Courses in Bioinformatics to enhance Entrepreneurship/ communication skill development

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	2

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.

12.	Queues	2

Concept, Realization of Queues Using Arrays , Circular Queue, Advantages of using circular queues, operations on queue

13.	Sorting and Searching	2

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.

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 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:
    - o Table space
    - o Data files
    - o Blocks
    - o Extents
    - o Segments
    - o Oracle Background Processes
    - o Control files
    - o Oracle Memory Management
    - o Rollback
    - o 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.   ➤ CRVD commands   **10**
- ➤ 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 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   2

- ➤ DO loop processing
- ➤ conditional DO loop processing
- ➤ SAS array processing
- ➤ using SAS arrays

9. Restuction / Rotating SAS Data Set   1

10. Creating Summary Reports   2

- ➤ Proc Print
- ➤ Proc Freq
- ➤ Proc Report
- ➤ Proc Tabulate
- ➤ Report Enhancement

UNIT II

11. SAS Macros Language   1

- ➤ Purpose of Macro Facility
- ➤ Program Flow

12. Macro Variables introduction to macro variables   2

- ➤ Automatic macro variables
- ➤ Macro variable references
- ➤ User-defined macro variables
- ➤ Delimiting macro variable references

13. Macro Definitions defining and calling a macro   2

- ➤ Macro parameters
- ➤ DATA Step and SQL Interfaces creating macro variables in the DATA step

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 208: Structural Biology & Molecular Modeling   **Total**

Core Course –Theory; 3 Credits   **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 302: Chemoinformatics and Drug Designing   **Total**

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

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.

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 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 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 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,   2
hypersensitivity, antagonism, synergism, potentiation, tachyphylaxis,
adverse drug reactions and its monitoring.

3.   Clinical evaluation of new drugs, organizations, types of clinical research   2

4.   Phases of clinical research: pre-clinical studies, human volunteer studies   2
(Phase I), clinical trials (Phase II and III studies), post-marketing
surveillance (Phase IV studies), prescription-event monitoring and
registries.

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,
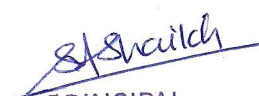randomization blinding and controls.

UNIT II

8.	**Concept of Pharmacovigilance**	2
	Standard terms and terminologies in Pharmacovigilance.	
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.
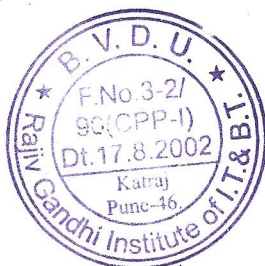
PRINCIPAL
Bharati Vidyapeeth
(Deemed to be University)
Rajiv Gandhi Institute of I.T. & B.T.
Pune Satara Road, Pune - 411 046.