



**JAI HIND COLLEGE
BASANTSING INSTITUTE OF SCIENCE
&
J. T. LALVANI COLLEGE OF COMMERCE
(AUTONOMOUS)**

"A" Road, Churchgate, Mumbai - 400 020, India.


**Affiliated to
University of Mumbai**

Program: M.Sc. Big Data Analytics

Semester II

**Credit Based Semester and Grading System (CBSGS) with
effect from the Academic year 2020-21**




**PRINCIPAL
JAI HIND COLLEGE**

M.Sc. Big Data Analytics Sem II Syllabus

| | |
|--------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Objectives: | <ul style="list-style-type: none">• To understand the functionality of the various data mining and data warehousing component• To pre-process and analyze the characteristics of different types of standard data,• To identify the appropriate statistical analysis technique for a business problem,• To understand fundamentals of Graph data and apply them to Relational Databases.• To recognize explicit and tacit assumptions and their consequences. |
| Outcomes: | <ul style="list-style-type: none">• Compare different approaches of data ware housing and data mining with various technologies.• Build skills to implement different classification and clustering techniques as per requirement to extract valuable information from any type of data set.• Develop novel solutions to identify significant features in data e.g., identify the feedback of potential buyers over online markets to increase the popularity of different products.• Implement the statistical method using R and Excel.• Understanding the concept of dimensions reduction in big data• Able to identify problems, and explain, analyze, and evaluate various cloud computing solutions.• Identify, evaluate and synthesize information (obtained through library, world-wide web, and other sources as appropriate) in a collaborative environment. |




PRINCIPAL
JAI HIND COLLEGE



**JAI HIND COLLEGE
BASANTSING INSTITUTE OF SCIENCE
&
J. T. LALVANI COLLEGE OF COMMERCE
(AUTONOMOUS)**

"A" Road, Churchgate, Mumbai - 400 020, India.

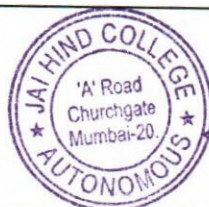
**Affiliated to
University of Mumbai**

Program: M.Sc. Big Data Analytics

Course: Enabling Technologies for Data Science-I

Semester II

**Credit Based Semester and Grading System (CBSGS) with
effect from the Academic year 2020-21**




**PRINCIPAL
JAI HIND COLLEGE**

M.Sc. Enabling Technologies for Data Science-I Syllabus

Semester II

| Course Code | Course Title | Number of Lectures | No. of Credits |
|-------------|------------------------------------------|--------------------|----------------|
| SBDA201 | Enabling Technologies for Data Science-I | 45 | 3 |

DATA MINING:

- a) **Introduction:**
Knowledge discovery from databases, scalability issues.
- b) **Data Warehousing:**
General principles, modeling, design, implementation and optimization, Cloud Computing, OLAP.
- c) **Data Preparation:**
Pre-processing, sub-sampling, feature selection.
- d) **Classification and Prediction:**
Bayes learning, decision trees, CART, neural learning, support vector machines, associations, dependence analysis, rule generation.
- e) **Cluster Analysis and Deviation Detection** Partitioning algorithms, Density bases algorithm, Grid based algorithm, Graph theoretic clustering.
- f) **Temporal and spatial data mining.**

References:

1. Data Mining Techniques: A. K. Pujari, Sangam Books Ltd., 2001
2. Mastering Data Mining: M. Berry and G. Linoff, John Wiley & Sons., 2000
3. Data Mining Cookbook: Modeling Data for Marketing, Risk, and Customer Relationship Management, Olivia Parr Rud, Wiley

Evaluation: Theory: 50% + Practical/Lab: 50%



(Signature)
**PRINCIPAL
JAI HIND COLLEGE**



**JAI HIND COLLEGE
BASANTSING INSTITUTE OF SCIENCE
&
J. T. LALVANI COLLEGE OF COMMERCE
(AUTONOMOUS)**

"A" Road, Churchgate, Mumbai - 400 020, India.

**Affiliated to
University of Mumbai**

Program: M.Sc. Big Data Analytics

Course: Machine Learning - I

Semester II

**Credit Based Semester and Grading System (CBSGS) with
effect from the Academic year 2020-21**





**PRINCIPAL
JAI HIND COLLEGE**

M.Sc. Machine Learning – I Syllabus

Semester II

| Course Code | Course Title | Number of Lectures | No. of Credits |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|--------------------|----------------|
| SBDA202 | Machine Learning - I | 45 | 3 |
| <p>a) Linear Regression Linear Regression with Multiple variables, applications.</p> <p>b) Logistic Regression: Model, Classification, Problem of over-fitting, Applications.</p> <p>c) Neural Networks: Representation Learning, Different Models like single and multi-layer perceptron, back propagation, Application.</p> <p>d) Machine Learning System Design: Evaluating a learning algorithms, handling skewed data, using large data sets.</p> <p>e) Support Vector Machines: Model, Large Margin Classification, Kernels, SVMs in practice.</p> <p>f) Unsupervised Learning.</p> <p>g) Dimensionality Reduction.</p> <p>h) Anomaly Detection.</p> | | | |
| <p>References:</p> <p>1. Machine Learning: Tom Mit chell</p> | | | |
| <p>Evaluation: Theory: 50% + Practical/Lab: 50% (Lab: using R and/or Python)</p> | | | |




**PRINCIPAL
JAI HIND COLLEGE**



**JAI HIND COLLEGE
BASANTSING INSTITUTE OF SCIENCE
&
J. T. LALVANI COLLEGE OF COMMERCE
(AUTONOMOUS)**

"A" Road, Churchgate, Mumbai - 400 020, India.

**Affiliated to
University of Mumbai**

Program: M.Sc. Big Data Analytics

Course: Advanced Statistical Methods

Semester II

**Credit Based Semester and Grading System (CBSGS) with
effect from the Academic year 2020-21**





**PRINCIPAL
JAI HIND COLLEGE**

M.Sc. Advanced Statistical Methods Syllabus

Semester II

| Course Code | Course Title | Number of Lectures | No. of Credits |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|--------------------|----------------|
| SBDA203 | Advanced Statistical Methods | 45 | 3 |
| <p>a) Estimation: Unbiasedness, Consistency, UMVUE, Maximum likelihood estimates.</p> <p>b) Test of Hypotheses: Two types of errors, test statistic, parametric tests for equality of means & variances.</p> <p>c) Linear Model: Gauss Markov Model, least square estimators, Analysis of variance.</p> <p>d) Regression: Multiple linear regression, forward, backward & stepwise regression, Logistic Regression.</p> | | | |
| <p>References:</p> <ol style="list-style-type: none">1. Statistical Inference: P. J. Bickel and K. A. Docksum, 2nd Edition, Prentice Hall.2. Introduction to Linear Regression Analysis: Douglas C. Montgomery | | | |
| <p>Evaluation: Theory: 70% + Practical/Lab: 30% (Lab work: Using R and Python)</p> | | | |




PRINCIPAL
JAI HIND COLLEGE



**JAI HIND COLLEGE
BASANTSING INSTITUTE OF SCIENCE
&
J. T. LALVANI COLLEGE OF COMMERCE
(AUTONOMOUS)**

"A" Road, Churchgate, Mumbai - 400 020, India.

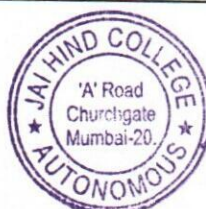
**Affiliated to
University of Mumbai**


Program: M.Sc. Big Data Analytics

Course: Foundations of Data Science

Semester II

**Credit Based Semester and Grading System (CBSGS) with
effect from the Academic year 2020-21**




**PRINCIPAL
JAI HIND COLLEGE**

M.Sc. Foundations of Data Science Syllabus

Semester II

| Course Code | Course Title | Number of Lectures | No. of Credits |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|--------------------|----------------|
| SBDA204 | Foundations of Data Science | 45 | 3 |
| <p>a) Graph Theory: Basic Concepts, Algorithms for connectedness, shortest path, Minimum Spanning Tree, Lab: Graph Databases, Java/Python Programming</p> <p>b) High Dimensional Space: Properties, Law of large numbers, Sphere and cube in high dimension, Generating points on the surface of a sphere, Gaussians in High dimension, Random projection, Applications. Lab: Graph Databases, Java/Python Programming</p> <p>c) Random Graphs: Large graphs, $G(n,p)$ model, Giant Component, Connectivity, Cycles, Non-Uniform models, Applications. Lab: Graph Databases, Java/Python Programming</p> <p>d) Singular Value Decomposition (SVD): Best rank k approximation, Power method for computing the SVD, Applications. Lab: R and Python Programming (Optional: Matlab/Octave)</p> <p>e) Random Walks: Reflection Principle, Long leads, Changes of Sign, Illustrations. Lab: R and Python Programming (Optional: Matlab/Octave)</p> <p>f) Algorithm for Massive Data Problems: Frequency Moments of data streams, matrix algorithms. Lab: R and Python Programming (Optional: Spark, Matlab/Octave)</p> | | | |
| References: 1. Foundations of Data Science: John Hopcroft & Ravindran Kannan. | | | |
| Evaluation: Theory: 40% + Practical/Lab: 60% (GraphDB, R, Python, Java) | | | |




PRINCIPAL
JAI HIND COLLEGE



**JAI HIND COLLEGE
BASANTSING INSTITUTE OF SCIENCE
&
J. T. LALVANI COLLEGE OF COMMERCE
(AUTONOMOUS)**

"A" Road, Churchgate, Mumbai - 400 020, India.

**Affiliated to
University of Mumbai**

Program: M.Sc. Big Data Analytics

Course: Cloud Computing

Semester II

**Credit Based Semester and Grading System (CBSGS) with
effect from the Academic year 2020-21**




**PRINCIPAL
JAI HIND COLLEGE**

M.Sc. Cloud Computing Syllabus

Semester II

| Course Code | Course Title | Number of Lectures | No. of Credits |
|-------------|-----------------|--------------------|----------------|
| SBDA205C | Cloud Computing | 45 | 3 |

Introduction to Cloud computing, Cloud service methods, IaaS, PaaS, SaaS, fundamentals of cloud Architecture (load distribution, resource pooling, scalability, load balancing, redundancy, etc), Introduce DevOps, CICD. Hands-on practice either on AWS/Azure/Google Cloud Platform (Course content to be developed)




PRINCIPAL
JAI HIND COLLEGE



**JAI HIND COLLEGE
BASANTSING INSTITUTE OF SCIENCE
&
J. T. LALVANI COLLEGE OF COMMERCE
(AUTONOMOUS)**

"A" Road, Churchgate, Mumbai - 400 020, India.

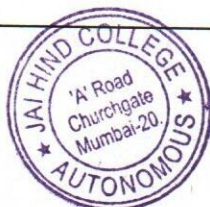
**Affiliated to
University of Mumbai**


Program: M.Sc. Big Data Analytics

Course: Value Thinking

Semester II

**Credit Based Semester and Grading System (CBSGS) with
effect from the Academic year 2020-21**




**PRINCIPAL
JAI HIND COLLEGE**

M.Sc. Value Thinking Syllabus

Semester II

| Course Code | Course Title | Number of Lectures | No. of Credits |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|--------------------|----------------|
| SBDA206 | Value Thinking | 30 | 1 |
| <p>This course involves watching few movies (list provided below) and reading few books (list provided below) that deals mostly with argumentative logic, evidence, drawing inference from evidences. After watching the movies and reading the books, there will be general discussion amongst the students. Couple of case studies that involve mostly logical thinking will also be presented. Each student will prepare a term paper. Evaluation will be on the basis of this term paper and participation in group discussion.</p> <p>Movies:</p> <ol style="list-style-type: none">1. Twelve Angry Men2. Roshoman by Kurosawa3. Trial of Nuremberg4. Mahabharata by Peter Brook <p>Books:</p> <ol style="list-style-type: none">1. The Hound of the Baskervilles by Arthur Conan Doyle2. Five Little Pigs by Agatha Christie3. The Purloined Letter by Edger Allan Poe4. The Case of the Substitute Face <p>Evaluation: Case Studies</p> | | | |




PRINCIPAL
JAI HIND COLLEGE



**JAI HIND COLLEGE
BASANTSING INSTITUTE OF SCIENCE
&
J. T. LALVANI COLLEGE OF COMMERCE
(AUTONOMOUS)**

"A" Road, Churchgate, Mumbai - 400 020, India.

**Affiliated to
University of Mumbai**

Program: M.Sc. Big Data Analytics

Course: Practical-I

Semester II

**Credit Based Semester and Grading System (CBSGS) with
effect from the Academic year 2020-21**




**PRINCIPAL
JAI HIND COLLEGE**


M.Sc. Practical-I Syllabus

Semester II

| Course Code | Course Title | Number of Practical Session | No. of Credits |
|-------------|-------------------------------------------------------------------------------------------------------|-----------------------------|----------------|
| SBDA201PR | Practical-I based on SBDA201(Enabling Technologies for Data Science-I), SBDA202(Machine Learning - I) | 40 | 4 |

- a) Data Warehousing:
General principles, modeling, design, implementation and optimization, Cloud Computing, OLAP.
- b) Data Preparation:
Pre-processing, sub-sampling, feature selection.
- c) Classification and Prediction:**
Bayes learning, decision trees, CART, neural learning, support vector machines, associations, dependence analysis, rule generation.
- d) Cluster Analysis and Deviation Detection:
Partitioning algorithms, Density bases algorithm, Grid based algorithm, Graph theoretic clustering.
- e) Temporal and spatial data mining.
- f) Linear Regression**
Linear Regression with Multiple variables, applications.
- g) Logistic Regression:
Model, Classification, Problem of over-fitting, Applications.
- h) Neural Networks:**
Representation Learning, Different Models like single and multi-layer perceptron, back propagation, Application.
- i) Machine Learning System Design:**
Evaluating a learning algorithm, handling skewed data, using large data sets.
- j) Support Vector Machines:**
Model, Large Margin Classification, Kernels, SVMs in practice.
- k) Unsupervised Learning.
- l) Dimensionality Reduction.
- m) Anomaly Detection.




PRINCIPAL
JAI HIND COLLEGE



**JAI HIND COLLEGE
BASANTSING INSTITUTE OF SCIENCE
&
J. T. LALVANI COLLEGE OF COMMERCE
(AUTONOMOUS)**

"A" Road, Churchgate, Mumbai - 400 020, India.

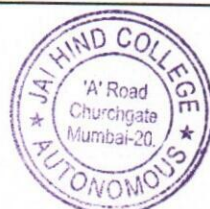
**Affiliated to
University of Mumbai**

Program: M.Sc. Big Data Analytics

Course: Practical-II

Semester II

**Credit Based Semester and Grading System (CBSGS) with
effect from the Academic year 2020-21**




**PRINCIPAL
JAI HIND COLLEGE**

M.Sc. Practical-II Syllabus

Semester II

| Course Code | Course Title | Number of Practical Session | No. of Credits |
|-------------|---------------------------------------------------------------------------------------------------|-----------------------------|----------------|
| SBDA202PR | Practical-II based on SBDA203(Advanced Statistical Methods), SBDA204(Foundations of Data Science) | 40 | 4 |

a) Graph Theory :

Basic Concepts, Algorithms for connectedness, Shortest path, Minimum Spanning Tree, Lab: Graph Databases, Java/Python Programming

b) High Dimensional Space:

Properties, Law of large numbers, Sphere and cube in high dimension, Generating points on the surface of a sphere, Gaussians in High dimension, Random projection, Applications.

Lab: Graph Databases, Java/Python Programming

c) Random Graphs :

Large graphs, $G(n,p)$ model, Giant Component, Connectivity, Cycles, Non-Uniform models, Applications.

Lab: Graph Databases, Java/Python Programming

d) Singular Value Decomposition (SVD):

Best rank k approximation, Power method for computing the SVD, Applications.

Lab: R and Python Programming (Optional: Matlab/Octave)

e) Random Walks:

Reflection Principle, Long leads, Changes of Sign, Illustrations.

Lab: R and Python Programming

f) Algorithm for Massive Data Problems:

Frequency Moments of data streams, matrix algorithms.

Lab: R and Python Programming

g) Estimation:

Unbiasedness, Consistency, UMVUE, Maximum likelihood estimates.

h) Test of Hypotheses:

Two types of errors, test statistic, parametric tests for equality of means & variances.

i) Linear Model:

Gauss Markov Model, least square estimators, Analysis of variance.

j) Regression:

Multiple linear regression, forward, backward & stepwise regression, Logistic Regression.




18
PRINCIPAL
JAI HIND COLLEGE



**JAI HIND COLLEGE
BASANTSING INSTITUTE OF SCIENCE
&
J. T. LALVANI COLLEGE OF COMMERCE
(AUTONOMOUS)**

"A" Road, Churchgate, Mumbai - 400 020, India.

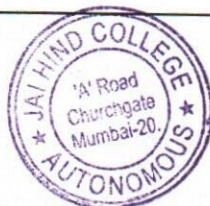
**Affiliated to
University of Mumbai**

Program: M.Sc. Big Data Analytics

Course: Practical-III

Semester II

**Credit Based Semester and Grading System (CBSGS) with
effect from the Academic year 2020-21**




**PRINCIPAL
JAI HIND COLLEGE**

M.Sc. Practical-III Syllabus

Semester II

| Course Code | Course Title | Number of Practical Session | No. of Credits |
|-------------------------------------------------------------|---------------------------------------------------|-----------------------------|----------------|
| SBDA203PR | Practical-III based on SBDA205C (Cloud Computing) | 20 | 2 |
| Hands-on practice either on AWS/Azure/Google Cloud Platform | | | |



Prd
**PRINCIPAL
JAI HIND COLLEGE**