Evolutionary Trace Server **2P**

14. Exploring Genetic Disorders Databases: OMIM, OMIA, Genetic Association Database, Genetic Disorder Guide, IGDD, DisGenet, Genetic Disorder UK for different aspect of genetic disorders. **2**P

15. Exploring and using the derived databases: PROSITE, PRINTS, BLOCKS, ProDom, Pfam SBASE for pattern searching, domain searches etc. **2P**

16. Exploring Structural Databases: PDB, SCOP, CATH, NDB, CCSD, CSD for detailed knowledge on biomolecules. **2P**

17. Studying the format & content of structural databases & visualization of structures using Rasmol, Cn3D and other utilities **1P**

18. Exploring Primer Designing Tools: Primer, Primer3, NetPrimer, Primerfinder for utilization in research. **1P**

19. Exploring DNA/RNA Sequence Analysis: CENSOR, Gene Finder, GENEID, GenHunt, GENIE, GRAIL, ORD ID, ORF Finder, ORFGene, Pol3Scan, tRNAscan to get familiar with these databases. **1P**

20. Exploring Translation Tools: Translation Tool, The Protein Machine, 6 Frame Translation Tool, Reverse Translation Tools to get familiar with these databases. **1P**

21. Exploring Restriction Analysis Tools: WEB Cutter, ENZFINDER, TACG to get familiar with these databases. **1P**

22. Exploring Molecular visualization tools for their respective applications: RasMol, Cn3D, SPDBV, Chime, Mol4D, etc **1P**

23. Databases and search methods for chemical compounds: PubChem Compound, PubChem Substance, ChEBI,ChEMBL, PDBeChem, RESID, EuroCarbDB **1P**

References:

- Bioinformatics: A Practical Guide to the analysis of Genes and Proteins (3rd Ed.) by Baxevanis, A.D. & Ouellettee, B., F. F., New York, John Wiley & Sons, Inc. Publications, 2004.
- Practical Bioinformatics, Michael Agostino, 1stEdition, ISBN:9780815344568 , September 26, 2012
- Bioinformatics: A Practical Approach, Shui Qing Ye, CRC Press, 20-Aug-2007
- Bioinformatics: Methods Express, Paul Dear (Editor) , Publication Date: September 18, 2007 | ISBN-10: 190484216X | ISBN-13: 978-1904842163 | Edition: 1st Edition
- Current Protocols in Bioinformatics by Baxevanis, A.D., Davison, D.B., Page, R. D. M. & Petsko, G.A., New York, John Wiley & Sons Inc., 2004.
- Bioinformatics: Sequence, Structure and Databanks: A Practical Approach [Paperback] Des Higgins (Editor), Willie Taylo, Publication Date: October 15, 2000 | ISBN-10: 0199637903 | ISBN-13: 978-0199637904 | Edition: 1

SEMESTER II

MBI 201: Java and Biojava Programming

Core Course **(4 Credits, 3L+1T, 60L)**

Unit 1: **10L**

1.1. Introduction to Java programming
1.2. Introduction to classes and Methods
1.3. Java Basics
1.4. Data types in JAVA
1.5. Arithmetic and Logical operations

Unit 2: **10L**

2.1.Arrays in JAVA
2.2.Conditional statement and Loops

Unit 3: **10L**

3.1. Exception Handling
3.2. Try---catch block
3.3. Throw
3.4. Throws
3.5. Finally
3.6. Inheritance, Use of Super keyword

Unit 4: **10L**

4.1. Modifiers, Access Control and Class Design
4.2. Packages
4.3. Interfaces
4.4. Multithreading

Unit 5: **20L**

5.1. Java Applets Basics
5.2. Working with Graphics, Fonts and Color
5.3. Simple Animation
5.4. Event handling
5.5. Mouse events

5.6. Keyboard events
5.7. Creating User Interfaces
5.8. AWT controls
5.9. Windows, Networking and other Tidbits

References:

- The Complete Reference by Patrick Naughton and Herbert Schildt, McGraw Hill.
- Computer Programming – II, Mumbai University by Bhakti Raul and Vinayaa Rwool, Tech-Max publication.
- Sams Teach Yourself Java in 21 Days (Covering Java 7 and Android) by Rogers Cadenhead, Sams publication.
- Java In a Nutshell by David Flanagan, Oreilly Publications.
- Java Examples in a Nutshell by David Flanagan, Oreilly Publications.

Semester II

MBI 202: Structural Biology & Molecular Modeling

Core Course **(4 Credits, 3L+1T, 60L)**

Unit 1: **10L**

1.1 Algorithms of visualization tools of tertiary structures, quaternary structures, architectures and topologies of proteins and DNA, such as RasMol, Cn3D, SPDBV, Chime, PyMOL etc.,
1.2 Secondary structure prediction using Chou Fasman, GOR methods; analysis of results and measuring the accuracy of predictions using Q3, Segment overlap, Mathew's correlation coefficient.

Unit 2: **10L**

2.1 Fundamentals of the methods for 3D structure prediction (sequence similarity/identity of target proteins of known structure, fundamental principles of protein folding etc.)
2.2 Homology Modeling, fold recognition, and ab-initio structure prediction methods - protocols/algorithms.
2.3 Prediction of protein structure: PHYRE and PSI-PRED methods

Unit 3: **10L**

3.1 3-D structure comparison and concepts.
3.2 Algorithms such as FSSP, CE, VAST and DALI, Fold Classes.
3.3 Databases of structure-based classification: CATH and SCOP.

Unit 4: **10L**

4.1 General features- bond stretching, angle bending, improper torsions, cross terms, non-bonded interactions, point charges, calculation of atomic charges, polarization, van der waals interactions, hydrogen bond interactions, Water models.
4.2 Types of force field - all atoms force field, united atom force field, etc.

Unit 5: **10L**

5.1 Molecular Energy minimization methods such as steepest descent, conjugate gradient – derivatives, First order steepest decent and conjugate gradients.
5.2 Second order derivatives Newton-Raphson, Minima, maxima saddle points and convergence criteria.-non derivatives minimization methods, the simplex.

Unit 6: **10L**

6.1 Molecular Dynamics Simulation methods: Newtonian dynamics, Integration algorithm, Periodic boundary conditions and minimum image convention, Potential truncation and shifted-force potentials, Neighbor list, Force calculations, Long range interactions, MD code for liquid Argon.

6.2 Classical Monte Carlo: Random numbers, Evaluating integrals using random numbers, Importance sampling, Metropolis algorithm.

6.3 Analysis of simulated trajectories: Radial distribution functions, Self diffusion coefficient, Time correlation functions

References:

- Molecular Modeling Principles and Applications (2nd Ed.) by Andrew R. Leach, 2001, Prentice Hall, USA.
- Principles of Protein Structure by G. E. Schulz, 2009, Springer
- Textbook of Structural Biology by Anders Liljas, Lars Liljas, Jure Piskur and GÃran Lindblom, World Scientific Publishic Co. Pvt. Ltd.
- Structural Bioinformatics, Edited by Jenny Gu, Philip E. Bourne, Wiely Blackwell
- Computational Structural Biology: Methods and Applications by orsten Schwede, Torsten Schwede, Manuel C. Peitsch, 2008, World Scientific Publishing Company.
- Molecular Modeling Of Nucleic Acids, Leontis N.B. and SantaLucia J. Jr. (eds), 1998, American Chemical Society.
- Molecular Modeling of Proteins by Andreas Kukol, 2008, Humana Press.
- Molecular Modeling and Simulation: An Interdisciplinary Guide by Schlick T., 2012, Springer.
- Biological Modeling and Simulation: A Survey of Practical Models, Algorithms, and Numerical Methods (Computational Molecular Biology) by Russell Schwartz, The MIT Press, Cambridge, Massachusetts, London, England

Semester II

MBI 203: Genomics & Proteomics

Core Course **(4 Credits, 3L+1T, 60L)**

Genomics

Unit 1: **10L**

1.1. Basic genetics (covered in MBI 102)
1.2. Methods of preparing genomic DNA; DNA sequence analysis methods: Sanger Dideoxy method and Fluorescence method.
1.3. Genome sequencing: strategies & approaches
1.4. Genome Assembly
1.5. Gene variation and Single Nucleotide Polymorphisms (SNPs); Expressed sequenced tags (ESTs); Gene-disease association OMIM; Genome Databases STS, GSS, HSS.
1.6. Recombinant DNA technology: DNA cloning basics, Polymerase chain reaction, DNA fingerprinting, Human genome project (HGP) and the genetic map.

Unit 2: **10L**

2.1 Tools for Genomic Data Mining: Basic Aspects of Genome Annotation
2.2 Database Search Engines: Special tools for searching genomic data
2.3 Prediction of ORFs and Genes; Gene Modeling
2.4 Prediction of Signal sequences (Promoters, Primers, splice sites, UTRs etc.), Operons
2.5 Identification of Disease Genes: Identification of Drug Targets, Metabolic diseases and Pathogenic diseases, Gene Expression Analysis
2.6 Structural Genomics and Functional Genomics

Unit 3: **10L**

3.1 Pharmacokinetics: Classification, Case study, Comparative screening, ADMET: Drug metabolism; Role of cytochrome P450; Elimination half-life; Toxicity screening
3.2 Pharmacogenetics:
  - ➢ The genetics of drug metabolism
  - ➢ The genetics of therapeutic targets
3.3 Genome Alignments: BLAST2, MUMmer , PipMaker , VISTA
3.4 Comparison of Gene Order: GeneOrder , Gene synteny
3.5 Comparative Genomics of organisms: Viruses, Microbes, Pathogens, Eukaryotes
3.6 Comparative Genomics Databases : COG, VirGen, CORG, HOBACGEN, Homophila, XREFdb

Proteomics

Unit 4: **10L**

4.1 Overview of Proteomics: Introduction and scope of proteomics; Protein separation techniques: ion-exchange, size-exclusion and affinity chromatography techniques.
4.2 Polyacrylamide gel electrophoresis; *Isoelectric focusing (IEF)*; Two dimensional PAGE for proteome analysis; Image analysis of 2D gels

Unit 5: **10L**

5.1 Introduction to mass spectrometry; Strategies for protein identification; Protein sequencing; Protein modifications and proteomics; Applications of proteome analysis to drug; Protein-protein interaction (Two hybrid interaction screening).
5.2 Protein engineering; Protein chips and functional proteomics; Clinical and biomedical application of proteomics; Proteome database; Proteomics industry.

Unit 6: **10L**

3.1 Protein-Protein Interaction Networks, databases and software:
- DIP (Database of Interacting Proteins)
- PPI Server
- BIND - Biomolecular Interaction Network Database
- PIM –Hybrigenics
- PathCalling Yeast Interaction Database
- MINT - a Molecular Interactions Database
- GRID - The General Repository for Interaction Datasets
- InterPreTS - protein interaction prediction through tertiary structure

References*:*

- Guide to Human Genome Computing by Martin J. Bishop, Academic Press. ISBN 0-12-102051-7.
- From Genome to Therapy: Integrating new technologies with drug development by Novartis Foundation, John Wiley. ISBN 0-471-62744-5.
- Genome mapping and sequencing By Ian Dunham, Horizon, ISBN1-898486-50-6.
- The Genome by Ram S. Verma, VCH, ISBN 1-56081-043-2.
- Bioinformatics - from genomes to drugs (vol. 1), basic technologies (vol.1) by Lengauer, T., Germany, Wiley-VCH, 2002.
- Vaccines: New Approaches to Immunological problems by Ronald W. Ellis, BH. ISBN 0-7506-9265-0.
- Human, Vaccines and Vaccination by M. Mackett and J.D. Williamson, BIOS. Scientific, ISBN 1-872748-7-5.
- Principles of Genome Analysis And Genomics (3rd Ed.) by Primrose, S.B. & Twyman, R.M., UK. Blackwell Publishing Company, 2003.
- Bioinformatics approach Guide to the analysis of genes and proteins by Andceas Baxevanis and B.F. Francis Ouellettee. John Wiley 2004.
- Fogel, G.B. and Corne, D.W., Evolutionary Computation in Bioinformatics.
- Patterson, B.K., Techniques in Quantification and Localization of Gene Expression.
- Singer, M. and Barg, P. Exploring Genetic Mechanism.
- Bowtell, D. and Sambrook, J. DNA Microarrays.
- Fundamentals of Data Mining in Genomics and Proteomics, By Werner Dubitzky, Martin Granzow, Daniel P. Berrar, 2007, Springer Science + Bussiness Media, LLC.
- Protein Arrays, Biochips and Proteomics: The Next Phase of Genomic Discovery edited by Joanna S. Albala, Ian Humphery-Smith, ISBN-0-8247-4212-1, 2003, Marcel Dekker.

Semester II

MBI 204: Perl and Bioperl Programming

Core Course **(4 Credits, 3L+1T, 60L)**

Unit 1: **10L**

1.1. Introduction to Perl programming
1.2. Numbers, Strings
1.3. Arithmetic and Logical operators
1.4. Control statements and Loops
1.5. Functions (push, pop, length, log etc)

Unit 2: **10L**

2.1.Scalar variables
2.2.Working with Arrays
2.3.Input/output in Perl

Unit 3: **10L**

3.1.Subroutine
  - Defining subroutine
  - Invoking subroutine
  - Passing arguments to subroutine
  - Return values

3.2.Working with files
  - Filehandle
  - Opening file
  - Reading contents of file
  - Writing into files etc.

Unit 4: **10L**

4.1. Regular Expression
4.2. Pattern Matching
4.3. Hashes

Unit 5: **10L**

5.1. Object-Oriented Programming in Perl

- Objects, Methods, and Classes in Perl

5.2. Perl Modular Programming

- Use of Perl Modules
- Namespaces
- Packages
- Defining modules
- Storing modules
- First Perl module
- Using modules
- CPAN modules
- Bioperl module
  - Installing Bioperl
  - Testing Bioperl
  - Using Bioperl Modules

Unit 6: **10L**

6.1. How to Document a Perl Class with POD

6.2. Data Structures

- Complex Data Structures
- Printing Complex Data Structures

6.3. AUTOLOAD

- Use of AUTOLOAD and how it works

6.4. Perl and Web

- How the Web Works
- The Common Gateway Interface

References:

- Beginning Perl for Bioinformatics by James Tisdall, O-Reilly publication.
- Perl: Complete Reference Perl by Martin C. Brown, McGraw-Hill publication.
- Mastering Perl for Bioinformatics by James D. Tisdall, O-Reilly Publication.
- Teach Yourself Perl 5 in 21 days by David Till, Sams publishing.
- Mastering Algorithms with Perl by Jon Orwant, Jarkko Hietaniemi and John Macdonald, O-Reilly Publication.
- Professional Perl Programming by Peter Wainwrigh, Published by Wrox Press Ltd.
- Beginning Web Development with Perl by Steve Suehring, Apress publication.

Semester II

MBI 205: Bio-ethics & IPR

Elective Course II **(2 Credits, 1L+1T, 30L)**

Bio-ethics:

Unit 1: **10L**

1.1 Regulatory Procedures: Good laboratory practice, Good manufacturing practice and FDA regulations
1.2 Regulations for recombinant DNA research and manufacturing process
1.3 Biosafety and Bioethics - Regulations for clinical trials,
1.4 Documentation and Compliance, in India and selected countries
1.5 Rules for import and export of biological materials.

IPR

Unit 2: IPR- Patents:

10L

2.1 Introduction
2.2 Patentable Invention
2.3 Procedure for obtaining Patent
2.4 Rights conferred on a Patentee
2.5 Transfer of Patent
2.6 Case Studies
2.7 Case law on Patent.

Unit 3: IPR- Trade Marks:

10L

3.1 Introduction
3.2 Statutory Authorities
3.3 Principles of registration of Trade Marks,Procedure of registration and duration
3.4 Licensing in Trade Marks
3.5 Case law on Trade Marks
3.6 **IPR- Copyright:** Introduction
3.7 Author and Ownership of Copyright
3.8 Rights conferred by Copyright, Assignment/ Licence of Copyright
3.9 Case Law on Copyright

References:

- A life-centered approach to bioethics : biocentric ethics by Lawrence E Johnson, 2007, Georgetown University Press
- Biethics by Ellen Frankel Paul, Fred D. Miller Jnr, Jeffrey Paul, 2002, Cambridge University Press
- Bioethics Introduction by Marianne Talbot, 2012, Cambridge University Press
- Ethical Issues in Biotechnology , Richard Sherlock, John D. Morrey, Publication Date: December 2002 | ISBN-10: 0742513777 | ISBN-13: 978-0742513778.
- Law relating to Intellectual Property, Fourth edition, Dr. B.L. Wadehra, Universal law publishing co. pvt. Ltd., 2010, ISBN: 978-81-7534-588-1.

Semester II

MBI 205: Immunoinformatics

Elective Course II **(2 Credits, 1L+1T, 30L)**

Unit 1: Immunology 10L

1.1 Basic concepts of immunology: Antigens, Antibodies, The Compliment System, Antigen- Antibody Reactions, The Major Histocompatibility Complex, allele selection, Antigen Presentation,TAP, T cell receptors.

1.2 Immune Effector Mechanisms: Cytokines and Chemokines, Hypersensitive Reactions. The Immune system in Health and Disease: Autoimmunity, Transplantation Immunology, Classification of Vaccine.

Unit 2: Immunoinformatics 10L

2.1 HLA System in model organism:Defining HLA Supertypes in: HLA–Structural and Modeling Principles, HLA Supertypes by GRID/CPCA and Hierarchical Clustering Methods, Structural basis for HLA-A2 Supertypes, Clustering of MHC Peptide-Binding Repertoires,HLA Alleles-Electrostatic Distribution Maps.

2.2 *In Silico* Prediction of Immunogenicity:

- ➢ Databases Searching: An Introduction, MHCDB(NCBI), IMGT, The IMGT/HLA
- ➢ Database, IPD: The Immuno Polymorphism Database, SYFPEITHI: Database for
- ➢ Searching and T-Cell Epitope Prediction, Mapping of T-Cell Epitopes, MHC
- ➢ Binders,and TAP Binders, Mapping of Linear\B-Cell Epitopes in Bcipep
- ➢ Database, Haptens, Carrier Proteins, and Anti-Hapten Antibodies

Unit 3: Predictions 10L

3.1 Predicting Peptide-MHC Binding: Peptide–MHC Binding Using Profiles

3.2 Machine Learning Techniques for MHC Binders.

3.3 Artificial Intelligence Methods for Predicting T-Cell Epitopes, MHC-Class I and II Binding Affinity. MHC–Molecular Affinity and QSAR Models.

3.4 Support Vector Machine for MHC Binding Peptides, Static Energy Analysis of MHC Class I and Class II Peptide-Binding Affinity.

3.5 Nonlinear Predictive Modeling of MHC Class II–Peptide Binding Using Bayesian, Neural Networks

References:

- Immunoinformatics: bioinformatic strategies for better understanding of immune function, Volume 254, Gregory Bock, Jamie Goode, Wiley, 2003, ISBN0470853565, 9780470853566.
- Immunoinformatics: Predicting Immunogenicity in Silico, Volume 409 of Methods in Molecular Biology, Darren R. Flower, Springer, 2007, ISBN160327118X, 9781603271189.
- Immunobiology: the immune system in health and disease, Charles Janeway, Edition4, Current Biology Publications, ISBN0815332173, 9780815332176.
- Immunology for medical students, Roderick Nairn, Matthew Helbert, Edition4, Mosby, 2002, ISBN0723431906, 9780723431909

Semester II

MBI 206: Software Engineering

Elective Course III **(2 Credits, 1L+1T, 30L)**

Unit 1: **10L**

1.1 Introduction of Software: The evolving Role of Software, Software characteristic, Software Application, Software Crisis, Software Myths.
1.2 Software Engineering Approach: Software Engineering Approach, A Generic View of Software Engineering, Software Process.
1.3 Software Process Models - Waterfall Model, Prototype Model, Incremental Model, Spiral Model, COCOMO Model.

Unit 2: **10L**

2.1 Software Process and Project Metrics: Measures, Metrics and Indicators, Metrics in the Process and Project Domains, Software Measurement, Metrics for Software Quality, Integrating Metrics within the Software Engineering Process.
2.2 Software Project Planning: Project Planning Objectives, Software Scope, Resources, Software Project Estimation, Decomposition Techniques, Empirical Estimation Models, the Make-Buy Decision.
2.3 Risk Management: Software Risks, Risk Identification, Risk Projection, Risk Mitigation, Monitoring, and Management.
2.4 Software Requirement Definition: Software requirement Specification, Formal Specification Techniques, Languages and Processors for Requirement Specification.

Unit 3: **10L**

3.1 Software Design: Fundamental Design Concepts, Modules and Modularization Criteria, Design Notation, Design Techniques, Detailed Design Consideration.
3.2 Implementation Issues: Structured Coding Techniques, Coding Styles, Standards and Guidelines, Documentation Guidelines.
3.3 Verification and Validation: Quality Assurance, Walkthroughs and Inspections, Symbolic Execution, unit testing and Debugging, System Testing, Formal Verification.
3.4 Maintenance: Introduction, Enhancing Maintainability during Development, Configuration Management, Managerial Aspects of Software Maintenance, Source-Code Metrics, Other Maintenance Tools and Techniques.

References:

- Software Engineering R. Pressmen – TMH,7th Ed.
- Software Engineering Concepts :Richard Farley
- Introducing Software Testing Louise Tamres
- Effective Methods for software Testing William Perry, Wiley Pub,3rd Ed.
- Software Testing in Real World Edward Kit, Pearson Pub.
- Software Testing Techniques Boris Beizer, dreamTech pub,2nd Ed.
- Software Testing By Ron Patton, TechMedia Pub.

Websites:

- www.effectivesoft.com
- www.sei.cmu.edu
- www.softwarerisk.com
- www.iist.org

Semester II

MBI 206: Emerging Trends in Information Technology

Elective Course III **(2 Credits, 1L+1T, 30L)**

Unit 1: **10L**

1.1 Social Networking: Definition, Overview of Social Networking Sites, Types of Social Networking Sites: General purpose, Niche Advantages of Social Networking Sites, Drawbacks of Social Networking Sites, Features and Need of Social Networking, Security Issues with Social Networking Sites, Examples

1.2 Cloud Computing: Definition, Cloud Architecture, Cloud Storage, Cloud Types: The NIST Model, The Cloud Cube Model, Deployment Models, Service Models.

1.3 Cloud Computing Service Models:
- ➢ Infrastructure as a Service(IaaS)
- ➢ Platform as a Service(PaaS)
- ➢ Software as a Service(SaaS)

1.4 Benefits of Cloud Computing, Disadvantages of Cloud Computing, Cloud Security

Unit 2: **10L**

2.1 Enterprise Content Management: ECM Introduction, Definition, Process, Types of Content, Examples

2.2 Content Management System(CMS) :Overview and examples

2.3 Electronic Document Management(EDM) :Introduction, Need, Examples

2.4 e-Learning: Definition, Introduction, Types of e-Learning:
- ➢ Learner-led e-Learning
- ➢ facilitated e-Learning
- ➢ Instructor-led e-Learning
- ➢ Embedded e-Learning
- ➢ Telemonitoring And e-Coaching

2.5 e-Learning Models:
- ➢ WBT
- ➢ CBT
- ➢ LMS
- ➢ LCMS

2.6 Virtual School Systems

2.7 e-Learning Tools And Technologies:E- Mail,Online Discussion, Chat and Instant Messaging, Voting, Whiteboard, Application Sharing, Conferencing, Online Meeting Tools

2.8 Standards for e-Learning

2.9 Case Study

Unit 3: **10L**

3.1 e-Commerce definition, Models of e- Commerce, Electronic Payment Systems: Credit/Debit Cards, Smart Cards, Paypal, e-Billing,e- Micropayments Point Of Sales System(POS): Meaning.
3.2 Uses m-Commerce: Overview of mobile- Commerce, Attributes of m-Commerce, Drivers of m-Commerce, m-Commerce Security issues
3.3 Mobile ATM(ICICI Bank Case Study) Applications of m-Commerce:
  - Mobile Financial Applications, m-wallet
  - Mobile Shopping
  - Advertising And Content provision

3.4 Case-Study

References:

- E-Learning Tools and Technologies, William Hortan, Katherine Hortan,Wiley Pub.
- Cloud Computing Bible, Barrie Sosinsky,Wiley India pub
- E-Commerce, C.S.V. Murthy,Himalaya Pub.
- E-World (Excel Publications), Arpita Gopal and Chandrani Singh
- Electronic Commerce A Managerial Perspective, Efraim Turban, Pearson Pub.
- Decision Support Systems and Intelligent Systems, Efraim Turban, Jay Aronson, Pearson,7th Ed
- Cloud computing, Michael Miller, Pearson Pub.
- Internet (Use of Search Engines Google & yahoo etc)

Semester II

MBI 207: Java and Bio-java Lab V

Core Course: Practical Course in Java and Bio-java **(2 Credits, 2P=6Hrs.)**

Writing 30 Programs in Java based on following assignment topics. **30P**

- Data types in JAVA
- Arithmetic and Logical operations
- Arrays in JAVA
- Conditional statement and Loops
- Exception Handling
- Inheritance
- Use of Super keyword
- Modifiers, Access Control and Class Design
- Packages
- Interfaces
- Multithreading
- Java Applets
- Event handling
- Simple Animation

References:

- The Complete Reference by Patrick Naughton and Herbert Schildt, McGraw Hill.
- Computer Programming – II, Mumbai University by Bhakti Raul and Vinayaa Rwool, Tech-Max publication.
- Sams Teach Yourself Java in 21 Days (Covering Java 7 and Android) by Rogers Cadenhead, Sams publication.
- Java In a Nutshell by David Flanagan, Oreilly Publications.
- Java Examples in a Nutshell by David Flanagan, Oreilly Publications.

Semester II

MBI 208: Structural Biology and Molecular Modeling Lab VI

Core Course: Practical Course in Structural Biology and Molecular Modeling

(2 Credits, 2P=6 Hrs,)

Structural Biology: **15P**

- **Visualization Tools:** RasMol, Cn3D, SPDBV, Chime, PyMOL
- **Secondary Structure Prediction Tools:** PHYRE, PSI-PRED, PREDIC PROTEIN,APSSP2, YASPIN SECONDARY STRUCTURE PREDICTION TOOLS, JPRED3
- **Structural Databases:** FSSP, VAST, DALI, CATH, SCOP

Molecular Modeling: **15P**

- Conformational Analysis
- BABEL, MOPAC
- Molecular dynamics simulations
- Molecular Dynamics Visualization
- Homology Modeling

References:

- Molecular Modeling Principles and Applications (2nd Ed.) by Andrew R. Leach, 2001, Prentice Hall, USA.
- Principles of Protein Structure by G. E. Schulz, 2009, Springer
- Textbook of Structural Biology by Anders Liljas, Lars Liljas, Jure Piskur and GÃran Lindblom, World Scientific Publishic Co. Pvt. Ltd.
- Structural Bioinformatics, Edited by Jenny Gu, Philip E. Bourne, Wiely Blackwell
- Computational Structural Biology: Methods and Applications by orsten Schwede, Torsten Schwede, Manuel C. Peitsch, 2008, World Scientific Publishing Company.
- Molecular Modeling Of Nucleic Acids, Leontis N.B. and SantaLucia J. Jr. (eds), 1998, American Chemical Society.
- Molecular Modeling of Proteins by Andreas Kukol, 2008, Humana Press.
- Molecular Modeling and Simulation: An Interdisciplinary Guide by Schlick T., 2012, Springer.
- Biological Modeling and Simulation: A Survey of Practical Models, Algorithms, and Numerical Methods (Computational Molecular Biology) by Russell Schwartz, The MIT Press, Cambridge, Massachusetts, London, England

Semester II

MBI 209: Genomics & Proteomics Lab VII

Core Course: Practical Course in Genomics & Proteomics **(2Credits, 2P=6 Hrs.)**

1. DNA cloning (covered in MBI 101)

2. Polymerase chain reaction **1P**

3. Browsing & viewing genome data **2P**

    - Ensembl@EBI
    - MapViewer@NCBI

4. Viewing regions exhibiting Synteny **2P**
    - Genome assembly
    - Genome Annotation:
      Using integrated genome annotation servers such as the server developed at IMTech, Chandigarh (http://imtech.res.in/raghava/gp.html)

5. Gene Prediction and Gene Modelling **5P**
    - Prediction of genes and gene structures (gene modeling) using online (web) servers of different methods tailored for prokaryotic and eukaryotic organisms such as GLIMMER, GeneMark, Grail, GENSCAN etc. Interpretation of results and comparison with known gene models (where available). Evaluation of accuracy of the methods.
    - Prediction of promoters using methods such as Neural Network Promoter Prediction (NNPP) at Berkeley Drosophila Genome Project server, Genome inspector for combined analysis of multiple signals in genomes etc. Using Promoter databases.
    - Prediction of alternate splice sites using methods such as Splice Site Prediction by Neural Network (at Berkeley Drosophila Genome Project server), GenScan, NetGene2 GeneSplicer etc. Prediction of PCR primers using Primer 3, ePCR etc.

6. Functional Genomics **5P**
    - Using primary databases (such as UniProt) and derived databases such as InterPro, PRINTS, BLOCKS, PRODOM, Pfam etc. along with advanced sequence analysis tools such as profiles searches, pattern searches for function annotation of genomic sequences. Validation and versification of results for known case studies
    - Using sequence-based and structure-based Function Annotation Servers such as: ProKnow (http://www.doe-mbi.ucla.edu/Services/ProKnow/)

- Joined Assembly of Function Annotations (JAFA) at http://jafa.burnham.org/learnMore.html etc. which are integrated services for function annotation
- ProFunc (http://www.ebi.ac.uk/thornton-srv/databases/ProFunc)

7. Isolation of Proteins **3P**

8. Affinity chromatography techniques **2P**

9. PAGE for proteome analysis & Image analysis of 2D gels **3P**

10. Mass Spectrometry **DP**

11. Proteomics databases: **3P**

- Trans-Proteomic Pipeline (TPP)
- PeptideProphet
- iProphet
- ProteinProphet
- Xpress & ASAPRatio
- SpectraST
- Corra & PIPE2
- PeptideAtlas & SRMAtlas
- PIPE2, TIQAM, & ATAQS

12. Protein-Protein Interaction Networks, databases and software: **3P**

- DIP (Database of Interacting Proteins)
- PPI Server
- BIND - Biomolecular Interaction Network Database
- PIM –Hybrigenics
- PathCalling Yeast Interaction Database
- MINT - a Molecular Interactions Database
- GRID - The General Repository for Interaction Datasets
- InterPreTS - protein interaction prediction through tertiary structure

References*:*

- Molecular cloning: a laboratory manual, Volume 1, Joseph Sambrook, E. F. Fritsch, Tom Maniatis, Edition2, Cold Spring Harbor Laboratory, ISBN0879693096, 9780879693091.
- Guide to Human Genome Computing by Martin J. Bishop, Academic Press. ISBN 0-12-102051-7.
- From Genome to Therapy: Integrating new technologies with drug development by Novartis Foundation, John Wiley. ISBN 0-471-62744-5.
- Genome mapping and sequencing By Ian Dunham, Horizon, ISBN1-898486-50-6.
- The Genome by Ram S. Verma, VCH, ISBN 1-56081-043-2.
- Bioinformatics - from genomes to drugs (vol. 1), basic technologies (vol.1) by Lengauer, T., Germany, Wiley-VCH, 2002.
- Principles of Genome Analysis And Genomics (3rd Ed.) by Primrose, S.B. & Twyman, R.M., UK. Blackwell Publishing Company, 2003.
- Bioinformatics approach Guide to the analysis of genes and proteins by Andceas Baxevanis and B.F. Francis Ouellettee. John Wiley 2004.
- http://www.proteomecenter.org
- Protein Microarray Technology , Kambhampati, D. (ed) (2004) Front Matter, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG. doi: 10.1002/3527601554.
- Bioinformatics in the Post-Genomic Era: Genome, Transcriptome, Proteome, and Information-Based Medicine, Jeff Augen Addison-Wesley Professional , 2004 ISBN:0321173864.
- Modern Protein Chemistry: Practical Aspects edited by Gary C. Howard, William E. Brown, 2002. CRC Press.
- Fundamentals of Data Mining in Genomics and Proteomics, By Werner Dubitzky, Martin Granzow, Daniel P. Berrar, 2007, Springer Science + Bussiness Media, LLC.
- Protein Arrays, Biochips and Proteomics: The Next Phase of Genomic Discovery edited by Joanna S. Albala, Ian Humphery-Smith, ISBN-0-8247-4212-1, 2003, Marcel Dekker.

Semester II

MBI 210: Perl and Bioperl Programming Lab VIII

Core Course: Practical Course in Perl and Bioperl Programming (2 Credits, 2P=6 Hrs.)

Practical based on following topics

1. Writing programs in Perl based on following assignment topics: **20P**
   - Numbers, Strings
   - Arithmetic and Logical operators
   - Control statements and Loops
   - Scalar variables
   - Working with Arrays
   - Input/output in Perl
   - Subroutine
   - Working with files
   - Regular Expression
   - Pattern Matching
   - Hashes
   - Object-Oriented Programming in Perl
   - Perl Modular Programming
   -
2. Writing programs for following: **10P**
   - Programs for Transcription. (DNA to RNA sequence)
   - Programs for Translation. (RNA to Protein Sequences)
   - Programs for finding Open Reading Frame.
   - Programs for finding consensus Sequence.
   - Programs for finding Triplet Codon
   - Interconversion of different sequence file Format

References:

- Beginning Perl for Bioinformatics by James Tisdall, O-Reilly publication.
- Perl: Complete Reference Perl by Martin C. Brown, McGraw-Hill publication.
- Mastering Perl for Bioinformatics by James D. Tisdall, O-Reilly Publication.
- Teach Yourself Perl 5 in 21 days by David Till, Sams publishing.
- Mastering Algorithms with Perl by Jon Orwant, Jarkko Hietaniemi and John Macdonald, O-Reilly Publication.
- Professional Perl Programming by Peter Wainwrigh, Published by Wrox Press Ltd.
- Beginning Web Development with Perl by Steve Suehring, Apress publication.

Semester II

MBI 211: HTML and XML Lab IX

Elective Lab I: Practical Course in HTML and Related Languages (2 Credits, 2P=6 Hrs.)

Unit 1 **15P**

- Introduction to HTML
- Frames
- Forms
- Tables
- Text formatting
- CSS

Unit 2 **15P**

- Introduction to XML
- Document type definition
- XSD
- XSLT
- DOM

References:

- Learning Web Design- a beginner's guide to HTML, graphics, and beyond by Jennifer Niederst, Publisher: O'reilly digital studio.
- HTML Black Book by Steven Holzner, Publisher: Dreamtech Press.
- Head First HTML with CSS & XHTML, by Elisabeth Robson and Eric Freeman, Publisher: O'reilly Media.
- XML Pocket Reference by Robert Eckstein and Michel Casabianca, Publisher: O'reilly.
- XML Schema by Eric van der Vlist, Publisher: O'reilly.
- Learning XML by Erik T. Ray, Publisher: O'reilly.

Semester II

MBI 211: Introduction to Python Lab IX

Elective Lab I: Practical Course in Introduction to Python (2 Credits, 2P = 6Hrs.)

Unit 1: 15P

- Introduction to Python.
- Using the Python Interpreter
- Invoking the Interpreter
- The Interpreter and Its Environment
- Numbers
- Strings
- Lists

Unit 2: 15P

- First Steps Towards Programming
- Program Organization
- Control Flow Tools
  - if Statements
  - for Statements
  - The range() Function
  - break and continue Statements, and else Clauses on Loops
- Functions
- Modules and Libraries

References:

- Python Programming for the Absolute Beginner by Michael Dawson, Course Technology PTR
- Dive Into Python by Mark Pilgrim, Apress, 2009.
- Learning Python, 4th Edition Powerful Object-Oriented Programming By Mark Lutz, Publisher: O'Reilly Media.
- Learning Python, By Mark Lutz and David Ascher, Publisher: O'Reilly Media.
- http://www.python.org/

SEMESTER III

MBI 301: Scientific Writing Skills

General Course **(2 Credits, 1L+1T, 30L)**

Course Statement/Objectives:

This course aims to demystify the writing process and teach the fundamentals of effective scientific writing. Instruction will focus primarily on the process of writing and publishing scientific manuscripts but grant writing will also be addressed. The course will be presented in two segments:
Part (1) teaches students how to write effectively, concisely, and clearly.
part (2) takes them through the preparation of an actual scientific manuscript or grant.

The course may be taken for 2 units or 3 units. Students taking the class for 2 units will be asked to attend a weekly lecture and to complete short writing and editing exercises, including writing a real "Letter to the Editor" piece. Students taking the class for 3 units will additionally develop a manuscript of their choice and will meet periodically with the instructor for individual editing sessions.

Course Outline

Unit 1: Demystifying the Writing Process 10L

1.1. What makes good writing? Are there "good writers" and "bad writers"?
1.2. Words, word choice, the basic elements of sentences and sentence structure. Writing in the active voice. Introduction to Bioinformatics
1.3. Dissecting the news article:News-writing is the art of maximizing information and minimizing words; it's the barest-bones form of writing. The fundamentals of good writing can be learned by dissecting news articles.
1.4. In-Class Exercise: Sorting through news articles.
1.5. Writing Basics I:Punctuation and Parallelism. Tricks for clarity, brevity, and finesse.
1.6. In-Class Exercise: Peer interviews and write-up mini-profiles
1.7. Writing Basics II:Paragraphs, logic, and organization. Organizational strategies.
1.8. In-Class Exercise: Peer interviews and write-up mini-profiles (swap)
1.9. Writing Basics III:Putting it all together…
1.10. In Class exercise: group rewrites of hard-to-read scientific snippets

Unit 2: Good Writing Applied: The Scientific Manuscript 10L

2.1 Methods and Results Sections: How to present data effectively. How to write prose that complements a table or figure.
2.2 In-Class Exercise: Discuss a variety of journal articles that present data in different ways; rewrite a results paragraph.

2.3 The Abstract, Introduction, and Discussion: Getting to the main point and summarizing effectively. How to conduct literature reviews. Writing an effective discussion.

2.4 Wrap-up scientific manuscripts plus Overview of grant writing:
- Submission and authorship for scientific manuscripts.
- Overview of the NIH grant writing process.

Unit 3: Communicating effectively with the media and lay public and peer review 10L

3.1 How to write articles for the lay public. How to deal with the media. Ethics. Also: how to write a peer review.

3.2 In-Class Exercise: Practice writing a peer review

3.3 SUBMIT completed (or near-completed) manuscript

Exercises:

1 Exercise One:
- Read chapters 1-4 *Sin and Syntax* (pp. 1-87)
- Read Chapter 6 of *Successful Scientific Writing*
- Mini editing exercise 1: revise sentences for clarity and brevity.

1 Exercise Two:
- Exercise: pick a lengthy feature article (3000+ words) from a popular magazine, on any subject (have fun with it!) and rewrite it as a bare-bones news article (500 words maximum for 3-unit students).
- Read chapters 5-8 of *Sin and Syntax* (pp. 88-128)

2 Exercise Three:
- Read chapters 9-10 of *Sin and Syntax* (pp. 129-168)
- Read Chapter 7 of *Successful Scientific Writing*
- Revise news article based on our edits
- Exercise: more sentence re-writing exercises

4. Exercise Four:
- Read chapters 11-12 of *Sin and Syntax* (pp. 169-195)
- Exercise: paragraph re-writing practice
- Choose an article for the “Letter to the Editor” assignment

5. Exercise Five:
- Read chapters 13-16 of *Sin and Syntax* (pp. 197-finish)
- Read Chapter 5 of *Successful Scientific Writing*
- Work on your chosen manuscript, paper, thesis, or grant proposal in preparation for the 2nd half of the course.
- Exercise: Write a real “Letter to the Editor”.

6. Exercise Six:
- Read chapter 3-4 of *Successful Scientific Writing*

- Exercise: Edit a peer's "Letter to the Editor"
- Prepare tables, figures, results, and methods sections.

7. Exercise Seven:

   - Read chapter 2 of *Successful Scientific Writing*
   - Exercise: tables, figures, results, and methods sections (as your data permit)
   - Exercise: Write the introduction or background section for your ongoing manuscript.

8. Exercise Eight:
   - Review chapter 2 of *Successful Scientific Writing*
   - Complete final revisions on your Letter to the Editor
   - Exercise: the introduction or background section of your ongoing manuscript/grant/thesis/paper
   - *****Conference**: Sign up for a time to meet with the instructor for 30-minute individual copy-editing session*. Following these sessions, revise your tables, figures, results, methods, and introduction/background.
   - Exercise: Write the discussion section for your ongoing manuscript.

9. Exercise Nine and Final!:

   - Read chapter 8 of *Successful Scientific Writing*
   - Exercise: final version of your Letter to the Editor; Discussion section of your ongoing manuscript
   - ****Conference**-- *Students will set up times to meet with the instructor for 45-minute individual copy-editing sessions. Revise discussion following sessions + add abstract and assemble manuscript.*
   - ***SUBMIT completed (or near-completed) manuscript.***

References:

1. Sin and Syntax, Constance Hale
2. Successful Scientific Writing: A step-by-step guide for biomedical scientists, Matthews and Bowen

Semester III

MBI 302: Cheminformatics and Drug Design

Advanced Course **(4 Credits, 3L+1T, 60L)**

Unit 1: **10L**

1.1 Introduction to cheminformatics: aims, scope, role of cheminformatics in pharmaceutical/chemical research.
1.2 Chemical Structure representation file format: 1D, 2D and 3D structures, molecular file formats (SMILES, PDF, SDF, MOL).
1.3 Molecular patterns: SMARTS and SMIRKS.
1.4 Molecular Descriptors: 1D, 2D, 3D, topological (Molecular connectivity, electrotoplogical and refractotopological) and shape indices descriptors.

Unit 2: **10L**

2.1 Introduction to graph theory, vertex partitioning algorithms, Molecular Similarity and Molecular Diversity Analysis.
2.2 Similarity metrics: Tanimoto Coefficient, Euclidean distance and Tversky Index.
2.3 Chemical Databases: Design, storage and retrieval methods, molecular database screening, Lipinski rule, Drug/Lead like molecules, chemical structure based search techniques
2.4 Clustering and regression methods: Multiple linear regression (MLR), Partial least square (PLS), Principal component analysis (PCA), k-nearest neighbor (kNN).

Unit 3: **10L**

3.1 Introduction to molecular properties: Hydrophobicity, molecular refractivity, electronic charges (Huckel charge, Wang-ford charge), principal moment of inertia.
3.2 Activities and Toxicities: In-silico ADMET Studies
3.3 Design & Analysis of combinatorial libraries: Molecular scaffolds, linkers and Functional groups, combinatorial library generation

Unit 4: **10L**

4.1 Drug discovery process, role of Bioinformatics in drug design.
4.2 Target identification and validation, lead optimization and validation.
4.3 Structure-based drug design and ligand based drug design.
4.4 Modeling of target-small molecule interactions.

Unit 5: 10L

5.1 Structure Activity Relationship: QSARs and QSPRs, QSAR Methodology.

5.2 Various descriptors used in QSARs: Electronics; Topology; Quantum Chemical based descriptors.

5.3 3D QSAR techniques: CoMFA and CoMSIA.

5.4 Training data, test data and external validation data, applicability domain in QSAR, Cross validation techniques, Pubchem BioAssay data for QSAR studies.

Unit 6: 10L

6.1 Pharmacophore features, Pharmacophore model, Receptor-based and ligand-based pharmacophore modeling.

6.2 Virtual screening based on pharmacophore model.

6.3 Receptor site, molecular docking study, flexible docking, rigid docking, molecular interactions.

6.4 Scoring functions, correlation between ligand-based and receptor-based studies.

References:

- Chemoinformatics -A Textbook by Johann Gasteiger and Thomas Engel, Wiley-VCH 2003.
- Handbook of Chemoinformatics by Johann Gasteiger, Wiley-VCH 2003.
- An Introduction to Chemoinformatics by Andrew R. Leach, Valerie J. Gillet, Springer 2007.
- Burger's Medicinal Chemistry edited by Donald J. Abraham, Volume – I.
- Computer-Aided Molecular Design: Theory and Applications by Jean-Pierre Doucet, J. P. Doucet, Jacques Weber, Elsevier Science & Technology Books.
- Receptor-based Drug Design edited by Paul Leff, Marcel Dekker Inc., New York.
- Advanced Drug Design and Development: A Medicinal Chemistry Approach by P. N. Kourounakis, 1994, Taylor & Francis.
- Biopharmaceutical Drug Design and Development by Susanna Wu-Pong, Yon Rojanasakul, 2008, Humana Press.
- Combinatorial Library Design and Evaluation: Principles, Software, Tools, and Applications in Drug Discovery by Arup Ghose, Vellerkad Viswanadhan, 2001.
- Computer-Aided Drug Design and Delivery Systems by Ahindra Nag, Baishakhi De, 2010, McGraw-Hill Professional.

Semester III

MBI 303: Biological Data Mining

Advanced Course **(4 Credits, 3L+1T, 60L)**

Unit 1: **10L**

1.1 Quality of Biological Data & Data Accuracy:
  - General issues regarding Biological Databases
  - Representation of errors due to machines, 3D structural and sequence data of proteins and nucleic acid Proteomics and Microarray data

1.2 Optimization Techniques
  - Steepest Descent
  - Conjugate Gradient
  - Newton-Raphson
  - Simulated annealing in Biomolecular Structure Optimization

Unit 2: **10L**

2.1 Genetic Algorithms
  - Genetic Algorithms: An Overview
  - Genetic Algorithms in Problem Solving
  - Genetic Algorithms in Scientific Models
  - Theoretical Foundations of Genetic Algorithms
  - Implementing a Genetic Algorithm
  - Conclusions and Future Directions

Unit 3: **10L**

3.1 *Ab initio* methods for structure prediction
  - Lattice, SOM, etc.
  - Information theory, entropy and relative entropy
  - Stochastic Grammars & Linguistics

3.2 Clustering & Classification Algorithms
  - Hierarchical and non-hierarchical Clustering
  - K-Means clustering
  - Grid based clustering
  - Analysis of MD trajectories
  - Microarray and Protein Array data Analysis

Unit 4: 10L

4.1 Dynamic Programming & application in
  - Sequence Alignments
  - Structure Alignments

Unit 5: 10L

5.1 Foundations for Machine learning Techniques
  - Hidden Markov Model
  - Neural Network
  - Bayesian modeling
  - The Cox-Jaynes Axiomes

Unit 6: 10L

6.1 Support Vector machine & Ant colony optimization applied to
  - Multiple Sequence Alignments
  - Biomolecular Structure Prediction

6.2 Fuzzy logic system & application in
  - Clustering and classifications
  - Microarray and Protein Array data Analysis

References:

- Data Mining: Concepts and Techniques by Han and Kamber, Morgan Kaufmann.
- Machine Learning by Tom Mitchell, McGraw Hill.
- An Introduction to Genetic Algorithms, Melanie Mitchell, MIT Press, ISBN: 9780262631853 | 221 pp. | 6.8 x 9.9 in | March 1998
- Biological Data Mining, edited by Jake Y. Chen, Stefano Lonardi,CRC Press, 2010, ISBN:978-1-4200-8684-3
- Data Mining: Practical Machine Learning Tools and Techniques by Witten and Frank, Elsevier.
- Biological Sequence Analysis: probabilistic models of proteins and nucleic acids by Durbin, R., Eddy, S., Krogh, A. & Mitchison, G.Cambridge Univ. Press, 1998.
- Optimization Theory and Application by Rao, S.S., 1984.
- Discrete optimization by Parker, R. G. & Rardin, R. L., 1988.
- Stochastic simulation by Repley, Brian D, Wiley series, 1987.
- Methods of microarray data analysis III by Johnson, K.F. & Lin, S.M. Boston. Kluwer academic publishers, 2003.
- Exploration and analysis of DNA microarray and protein array data by Amaratunga, D. & Cabrera, J.New Jersey. John Wiley & Sons Inc., 2004.
- Ant colony optimization by Dorigo, Marco & Stutzle, Thomas New Delhi, Prentice-Hall of India Pvt Ltd, 2004.
- Data mining: introductory and advanced topics by Dunham, M.H.:New Delhi, Pearson Education, 2003.
- An introduction to bioinformatics algorithms by Jones, Neil.C. & Pevzner, Pavel A. New Delhi, Anne Books, 2005.
- Advances In Genetic Programming, Volume 1, Kenneth E. Kinnear, MIT Press, ISBN: 9780262515535, January 2003.
- An Introduction To Fuzzy Sets, Analysis and Design, Witold Pedrycz and Fernando Gomide, MIT Press, ISBN: 9780262161718, April 1998.
- Fuzzy sets and fuzzy logic: theory and applications by Klir, G.J. & Yuan Bo, New Delhi. Printice Hall of India, 2002. 81-203-1136-1.
- Fuzzy and neuro-fuzzy systems in medicine by Teodorrescu, Horia-Nicolai Kandel, Abraham. & Jain, Lakhmi, C New York, CRS Press, 1999.

Semester III

MBI 304: Systems Biology

Advanced Course **(4 Credits, 3L+1T, 60L)**

Unit 1: Biological Systems – their analysis and modeling **10L**

1.1 Systems and modeling, modeling types, model constraints, process of modeling, constraints and analysis categories

1.2 Qualitative modeling, principles, types, Forrester diagrams, formulations, simplification

1.3 Quantitative modeling, categories, formulation, physical process modeling

Unit 2: Models of mathematics in Systems Biology **10L**

2.1 Difference equation based modeling, linear models, non-linear models, discrete models, continuous models, logistic models

2.2 Interaction modeling, predator-prey models, stability, linearization, equilibria

2.3 Reaction modeling, stoichiometry, fundamental subspaces of stoichiometric matrices

Unit 3: Numerical techniques in analyzing bio-models **10L**

3.1 Numerical integration, instability, stiffness

3.2 Numerical differentiation, ODEs, PDEs

3.3 Regressions, parameter estimations, non-linear equations, calibrations

Unit 4: Validation techniques 10L

4.1 Model validation, techniques, and discrimination models

4.2 Implementation of models in real space, acquisition of data, dynamic validations

4.3 System-level validation

Unit 5: Advances in in-silico modeling – (Case Studies) 20L

5.1 Complexity and robustness analysis

5.2 Modularity-based studies

5.3 Machine-learning based modeling techniques

References:

- Elizabeth S. Allman, John A. Rhodes, Mathematical models in Biology: An Introduction, Cambridge University Press.
- Bernhard O. Palsson, Systems Biology: Properties of Reconstructed Networks, Cambridge University Press.
- James W. Haefner, Modeling Biological Systems: Principles and Applications Springer.
- Zoltan Szallasi, Jörg Stelling, Vipul Periwal, System Modeling in Cell Biology: From Concepts to Nuts and Bolts, MIT Press.

Semester III

MBI 305: Recent Trends in Bioinformatics

Elective Course IV **(4 Credits, 3L+1T, 60L)**

Hot Topics of bioinformatics should be covered in this Paper. The topics are subjected to the variation as per the recent trends.

Unit I: Linux: Linux Essentials, Hands-on Exercises **10L**

- Querying Biological Databases with SQL
- An Introduction to GraphPad Prism - statistics and graphing software
- Relational Databases for Biologists, Filemaker Pro and MySQL
- Tips and Tricks with Excel, Demo.xls
- Speed up your research with LSF
- Matlab Statistics Toolbox
- CERF: Electronic Lab Notebook

Unit II: NGS: Applications/Analysis **20L**

- RNA-seq Analysis in Galaxy
- NGS: QC'ing and Mapping Reads
- Analysis of ChIP-seq data in Galaxy
- RNA-seq: A practical guide to the analysis of differential gene expression
- RNA-Seq Methods and Applications
- Analysis of ChIP-seq experiments
- Mapping Next Generation Sequence Reads
- Analysis of next generation sequencing experiments with Galaxy
- An Introduction to Analyzing Big Data using Amazon Web Services

Unit III: Advanced Genomics: **10L**

- Visualizing Genomes: Browsers
- Visualizing Genomes: Circos Plots
- Integrative Genomics Viewer (IGV)
- UCSC Genome Browser
- Juggling Genome Coordinates
- Introduction to ENSEMBL
- Gene list enrichment analysis
- Identifying and displaying differentially expressed genes
- Cool BaRC Web Tools

Unit IV: Microarray Data Analysis: **10L**

- An introduction to R
- An introduction to R Graphics

- Assessing Sequence and Microarray Data Quality: R/Bioconductor short course
- Clustering and Displaying Microarray Data
- Depositing Microarray Data into ArrayExpress
- Enrichment Analysis

Unit V: Specialized Databases: 10L

- Visualizing Networks: Cytoscape
- BIOBASE Knowledge Library & Human Gene Mutation Database
- Creating Phylogenetic Trees with MEGA
- Searching for transcription factor binding sites with TRANSFAC
- GeneGo: Bioinformatics Technology for Systems Biology

References:

- Next-Generation DNA Sequencing Informatics,Stuart M. Brown, New York University School of Medicine *2013 ISBN 978-1-936113-87-3*
- Chip Technology, Volume 77,Jörg D. Hoheisel, Alvis Brazma, illustratedPublisherSpringer, 2002, ISBN3540432159, 9783540432159
- Microarray biochip technology, Mark Schena , illustratedPublisherEaton Pub., 2000, Original fromthe University of California, ISBN1881299376, 9781881299370
- Discovering Statistics Using R, Andy Field, Jeremy Miles, Zoë Field, SAGE, 2012, ISBN1446258467, 9781446258460
- The R Book, Michael J. Crawley, John Wiley & Sons, 2012ISBN1118448944, 9781118448946
- A Guide to MATLAB: For Beginners And Experienced Users, Brian R. Hunt, Ronald L. Lipsman, Jonathan M. Rosenberg, Edition2, illustratedPublisherCambridge University Press, 2006, ISBN1139452533, 9781139452533
- MATLAB Guide, Desmond J. Higham, Nicholas J. Higham, Edition2, illustratedPublisherSIAM, 2005ISBN0898715784, 9780898715781
- MATLAB Primer, , Kermit Sigmon, Timothy A. Davis, Edition7, CRC Press, 2004ISBN1420034952, 9781420034950
- Foundations of Comparative Genomics, Arcady R. Mushegian, Academic Press, 2010ISBN0080546099, 9780080546094
- Essentials of Genomic and Personalized Medicine, Geoffrey S. Ginsburg, Huntington F Willard, Academic Press, 2009ISBN0080958117, 9780080958118
- https://genome.edu.au/wiki/Galaxy_Tutorials
- http://jura.wi.mit.edu/bio/education/hot_topics
- GeneChip® Expression Analysis-Data Analysis Fundamentals (by Affymetrix) http://mmjggl.caltech.edu/microarray/data_analysis_fundamentals_manual.pdf http://www.stat.duke.edu/~mw/ABS04/RefInfo/data_analysis_fundamentals_manual.pdf

Semester III

MBI 305: Commercial Bioinformatics

Elective Course IV **(4 Credits, 3L+1T, 60L)**

Students need to collect information on current Bioinformatics Industries (e.g. working in the following areas) in India and their area of research. If possible they should visit some bioinformatics companies for an exposure to industrial thought & working process.

Unit 1: **10L**

1.1 Definition of Bioinformatics Company
1.2 Genome Technology: High throughput sequencing and assembly
1.3 Microarray and Genome wide expression analysis
1.4 Transcriptomes, Proteome
1.5 Genomics in medicine

Unit 2: **10L**

2.1 Disease Monitoring
2.2 Profile for therapeutic molecular targeting
2.3 Diagnostic drug discovery and genomics
2.4 Pharmacogenomics and its application
2.5 Proteomics in medicine and its applications

Unit 3: **10L**

3.1 Patenting and data generation from patent literature for commercial benefits
3.2 IPR and bioinformatics
3.3 Bioinformatics patents
3.4 Patenting and data generation from patent literature for commercial benefits

Unit 4: **10L**

4.1 Neuro bioinformatics
4.2 Epigenetics
4.3 Advanced Glycobiology
4.4 Nano –biotechnology & bioinformatics Aspects
4.5 System Biology
4.6 Agro bioinformatics
4.7 Metagenomics
4.8 NGS Technology
4.9 Big Data Analytics

Unit 5: **20L**

5.1 Bioinformatics companies working in India & International market with emphasis on their area of interest. e.g. Accelrys, Applied Maths, Astrid Research, BIOBASE, Biomatters Ltd., Biomax Informatics AG bioinformatics services, CLC Bio, DNASTAR, Gene Codes Corporation, Genedata, Geneious GeneTalk, Genomatix, Genostar, Ingenuity Systems, Inte:Ligand, Integromics, Invitrogen, Korea Computer Centre Sinhung Company, MacVector, Medha Innovation and Development, Ocimum Bio Solutions, Omixon BioComputing Solutions, PREMIER Biosoft, QFAB Bioinformatics, Qlucore, Rosetta Biosoftware, Significo Research, SimBioSys, SRA International, Strand Life Sciences, TimeLogic etc.

References:

- Next-Generation DNA Sequencing Informatics,Stuart M. Brown, New York University School of Medicine *2013 ISBN 978-1-936113-87-3*
- Chip Technology, Volume 77,Jörg D. Hoheisel, Alvis Brazma, illustratedPublisherSpringer, 2002, ISBN3540432159, 9783540432159
- Microarray biochip technology, Mark Schena , illustratedPublisherEaton Pub., 2000, Original fromthe University of California, ISBN1881299376, 9781881299370
- Discovering Statistics Using R, Andy Field, Jeremy Miles, Zoë Field, SAGE, 2012, ISBN1446258467, 9781446258460
- The R Book, Michael J. Crawley, John Wiley & Sons, 2012ISBN1118448944, 9781118448946
- A Guide to MATLAB: For Beginners And Experienced Users, Brian R. Hunt, Ronald L. Lipsman, Jonathan M. Rosenberg, Edition2, illustratedPublisherCambridge University Press, 2006, ISBN1139452533, 9781139452533
- MATLAB Guide, Desmond J. Higham, Nicholas J. Higham, Edition2, illustratedPublisherSIAM, 2005ISBN0898715784, 9780898715781
- MATLAB Primer, , Kermit Sigmon, Timothy A. Davis, Edition7, CRC Press, 2004ISBN1420034952, 9781420034950
- Foundations of Comparative Genomics, Arcady R. Mushegian, Academic Press, 2010ISBN0080546099, 9780080546094
- Essentials of Genomic and Personalized Medicine, Geoffrey S. Ginsburg, Huntington F Willard, Academic Press, 2009ISBN0080958117, 9780080958118
- https://genome.edu.au/wiki/Galaxy_Tutorials
- http://jura.wi.mit.edu/bio/education/hot_topics
- GeneChip® Expression Analysis-Data Analysis Fundamentals (by Affymetrix) http://mmjggl.caltech.edu/microarray/data_analysis_fundamentals_manual.pdf http://www.stat.duke.edu/~mw/ABS04/RefInfo/data_analysis_fundamentals_manual.pdf
- http://hari.hyderabadatoz.com/bio/Bioinf_companylist.html
- http://en.wikipedia.org/wiki/List_of_bioinformatics_companies

Semester III

MBI 306: Biodiversity Informatics

Elective Course V **(2 Credits, 1L+1T, 30L)**

Unit 1: **10L**

1.1 Basic Principles of Taxonomy & Phylogeny
1.2 Modern Taxonomical Methods
1.3 Molecular Systematics: Molecular data types, generation and analysis
1.4 Diversity of Life: Genetic, Species, Community and Landscape Diversity
1.5 Species Diversity: Measurement methodologies, Indices
1.6 Trends of Species Diversity: Gradients, Biodiversity Hot Spots

Unit 2: **10L**

2.1 Ecological / Ecosystem diversity
2.2 Urban Biodiversity
2.3 National, Regional and Global Diversity Information Systems and Networks.
2.4 Biodiversity Informatics Databases:

- Species 2000
- Tree of Life
- ATCC
- GBIF
- ICTV
- Species Analyst Collaboration
- Animal Virus information System
- Online Biodiversity and Ecosystem based Databases

Unit 3: **10L**

3.1 Software for identification
3.2 Accessing existing databases on the WWW
3.3 Software for identification of species
3.4 Probabilistic and deterministic identification
3.5 Delta MicroIS, AVIS, ICTV
3.6 Biocomplexity issues in biodiversity
3.7 Need of metadata standards & ontology

References:

- Recknagel, F. 2002 Ecological Informatics: Understanding Ecology by Biologically- Inspired Computation. Springer, New York.
- Phillipson, J. 1972 Ecological Energetics, Edward Arnold.
- Odum, E.P. 1983 Basic Ecology. Saunders International Edition, Japan.
- Atkinson, P.M. and Tate, N.J.(Eds.) 1999 Advances in remote sensing and GIS analysis., Wiley, New York.
- Gunther,O. 1998 Environmental Information Systems. Berlin, New York, Springer.
- Pankhurst, R.J. 1981 Practical taxonomic computing. Cambridge University Press, Cambridge, U.K.
- Judd, W.S., Campbell, C.S., Kellogg, E.A.,Stevens, P.F. and Donoghue, M.J.(2002) Plant Systematics: A Phylogenetic Approach, 2nd Ed., Sinauer Associates, Inc. Publishers, Sunderland, Masssachusetts, U.S.A.
- Nordenstam, B., El Gazaly, G. and Kassas, M. 2000 Plant Systematics for 21st Century. Portland Press Ltd.,London.

Semester III

MBI 306: Artificial Bio-intelligence

Elective Course V **(2 Credits, 1L+1T, 30L)**

Unit 1: **10L**

1.1 **Intelligent Agents:** PEAS Representation for an Agent, Agent Environments, Concept of Rational Agent, Structure of Intelligent agents, Types of Agents.

1.2 **Problem Solving**: Solving problems by searching, - DFS, BFS, A* searches, Performance Evaluation. Use of graphs in bioinformatics

1.3 **Knowledge and Reasoning**: A knowledge Based Agent, Introduction To Logic, First Order Logic: Syntax and Semantics, Inference in First Order Logic, Unification, Forward and backward chaining, Resolution.

Unit 2: **10L**

2.1 **Planning:** Planning problem, Planning with State Space Search, Partial Order Planning, Hierarchical Planning, Conditional Planning.

2.2 **Learning:** Introduction to neural networks, Perceptrons, Multilayer feed forward network, Application of ANN, Reinforcement learning: Passive & Active Reinforcement learning.

2.3 **Genetic Algorithms:**Single-objective genetic algorithms, Multi-objective genetic algorithms, GA – bioinformatics applications genetic programming

Unit 3: **10L**

3.1 **Introduction to Robotics:** Robot Classification, Robot Specification.

3.2 **Direct and Inverse Kinematics:** Coordinate Frames, Rotations, Homogeneous Coordinates, Arm Equation of four axis SCARA Robot, TCV, Direct Kinematics of Four Axis SCARA Robot Inverse Kinematics of Four Axis SCARA Robot

3.3 **Application**: DNA sequencing, hybridization, , RNA folding prediction using GA., genomics and proteomics, MAGE (*MicroArray* and *Gene Expression)*, Advances in telesurgery and surgical robotics.

References:

- Stuart Russell and Peter Norvig Artificial Intelligence A Modern Approach Edition, Pearson Publication.
- Robert J. Schilling, Fundamentals of Robotics Analysis and Control, PHI Publication.
- Intelligent Bioinformatics: The Application of Artificial Intelligence Techniques to Bioinformatics Problems Edward Keedwell, Ajit Narayanan, John Wiley & Sons Ltd,
- A Beginner's Guide to Microarrays: Eric M. Blalock, Springer
- Artificial Intelligence and Bioinformatics group(AIBIG), Afnizanfaizal Abdullah Springer
- Surgical Robotics Systems Applications and Visions : Rosen, Jacob; Hannaford, Blake; Satava, Richard M. (Eds.) Springer
- Robotic Surgery: Farid Gharagozloo, Farzad Najam, McGraw-Hill Professional Publishing
- Robin R Murphy Introduction to AI Robotics ISBN-81-203-2458-7 PHI Publication.

Semester III

MBI 307: Cheminformatics and Drug Design Lab X

Advanced Course: Practical Course in Cheminformatics and Drug Design

(2 Credits, 2P=6 Hrs.)

Cheminformatics: 15P

- Practical session for chemical structure representation and storage in special file formats.
- Importance of 3D structures and method of generation from 1D & 2D representations.
- A brief introduction to building molecular databases with special emphasis on retrieval using structure input.
- Substructure/Exact/similar structure based searching.
- Development of descriptors (2D descriptors, Refracto-topological and Electro-topological state indices, 3D descriptors).

Drug Design: 15P

- Development of 2D QSAR models.
  - ➢ Classical QSAR
  - ➢ kNN QSAR
  - ➢ GFA QSAR
- Development of 3D QSAR models.
  - ➢ CoMFA
  - ➢ CoMSIA
  - ➢ Topomar CoMFA
  - ➢ HQSAR
- Web-based pharmacophore modeling.
  - ➢ HypoGen
  - ➢ HipHop
  - ➢ DiscoTech
- Molecular docking study.
  - ➢ Rigid Docking
  - ➢ Flexible Docking
  - ➢ Self Docking
- Virtual Screening.
- Combinatorial library design.

References:

- Chemoinformatics -A Textbook by Johann Gasteiger and Thomas Engel, Wiley-VCH 2003.
- Handbook of Chemoinformatics by Johann Gasteiger, Wiley-VCH 2003.
- An Introduction to Chemoinformatics by Andrew R. Leach, Valerie J. Gillet, Springer 2007.
- Burger's Medicinal Chemistry edited by Donald J. Abraham, Volume – I.
- Computer-Aided Molecular Design: Theory and Applications by Jean-Pierre Doucet, J. P. Doucet, Jacques Weber, Elsevier Science & Technology Books.
- Receptor-based Drug Design edited by Paul Leff, Marcel Dekker Inc., New York.
- Advanced Drug Design and Development: A Medicinal Chemistry Approach by P. N. Kourounakis, 1994, Taylor & Francis.
- Biopharmaceutical Drug Design and Development by Susanna Wu-Pong, Yon Rojanasakul, 2008, Humana Press.
- Combinatorial Library Design and Evaluation: Principles, Software, Tools, and Applications in Drug Discovery by Arup Ghose, Vellerkad Viswanadhan, 2001.
- Computer-Aided Drug Design and Delivery Systems by Ahindra Nag, Baishakhi De, 2010, McGraw-Hill Professional.

Semester III

MBI 308: Biological Data Mining Lab XI

Advanced Course: Practical Course in Biological Data Mining (2 Credits, 2P=6Hrs.)

- **Neural networks** **10P**

- Use of neural network tools like BrainBox, MATLAB etc.
- Program in 'C' for implementing the perception learning rule and the delta learning rule (adaline network) for say logic gate patterns.
- Implementation of the Backpropagation learning algorithm for the XOR gate and for an application such as protein secondary structure prediction, promoter prediction etc.
- Estimator of transition probabilities for markov models based on various sample sizes.

- **Hidden Markov model implementation in C/ Java.** **10P**

- **Ant colony algorithm for the Travelling salesman problem (TSP) & Implementation of SVM** **05P**

- **Optimization Algorithms** **05P**

- Programming in 'C' for implementation of Golden section algorithm, Steepest descent, Newton Raphson, Conjugate gradient etc for energy minimization applications.
- Implementation of random walk and Monte Carlo algorithm.
- Simulated Annealing algorithm for energy minimization.

References:

- Data Mining: Concepts and Techniques by Han and Kamber, Morgan Kaufmann.
- Machine Learning by Tom Mitchell, McGraw Hill.
- Data Mining: Practical Machine Learning Tools and Techniques by Witten and Frank, Elsevier.
- Biological Sequence Analysis: probabilistic models of proteins and nucleic acids by Durbin, R., Eddy, S., Krogh, A. & Mitchison, G.Cambridge Univ. Press, 1998.
- Optimization Theory and Application by Rao, S.S., 1984.
- Discrete optimization by Parker, R. G. & Rardin, R. L., 1988.
- Stochastic simulation by Repley, Brian D, Wiley series, 1987.
- Methods of microarray data analysis III by Johnson, K.F. & Lin, S.M. Boston. Kluwer academic publishers, 2003.
- Exploration and analysis of DNA microarray and protein array data by Amaratunga, D. & Cabrera, J.New Jersey. John Wiley & Sons Inc., 2004.
- Ant colony optimization by Dorigo, Marco & Stutzle, Thomas New Delhi, Prentice-Hall of India Pvt Ltd, 2004.
- Data mining: introductory and advanced topics by Dunham, M.H.:New Delhi, Pearson Education, 2003.
- An introduction to bioinformatics algorithms by Jones, Neil.C. & Pevzner, Pavel A. New Delhi, Anne Books, 2005.
- Fuzzy sets and fuzzy logic: theory and applications by Klir, G.J. & Yuan Bo, New Delhi. Printice Hall of India, 2002. 81-203-1136-1.
- Fuzzy and neuro-fuzzy systems in medicine by Teodorrescu, Horia-Nicolai
  Kandel, Abraham. & Jain, Lakhmi, C New York, CRS Press, 1999.

Semester III

MBI 309: System Biology Lab XII

MBI 313: Practical Course in System Biology **(2 Credits, 2P=6Hrs.)**

Practical approach of system biology will be executed through these practical. **30 P**

- Qualitative modeling using dynamic model editors.
- Adaptive modeling using ODE tools.
- Numerical methods implementation.

References:

- Elizabeth S. Allman, John A. Rhodes, Mathematical models in Biology: An Introduction, Cambridge University Press.
- Bernhard O. Palsson, Systems Biology: Properties of Reconstructed Networks, Cambridge University Press.
- James W. Haefner, Modeling Biological Systems: Principles and Applications Springer.
- Zoltan Szallasi, Jörg Stelling, Vipul Periwal, System Modeling in Cell Biology: From Concepts to Nuts and Bolts, MIT Press.

Semester III

MBI 310: Recent trends in Bioinformatics Lab XIII

Elective Lab II: Practical Course in Recent trends in Bioinformatics

(2 Credits, 2P=6Hrs.)

- **Linux:** Advanced: Linux Essentials, Hands-on Exercises **6P**

- **NGS: Applications/Analysis** **6P**

  - RNA-seq Analysis in Galaxy
  - NGS: QC'ing and Mapping Reads
  - Analysis of ChIP-seq data in Galaxy
  - RNA-seq: A practical guide to the analysis of differential gene expression
  - RNA-Seq Methods and Applications
  - Analysis of ChIP-seq experiments
  - Mapping Next Generation Sequence Reads
  - Analysis of next generation sequencing experiments with Galaxy
  - An Introduction to Analyzing Big Data using Amazon Web Services

- **Advanced Genomics:** **6P**

  - Visualizing Genomes: Browsers
  - Visualizing Genomes: Circos Plots
  - Integrative Genomics Viewer (IGV)
  - UCSC Genome Browser
  - Juggling Genome Coordinates
  - Introduction to ENSEMBL
  - Gene list enrichment analysis
  - Identifying and displaying differentially expressed genes
  - Cool BaRC Web Tools

- **Microarray Data Analysis:** **6P**

  - An introduction to R
  - An introduction to R Graphics
  - Assessing Sequence and Microarray Data Quality: R/Bioconductor short course
  - Clustering and Displaying Microarray Data
  - Depositing Microarray Data into ArrayExpress
  - Enrichment Analysis

- **Specialized Databases:** **5P**

- ➢ Visualizing Networks: Cytoscape
- ➢ BIOBASE Knowledge Library & Human Gene Mutation Database
- ➢ Creating Phylogenetic Trees with MEGA
- ➢ Searching for transcription factor binding sites with TRANSFAC
- ➢ GeneGo: Bioinformatics Technology for Systems Biology

- **Industrial Visit** **1P**

References:

- Next-Generation DNA Sequencing Informatics,Stuart M. Brown, New York University School of Medicine *2013 ISBN 978-1-936113-87-3*
- Chip Technology, Volume 77,Jörg D. Hoheisel, Alvis Brazma, illustratedPublisherSpringer, 2002, ISBN3540432159, 9783540432159
- Microarray biochip technology, Mark Schena , illustratedPublisherEaton Pub., 2000, Original fromthe University of California, ISBN1881299376, 9781881299370
- Discovering Statistics Using R, Andy Field, Jeremy Miles, Zoë Field, SAGE, 2012, ISBN1446258467, 9781446258460
- The R Book, Michael J. Crawley, John Wiley & Sons, 2012ISBN1118448944, 9781118448946
- A Guide to MATLAB: For Beginners And Experienced Users, Brian R. Hunt, Ronald L. Lipsman, Jonathan M. Rosenberg, Edition2, illustratedPublisherCambridge University Press, 2006, ISBN1139452533, 9781139452533
- MATLAB Guide, Desmond J. Higham, Nicholas J. Higham, Edition2, illustratedPublisherSIAM, 2005ISBN0898715784, 9780898715781
- MATLAB Primer, , Kermit Sigmon, Timothy A. Davis, Edition7, CRC Press, 2004ISBN1420034952, 9781420034950
- Foundations of Comparative Genomics, Arcady R. Mushegian, Academic Press, 2010ISBN0080546099, 9780080546094
- Essentials of Genomic and Personalized Medicine, Geoffrey S. Ginsburg, Huntington F Willard, Academic Press, 2009ISBN0080958117, 9780080958118
- https://genome.edu.au/wiki/Galaxy_Tutorials
- http://jura.wi.mit.edu/bio/education/hot_topics
- GeneChip® Expression Analysis-Data Analysis Fundamentals (by Affymetrix) http://mmjggl.caltech.edu/microarray/data_analysis_fundamentals_manual.pdf
- http://www.stat.duke.edu/~mw/ABS04/RefInfo/data_analysis_fundamentals_manual.pdf

Semester III

MBI 310: Commercial Bioinformatics Lab XIII

Elective Lab II: Practical Course in Commercial Bioinformatics

(2 Credits, 2P=6Hrs.)

30 P will be conducted based on the theory portion.

References:

- Discovering Statistics Using R, Andy Field, Jeremy Miles, Zoë Field, SAGE, 2012, ISBN1446258467, 9781446258460
- The R Book, Michael J. Crawley, John Wiley & Sons, 2012ISBN1118448944, 9781118448946
- A Guide to MATLAB: For Beginners And Experienced Users, Brian R. Hunt, Ronald L. Lipsman, Jonathan M. Rosenberg, Edition2, illustratedPublisherCambridge University Press, 2006, ISBN1139452533, 9781139452533
- MATLAB Guide, Desmond J. Higham, Nicholas J. Higham, Edition2, illustratedPublisherSIAM, 2005ISBN0898715784, 9780898715781
- MATLAB Primer, , Kermit Sigmon, Timothy A. Davis, Edition7, CRC Press, 2004ISBN1420034952, 9781420034950
- Foundations of Comparative Genomics, Arcady R. Mushegian, Academic Press, 2010ISBN0080546099, 9780080546094
- Essentials of Genomic and Personalized Medicine, Geoffrey S. Ginsburg, Huntington F Willard, Academic Press, 2009ISBN0080958117, 9780080958118
- https://genome.edu.au/wiki/Galaxy_Tutorials
- http://jura.wi.mit.edu/bio/education/hot_topics
- GeneChip® Expression Analysis-Data Analysis Fundamentals (by Affymetrix)
  http://mmjggl.caltech.edu/microarray/data_analysis_fundamentals_manual.pdf
  http://www.stat.duke.edu/~mw/ABS04/RefInfo/data_analysis_fundamentals_manual.pdf
- http://hari.hyderabadatoz.com/bio/Bioinf_companylist.html
- http://en.wikipedia.org/wiki/List_of_bioinformatics_companies

Semester VI

MBI 401: Project

Project **(12 Credits, 2P= 6Hrs.)**

1. Student shall register for the project work in third semester with any of the teaching staff members of RGITBT or may register with external guide from the other recognized institutions imparting post graduate teaching and/or research in life sciences. The registration system involves identifying and finalizing the topic for project work should essentially have 60% bioinformatics component.

2. Student has to initiate literature search work in the third semester only.

3. Student has himself is required to identify the suitable guide and obtain necessary permissions for the project work. Necessary recommendation letter will be issued to the student incase he/she is desirous of working with external guide.

4. There shall be proper format for the final preparation of the project is indicated below:

   - Origin of the problem to be researched
   - Introduction
   - Survey of the literature for last five years
   - Material and methods with detailed description.
   - Results (Data must be analysed with suitable statistical methods to substantiate the results)
   - Discussion
   - References.

5. Student shall make four copies of the project dissertation and submit the same on the declared due date.

6. Student has to give presentation of the project at the end of the fourth semester. Each unit of the project dissertation carries specific marks as indicated in the assessment scheme.

7. 600 marks are allotted for the assessment of this dissertation work. The distribution of these marks is as follows;

i. 200 Marks for updated literature survey and framing of Aims & Objectives of the selected dissertation topic.
ii. 200 Marks for progress of work (based on day – to – day work)
iii. 100 Marks for fulfillment of objectives
iv. 100 Marks for presentation/ Viva (External evaluation)
