Course Code	BIO543
Course Name	Big Data Mining in Healthcare
Credits	4
Course Offered to	UG/PG
Course Description	There is a exponential growth of data in the field of biological, medical and clinical scineces. Aim of this course is to taught implementation of data mining techniques in healthcare, to solve health-related problems. In this course, students will learn techniques for managing and mining big data. It will be broadly divided in four parts; first part will cover various repositories or databases in the field of medical and biological data. In second part, students will be introduced with techniques comonly used to analyze and manage big data. Implementation of techniques using Python and R will be covered in third phase of this course. Finally, students will learn how to solve health-related problems using knowldge based discovery approach.

Pre-requisites		
Pre-requisite (Mandatory)	Pre-requisite (Desirable)	Pre-requisite(other)
None	None	Basic knowledge of statistics and programming

*Please insert more rows if required

Post Conditions*(For suggestions on verbs please refer the second sheet)			
CO1	**CO2**	**CO3**	**CO4**
Major source of medical and biological data and its importance in healthcare.	Managing, mining and annotating big data in genomic medicine.	Implementation of data mining techniques using R and Python.	Software for desigining drug, vaccine and disease biomarkers from genomic data

Weekly Lecture Plan			
Week Number	**Lecture Topic**	**COs Met**	**Assignment/Labs/Tutorial**
1	Introduction to genomics and proteomics	CO1	
2	Advances in the field of sequencing/expression (NGS, Single Cell Sequencing, RNAseq, Microarray)	CO1	
3	Cancer genomics, genetic biomarkers to detect cancer and its stage	CO1, CO2	
4	Challenges in managing and annotating biomedical data	CO2	
5	SQL and non-SQL techniques for managing big data	CO2	
6	Creation of databases/repositories to manage medical and clinical data	CO2	
7	Introduction to datamining techniques	CO2,CO3	
8	Implementation of datamining tools using Python	CO2, CO3	
9	Statistical analysis of biological data using R/Bioconductor	CO1,CO3	
10	Disease forcasting systems	CO4	

11	Development of models for designing disease biomarkers	CO2, CO3, CO4	
12	Case study of research papers based on datamining	CO1, CO2, CO3, CO4	
13	Project assignment to solve real-life problems in health	CO1, CO2, CO3, CO4	

*Please insert more rows if required

Assessment Plan	
Type of Evaluation	**% Contribution in Grade**
Mid-sem	2000%
End-sem	2000%
Project	3000%
Assignment	2000%
Paper presentation	1000%
*Please insert more row for other type of Evaluation	

Resource Material	
Type	**Title**
	Methods Mol Biol. 409
	Mining of Massive Datasets
Textbook	Handbook of Medical and Healthcare Technologies
Textbook	Big Data Analysis for Bioinformatics and Biomedical Discoveries
Textbook	Big Data Analytics in Genomics