Department of Information Technology (In-house)

Syllabus of B.Sc. in Information Technology (Data Science)

(Effective from academic session 2019-20)

Semester-VI

Name of the Course: B.Sc. in Information Technology (Data Science) **Subject:** Big Data Analytics & Big Data Analytics Lab	
Course Code: BITDS601 & BITDS691	**Semester: VI**
Duration: 36 Hrs	**Maximum Marks:100+100**
Teaching Scheme	**Examination Scheme**
Theory: 3 hrs./week	End Semester Exam:70
Tutorial: 0	Attendance: 5
Practical: 4 hrs./week	Continuous Assessment: 25
Credit: 3+2	Practical Sessional internal continuous evaluation: 40
	Practical Sessional external examination: 60

Aim:

Sl. No.	
1.	Understand big data for business intelligence
2.	Learn business case studies for big data analytics.
3.	Understand nosql big data management.
4.	Perform map-reduce analytics using Hadoop and related tools

Objective:

Sl. No.	
1.	Understand the fundamentals of Big cloud and data architectures.
2.	Understand HDFS file structure and Mapreduce frameworks, and use them to solve complex problems, which require massive computation power
3.	Use relational data in a Hadoop environment, using Hive and Hbase tools of the Hadoop Ecosystem..
4.	Understand the Comparison with traditional databases.

Department of Information Technology (In-house)
Syllabus of B.Sc. in Information Technology (Data Science)
(Effective from academic session 2019-20)

Pre-Requisite:			
Sl. No.			
1.	Database Management Systems.		
2.	Object Oriented Programming Through Java		
Contents		**3 Hrs./week**	
Chapter	**Name of the Topic**	**Hours**	**Marks**
01	**Introduction to big data** Introduction to Big Data Platform – Challenges of Conventional Systems - Intelligent data analysis – Nature of Data - Analytic Processes and Tools - Analysis vs Reporting.	**6**	**10**
02	**Mining data streams** Introduction To Streams Concepts – Stream Data Model and Architecture - Stream Computing - Sampling Data in a Stream – Filtering Streams – Counting Distinct Elements in a Stream – Estimating Moments – Counting Oneness in a Window – Decaying Window - Real time Analytics Platform(RTAP) Applications - Case Studies - Real Time Sentiment Analysis- Stock Market Predictions.	**10**	**20**
03	**Hadoop** History of Hadoop- the Hadoop Distributed File System – Components of Hadoop Analysing the Data with Hadoop- Scaling Out- Hadoop Streaming- Design of HDFS-Java interfaces to HDFS Basics- Developing a Map Reduce Application-How Map Reduce Works-Anatomy of a Map Reduce Job run-Failures-Job Scheduling-Shuffle and Sort – Task execution - Map Reduce Types and Formats-Map Reduce FeaturesHadoop environment.	**12**	**20**
04	**Frameworks** Applications on Big Data Using Pig and Hive – Data processing operators in Pig – Hive services – HiveQL – Querying Data in Hive - fundamentals of HBase and ZooKeeper - IBM InfoSphere BigInsights and Streams. Predictive Analytics- Simple linear regression- Multiple linear regression- Interpretation 5 of regression coefficients. Visualizations - Visual data analysis techniques- interaction	**8**	**20**

	techniques - Systems and applications.		
	Sub Total:	**36**	**70**
	Internal Assessment Examination & Preparation of Semester Examination	**4**	**30**
	Total:	**40**	**100**

Practical:

Skills to be developed:

Intellectual skills:

1. The HDFS file system, MapReduce frameworks are studied in detail.
2. Hadoop tools like Hive, and Hbase, which provide interface to relational databases, are also covered as part of this course work.
3. Ability to implement algorithms to perform various operations on Mapreduce,Pig,Hive

List of Practical:

1. Basic Linux command
2. Installation of Hadoop .
3. Create a directory in HDFS at given path(s).
4. Copy a file from/To Local file system to HDFS
5. Remove a file or directory in HDFS.
6. Display the aggregate length of a file.
7. Word Count Map Reduce program to understand Map Reduce Paradigm
8. Implementing Matrix Multiplication with Hadoop Map Reduce
9. Pig Latin scripts to sort,group, join,project, and filter your data.
10. Hive Databases,Tables,Views,Functions and Indexes

Assignments:

Based on the curriculum as covered by subject teacher.

List of Books

Text Books:

Name of Author	**Title of the Book**	**Edition/ISSN/ISBN**	**Name of the Publisher**
Tom White	Hadoop: The Definitive Guide	Third Edition	O'reilly Media

Chris Eaton, Dirk DeRoos, Tom Deutsch, George Lapis, Paul Zikopoulos	Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data		McGrawHill Publishing
Reference Books:			
Anand Rajaraman and Jeffrey David Ullman	Mining of Massive Datasets		CUP
Bill Franks	Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics		John Wiley& sons
Glenn J. Myatt	Making Sense of Data		John Wiley & Sons
Pete Warden	Big Data Glossary		O'Reilly
List of equipment/apparatus for laboratory experiments:			
Sl. No.			
1.	Computer with moderate configuration		
2.	Linux os or VM		
3.	Hadoop 2.x or higher and other software as required.		

End Semester Examination Scheme. **Maximum Marks-70.** **Time allotted-3hrs.**

Group	**Unit**	**Objective Questions** (MCQ only with the correct answer)		**Subjective Questions**			
		No of question to be set	Total Marks	No of question to be set	To answer	Marks per question	Total Marks
A	**1 to 5**	**10**	**10**				
B	**1 to 5**			**5**	**3**	**5**	**60**

C	**1 to 5**			**5**	**3**	**15**	

- Only multiple choice type question (MCQ) with one correct answer are to be set in the objective part.
- Specific instruction to the students to maintain the order in answering objective questions should be given on top of the question paper.

Examination Scheme for end semester examination:

Group	Chapter	Marks of each question	Question to be set	Question to be answered
A	**All**	**1**	**10**	**10**
B	**All**	**5**	**5**	**3**
C	**All**	**15**	**5**	**3**

Examination Scheme for Practical Sessional examination:

Practical Internal Sessional Continuous Evaluation

Internal Examination:

Continuous evaluation			**40**

External Examination: Examiner-

Signed Lab Note Book	**10**	
On Spot Experiment	**40**	
Viva voce	**10**	**60**

Department of Information Technology (In-house)

Syllabus of B.Sc. in Information Technology (Data Science)

(Effective from academic session 2019-20)

Name of the Course: B.Sc. in Information Technology (Data Science) **Subject:** Computer Vision & Image Processing & Computer Vision & Image Processing Lab	
Course Code: BITDS602 & BITDS692	**Semester: VI**
Duration: 36 Hrs.	**Maximum Marks: 200**
Teaching Scheme	**Examination Scheme**
Theory: 3 hrs./week	End Semester Exam: 70
Tutorial: 0	Attendance : 5
Practical:4 hrs./week	Continuous Assessment:25
Credit: 3+2	Practical Sessional internal continuous evaluation:40
	Practical Sessional external examination:60

Aim:	
Sl. No.	
1.	Students will learn basic principles of image formation, image processing algorithms and different algorithms for reconstruction and recognition from single or multiple images
Objective:	
Sl. No.	
1.	To implement fundamental image processing techniques required for computer vision
2.	Understand Image formation process
3.	Extract features form Images and do analysis of Images
	To develop applications using computer vision techniques
Pre-Requisite:	
Sl. No.	
1.	Programming
2.	Mathematic course

Contents	**Hrs./week**

Department of Information Technology (In-house)

Syllabus of B.Sc. in Information Technology (Data Science)

(Effective from academic session 2019-20)

Chapter	Name of the Topic	Hours	Marks
01	Overview, computer imaging systems, lenses, Image formation and sensing, Image analysis, pre-processing and Binary image analysis	**3**	**10**
02	Edge detection, Edge detection performance, Hough transform, corner detection	**6**	**10**
03	Segmentation, Morphological filtering, Fourier transform	**3**	**10**
04	Feature extraction, shape, histogram, color, spectral, texture, using CVIPtools, Feature analysis, feature vectors, distance /similarity measures, data preprocessing	**9**	**10**
05	**Pattern Analysis:** Clustering: K-Means, K-Medoids, Mixture of Gaussians Classification: Discriminant Function, Supervised, Un-supervised, Semisupervised Classifiers: Bayes, KNN, ANN models; Dimensionality Reduction: PCA, LDA, ICA, and Non-parametric methods.	**9**	**20**
06	Recent trends in Activity Recognition, computational photography, Biometrics	**6**	**10**
	Sub Total:	**36**	**70**
	Internal Assessment Examination & Preparation of Semester Examination	**4**	**30**
	Total:	**40**	**100**

Practical:

Skills to be developed:

Intellectual skills:

1. Ability to pre process the image
2. Ability to image feature identification
3. Can be able to apply recent machine learning methods for different purpose.

List of Practical:

Based on theory Paper

Assignments:

Based on the curriculum as covered by subject teacher.

List of Books

Text Books:

Name of Author	Title of the Book	Edition/ISSN/ISBN	Name of the Publisher
Richard Szeliski	Computer Vision: Algorithms and Applications		
Goodfellow, Bengio, and Courville	Deep Learning		

Reference Books:

Fisher et al	. Dictionary of Computer Vision and Image Processing		

List of equipment/apparatus for laboratory experiments:

Sl. No.	
1.	**Computer**
2.	**Matlab/python/R**

End Semester Examination Scheme. **Maximum Marks-70.** **Time allotted-3hrs.**

Group	Unit	Objective Questions (MCQ only with the correct answer)		Subjective Questions			
		No of question	Total Marks	No of question	To answer	Marks per question	Total Marks

		to be set		to be set			
A	**1 to 6**	**10**					
			10				**60**
B	**1 to 6**			**5**	**3**	**5**	
C	**1 to 6**			**5**	**3**	**15**	

- Only multiple choice type question (MCQ) with one correct answer are to be set in the objective part.
- Specific instruction to the students to maintain the order in answering objective questions should be given on top of the question paper.

Examination Scheme for end semester examination:

Group	Chapter	Marks of each question	Question to be set	Question to be answered
A	**All**	**1**	**10**	**10**
B	**All**	**5**	**5**	**3**
C	**All**	**15**	**3**	**3**

Examination Scheme for Practical Sessional examination:

Practical Internal Sessional Continuous Evaluation

Internal Examination:

Continuous evaluation			**40**

External Examination: Examiner-

Signed Lab Note Book	**10**	
On Spot Experiment	**40**	
Viva voce	**10**	**60**

Department of Information Technology (In-house)
Syllabus of B.Sc. in Information Technology (Data Science)
(Effective from academic session 2019-20)

Name of the Course: B.Sc. in Information Technology (Data Science) **Subject:** Machine Learning for Financial Modelling and Forecasting	
Course Code: BITDS603A	**Semester: VI**
Duration: 36 Hrs.	**Maximum Marks: 100**
Teaching Scheme	**Examination Scheme**
Theory: 3 hrs./week	End Semester Exam: 70
Tutorial: 0	Attendance : 5
Practical:0	Continuous Assessment:25
Credit: 3	Practical Sessional internal continuous evaluation:NA
	Practical Sessional external examination:NA

Aim:

Sl. No.	
1.	Aim of this study to predict supply/demand/inventory of the market, and improve business performance.

Objective:

Sl. No.	
1.	To acquire expertise in the mechanics of the most popular machine learning models, and their inter-relationship, in order to do proper model selection and fitting.
2.	To understand the behavior of financial time series, their statistical properties, and learn to design and assess financial forecasting models and investment strategies based on supervised learning models or other models that use different types (quantitative and qualitative) of information sets.

Pre-Requisite:

Sl. No.	
1.	Foundations of Data Science. Basic Statistics.
2.	Knowledge of R or Python

Contents	**Hrs./week**

Department of Information Technology (In-house)
Syllabus of B.Sc. in Information Technology (Data Science)
(Effective from academic session 2019-20)

Chapter	Name of the Topic	Hours	Marks
01	**Understanding Financial Time Series Data:** Asset's price and return. Basic statistics of returns. Measures of dependence. Stationarity. Forecasting. Volatility. Technical and Fundamental Financial indicators as information set.	**8**	**15**
02	**Financial Time Series Modeling:** Linear regression models and GARCH nonlinear model (quick review). Kernels in Statistical Machine Learning. Support Vector Regression. Neural Networks. Feed-forward networks. Multilayered Networks (Deep Learners). Recurrent networks. LSTM. Data preprocessing and Evaluation of Model Estimation.	**10**	**20**
03	**Optimization** Heuristics in Finance. Random search. Simulated Annealing, Genetic Programming, and other heuristics. Using heuristics for parameter estimation of GARCH, SVM, and Neural networks.	**8**	**15**
04	**Applications** Estimating and Forecasting Financial time series. Algorithmic trading. Porfolio selection. Portfolio optimization under different constraints sets. Credit scoring.	**10**	**20**
	Sub Total:	**36**	**70**
	Internal Assessment Examination & Preparation of Semester Examination	**4**	**30**
	Total:	**40**	**100**

Assignments:

Based on the curriculum as covered by subject teacher.

List of Books

Text Books:

Name of Author	Title of the Book	Edition/ISSN/ISBN	Name of the Publisher

A. Arratia	Computational Finance, An Introductory Course with R		Atlantis Press & Springer, 2014
P. Cortez	Modern Optimization with R	2014	
Reference Books:			
R. Tsay	Analysis of Financial Time Series		Wiley, 2013
Cover, T. A., and Thomas, J. A.,	Elements of Information Theory	Second ed.	(Wiley, 2006).

End Semester Examination Scheme. **Maximum Marks-70.** **Time allotted-3hrs.**

Group	**Unit**	**Objective Questions** (MCQ only with the correct answer)		**Subjective Questions**			
		No of question to be set	Total Marks	No of question to be set	To answer	Marks per question	Total Marks
A	**1 to 5**	**10**	**10**				**60**
B	**1 to 5**			**5**	**3**	**5**	
C	**1 to 5**			**5**	**3**	**15**	

- Only multiple choice type question (MCQ) with one correct answer are to be set in the objective part.
- Specific instruction to the students to maintain the order in answering objective questions should be given on top of the question paper.

Department of Information Technology (In-house)

Syllabus of B.Sc. in Information Technology (Data Science)

(Effective from academic session 2019-20)

Examination Scheme for end semester examination:				
Group	**Chapter**	**Marks of each question**	**Question to be set**	**Question to be answered**
A	**All**	**1**	**10**	**10**
B	**All**	**5**	**5**	**3**
C	**All**	**15**	**3**	**3**

Department of Information Technology (In-house)
Syllabus of B.Sc. in Information Technology (Data Science)
(Effective from academic session 2019-20)

Name of the Course: B.Sc. in Information Technology (Data Science) **Subject:** Applied Machine Learning for Medical Image Analysis			
Course Code: BITDS603B		**Semester: VI**	
Duration: 36 Hrs.		**Maximum Marks: 100**	
Teaching Scheme		**Examination Scheme**	
Theory: 3 hrs./week		End Semester Exam: 70	
Tutorial: 0		Attendance : 5	
Practical:0		Continuous Assessment:25	
Credit: 3		Practical Sessional internal continuous evaluation:NA	
		Practical Sessional external examination:NA	
Aim:			
Sl. No.			
	Familiarity with vision and medical image computing based on machine learning approach.		
Objective:			
Sl. No.			
1.	Each student will gain an understanding of the breadth of methods used in medical image segmentation		
2.	Each student will gain a detailed understanding of one particular approach.		
Pre-Requisite:			
Sl. No.			
1.	Digital image processing		
2.	Mathematical Knowledge		
Contents		**Hrs./week**	
Chapter	**Name of the Topic**	**Hours**	**Marks**
01	**Introduction**	**9**	**20**

	Similarity between images. Image preprocessing. Image matching and registration. Basics. Advanced image registration techniques. Applications of image registration. Evaluating image registration for medical applications.		
02	**Medical Image Segmentation and Applications:** Introduction to Computer Aided Detection (CADe). Image preprocessing. Clustering segmentation techniques. Region-based segmentation in 2D and 3D images. Free-form Segmentation and active contours. Deformable template matching and active shape models. Evaluation of detection algorithms for medical applications	**12**	**25**
03	**Computer Aided Diagnosis:** Introduction to diagnosis and CADx. Object and image characterization. Morphological, texture, and shape descriptors. Interest point detectors and descriptors. Classification and diagnosis. CADx evaluation. Applications through machine learning.	**15**	**25**
	Sub Total:	**36**	**70**
	Internal Assessment Examination & Preparation of Semester Examination	**4**	**30**
	Total:	**40**	**100**

Practical:

Skills to be developed:

Assignments:

Based on the curriculum as covered by subject teacher.

List of Books

Text Books:

Department of Information Technology (In-house)
Syllabus of B.Sc. in Information Technology (Data Science)
(Effective from academic session 2019-20)

Name of Author	Title of the Book	Edition/ISSN/ISBN	Name of the Publisher
Rafael C. Gonzalez	Digital Image Processing Using MATLAB	978-0130085191	
Oleg S. Pianykh (Author)	Digital Imaging and Communications in Medicine (DICOM)	978-3540745709	
Reference Books:			
Barton F. Branstetter	Practical Imaging Informatics: Foundations and Applications for PACS	978-1441904836	
Bettyann H. Kevles	Naked to the Bone		

End Semester Examination Scheme. Maximum Marks-70. Time allotted-3hrs.

Group	Unit	Objective Questions (MCQ only with the correct answer)		Subjective Questions			
		No of question to be set	Total Marks	No of question to be set	To answer	Marks per question	Total Marks
A	**1 to 7**	**10**	**10**				**60**
B	**1 to 7**			**5**	**3**	**5**	
C	**1 to 7**			**5**	**3**	**15**	

- Only multiple choice type question (MCQ) with one correct answer are to be set in the objective part.
- Specific instruction to the students to maintain the order in answering objective questions

should be given on top of the question paper.				
Examination Scheme for end semester examination:				
Group	**Chapter**	**Marks of each question**	**Question to be set**	**Question to be answered**
A	**All**	**1**	**10**	**10**
B	**All**	**5**	**5**	**3**
C	**All**	**15**	**3**	**3**

Department of Information Technology (In-house)

Syllabus of B.Sc. in Information Technology (Data Science)

(Effective from academic session 2019-20)

Name of the Course: B.Sc. in Information Technology (Data Science) **Subject:** Data Warehousing			
Course Code: BITDS603C		**Semester: VI**	
Duration: 36 Hrs.		**Maximum Marks: 100**	
Teaching Scheme		**Examination Scheme**	
Theory: 3 hrs./week		End Semester Exam: 70	
Tutorial: 0		Attendance : 5	
Practical:		Continuous Assessment:25	
Credit: 3		Practical Sessional internal continuous evaluation:NA	
		Practical Sessional external examination:NA	
Aim:			
Sl. No.			
	Understand the components, architecture and other important tools of data warehousing.		
Objective:			
Sl. No.			
1.	Be familiar with the concepts of data warehouse and data mining,		
2.	Be acquainted with the tools and techniques used for Knowledge Discovery in Databases.		
Pre-Requisite:			
Sl. No.			
1	Data Base Management System		
Contents		**Hrs./week**	
Chapter	**Name of the Topic**	**Hours**	**Marks**
01	Introduction to Data Warehousing; Data Mining: Mining frequent patterns, association and correlations; Sequential Pattern Mining concepts, primitives, scalable methods	**3**	**8**

02	Classification and prediction; Cluster Analysis – Types of Data in Cluster Analysis, Partitioning methods, Hierarchical Methods; Transactional Patterns and other temporal based frequent patterns	**6**	**10**
03	Mining Time series Data, Periodicity Analysis for time related sequence data, Trend analysis, Similarity search in Time-series analysis	**6**	**12**
04	Mining Data Streams, Methodologies for stream data processing and stream data systems, Frequent pattern mining in stream data, Sequential Pattern Mining in Data Streams, Classification of dynamic data streams, Class Imbalance Problem; Graph Mining; Social Network Analysis	**7**	**15**
05	Web Mining, Mining the web page layout structure, mining web link structure, mining multimedia data on the web, Automatic classification of web documents and web usage mining; Distributed Data Mining	**8**	**15**
06	Recent trends in Distributed Warehousing and Data Mining, Class Imbalance Problem; Graph Mining; Social Network Analysis	**6**	**10**
	Sub Total:	**36**	**70**
	Internal Assessment Examination & Preparation of Semester Examination	**4**	**30**
	Total:	**40**	**100**

Assignments:

Based on the curriculum as covered by subject teacher.

List of Books

Text Books:

Name of Author	Title of the Book	Edition/ISSN/ISBN	Name of the Publisher
Alex Berson and Stephen J. Smith	Data Warehousing, Data Mining, & OLAP	Second Edition	TataMcGraw Hill Education
Sam Aanhory & Dennis Murray	Data Warehousing in the Real World		Pearson Edn Asia

Reference Books:

Ralph Kimball	Data warehouse Toolkit		Wiley India
Paulraj Ponnaiah Wiley	Data Warehousing Fundamentals		
K.P.Soman,S.Diwakar, V.Ajay	Insight into Data Mining		PHI

End Semester Examination Scheme. Maximum Marks-70. Time allotted-3hrs.

Group	**Unit**	**Objective Questions** (MCQ only with the correct answer)		**Subjective Questions**			
		No of question to be set	Total Marks	No of question to be set	To answer	Marks per question	Total Marks
A	**1 to 6**	**10**	**10**				**60**
B	**1 to 6**			**5**	**3**	**5**	
C	**1 to 6**			**5**	**3**	**15**	

- Only multiple choice type question (MCQ) with one correct answer are to be set in the objective part.
- Specific instruction to the students to maintain the order in answering objective questions should be given on top of the question paper.

Examination Scheme for end semester examination:

Group	**Chapter**	**Marks of each question**	**Question to be set**	**Question to be answered**
A	**All**	**1**	**10**	**10**
B	**All**	**5**	**5**	**3**
C	**All**	**15**	**3**	**3**

Department of Information Technology (In-house)

Syllabus of B.Sc. in Information Technology (Data Science)

(Effective from academic session 2019-20)

Name of the Course: B.Sc. in Information Technology (Data Science) **Subject:** Natural Language Processing	
Course Code: BITDS604	**Semester: VI**
Duration: 48 Hrs.	**Maximum Marks: 100**
Teaching Scheme	**Examination Scheme**
Theory: 3 hrs./week	End Semester Exam: 70
Tutorial: 1 hr./week	Attendance : 5
Practical:0	Continuous Assessment:25
Credit: 4	Practical Sessional internal continuous evaluation:NA
	Practical Sessional external examination:NA

Aim:	
Sl. No.	
1.	Process the text data at syntactic and semantic level.
2.	Extract the ¬key information from Text data.
3.	Analyze the text content to provide predictions related to a specific domain using language models.

Objective:	
Sl. No.	
1.	To get introduced to language processing technologies for processing the text data
2.	To understand the role of Information Retrieval and Information Extraction in Text Analytics.
3.	To acquire knowledge on text data analytics using language models.

Department of Information Technology (In-house)
Syllabus of B.Sc. in Information Technology (Data Science)
(Effective from academic session 2019-20)

Pre-Requisite:			
Sl. No.			
1.	Programming Knowledge		
Contents		**Hrs./week**	
Chapter	**Name of the Topic**	**Hours**	**Marks**
01	**Regular Expressions and Automata Recap-** Introduction to NLP, Regular Expression, Finite State Automata **Tokenization –** Word Tokenization, Normalization, Sentence Segmentation, Named Entity Recognition, Multi Word Extraction, Spell Checking – Bayesian Approach, Minimum Edit Distance **Morphology** – Inflectional and Derivational Morphology, Finite State Morphological Parsing, The Lexicon and Morphotactics, Morphological Parsing with Finite State Transducers, Orthographic Rules and Finite State Transducers, Porter Stemmer	**12**	**20**
02	**Language Modeling** Introduction to N-grams, Chain Rule, Smoothing – Add-One Smoothing, Witten-Bell Discounting; Backoff, Deleted Interpolation, N-grams for Spelling and Word Prediction, Evaluation of language models. **Hidden Markov Models and POS Tagging** Markov Chain, Hidden Markov Models, Forward Algorithm, Viterbi Algorithm, Part of Speech Tagging – Rule based and Machine Learning based approaches, Evaluation.	**12**	**20**
03	**Text Classification** Text Classification, Naïve Bayes' Text Classification, Evaluation, Sentiment Analysis – Opinion Mining and Emotion Analysis, Resources and Techniques. **Context Free Grammar**	**12**	**20**

	Context Free Grammar and Constituency, Some common CFG phenomena for English, Top-Down and Bottom-up parsing, Probabilistic Context Free Grammar, Dependency Parsing		
04	**Computational Lexical Semantics** I ntroduction to Lexical Semantics – Homonymy, Polysemy, Synonymy, Thesaurus – WordNet, Computational Lexical Semantics – Thesaurus based and Distributional Word Similarity **Information Retrieval** Boolean Retrieval, Term-document incidence, The Inverted Index, Query Optimization, Phrase Queries, Ranked Retrieval – Term Frequency – Inverse Document Frequency based ranking, Zone Indexing, Query term proximity, Cosine ranking, Combining different features for ranking, Search Engine Evaluation, Relevance Feedback	**12**	**10**
	Sub Total:	**36**	**70**
	Internal Assessment Examination & Preparation of Semester Examination	**4**	**30**
	Total:	**40**	**100**

Assignments:

Based on the curriculum as covered by subject teacher.

List of Books

Text Books:

Name of Author	Title of the Book	Edition/ISSN/ISBN	Name of the Publisher
Jurafsky and Martin,	Speech and Language Processing		Pearson Education
Manning and Schutze	Foundation of Statistical Natural Language Processing		MIT Press
Reference Books:			
	Multilingual Natural Language Processing Applications from		Bikel, Pearson

	Theory to Practice		
Matthew A. Russell	Mining the Social Web		O'Reilly

End Semester Examination Scheme. Maximum Marks-70. Time allotted-3hrs.

Group	Unit	Objective Questions (MCQ only with the correct answer)		Subjective Questions			
		No of question to be set	Total Marks	No of question to be set	To answer	Marks per question	Total Marks
A	**1 to 7**	**10**	**10**				**60**
B	**1 to 7**			**5**	**3**	**5**	
C	**1 to 7**			**5**	**3**	**15**	

- Only multiple choice type question (MCQ) with one correct answer are to be set in the objective part.
- Specific instruction to the students to maintain the order in answering objective questions should be given on top of the question paper.

Examination Scheme for end semester examination:

Group	Chapter	Marks of each question	Question to be set	Question to be answered
A	**All**	**1**	**10**	**10**
B	**All**	**5**	**5**	**3**
C	**All**	**15**	**3**	**3**

Department of Information Technology (In-house)
Syllabus of B.Sc. in Information Technology (Data Science)
(Effective from academic session 2019-20)

Name of the Course: B.Sc. in Information Technology (Data Science) **Subject:** Grand Viva Voce	
Course Code: BITDS681	**Semester: VI**
Duration: 12Hrs	**Maximum Marks: 100**
Teaching Scheme	**Examination Scheme**
Theory: 0	**End Semester Exam: 100**
Tutorial: 0	**Attendance: 0**
Practical: 2 hrs./week	**Continuous Assessment: 0**
Credit: 1	**Practical Sessional internal continuous evaluation: NA**
	Practical Sessional external examination: NA
Contents	
Students will give a viva from all the subject that they have covered in the course.	

Name of the Course: B.Sc. in Information Technology (Data Science) **Subject:** Major Project -II	
Course Code: BITDS682	**Semester: VI**
Duration: 36 Hrs.	**Maximum Marks: 100**
Teaching Scheme	**Examination Scheme**
Theory: 0	**End Semester Exam: 100**
Tutorial: 0	**Attendance: 0**
Practical: 4 hrs./week	**Continuous Assessment: 0**
Credit: 2	**Practical Sessional internal continuous evaluation: 40**
	Practical Sessional external examination: 60
Contents	
Students will do projects on application areas of latest technologies and current topics of societal relevance.	