

NOIDA INTERNATIONAL UNIVERSITY



SCHOOL OF ENGINEERING & TECHNOLOGY
EVALUATION SCHEME & SYLLABUS

FOR

BACHELOR OF TECHNOLOGY

Computer Science & Engineering/Information Technology

(4th year Syllabus)

(4 Year Course)

W.E.F Session 2018-2019 onwards

SEVENTH SEMESTER

COURSE		Contact Hours/Week			Credit	% of Total Marks				
Code	Course Title	L	T	P		CA	TA	Int. Total	Ext.	Total
	Elective IV	3	0	0	3	20	20	40	60	100
	Elective V	3	0	0	3	20	20	40	60	100
	Elective VI	3	0	0	3	20	20	40	60	100
OEC002	HRD & OB *	3	0	0	3	20	20	40	60	100
PRACTICALS										
PROJ-CS70	Project-II**	0	0	6	3	20	20	40	60	100
Total		12	0	6	15					

EIGHTH SEMESTER

COURSE		Contact Hours/Week			Credit	% of Total Marks				
Code	Course Title	L	T	P		CA	TA	Int. Total	Ext.	Total
OEC003	Cyber Law & Ethics	3	0	0	3	20	20	40	60	100
OEC004	Introduction	3	0	0	3	20	20	40	60	100
PRACTICALS										
PROJ-CS80	Project III**	0	0	12	6	100	100	200	300	500
Total		6	0	12	12					

* HRD & OB- Human Resource Development & Organizational Behavior

** Project Synopsis Seminar

**The marks will be awarded on the basis of Industrial Project Training in 8th semester

DETAILED 4-YEAR CURRICULUM CONTENTS

Undergraduate Degree in Engineering & Technology

BRANCH/COURSE: COMPUTER SCIENCE AND ENGINEERING

AND

INFORMATION TECHNOLOGY

Course Code: PEC-CS-D 702

Course Name: Speech & Natural Language Processing

Course Credit Hour: 3hr

Total Contact Hour: 42hr

Course Objective: This course introduces the fundamental concepts and techniques of natural language processing (NLP). Students will gain an in-depth understanding of the computational properties of natural languages and the commonly used algorithms for processing linguistic information. The course examines NLP models and algorithms using both the traditional symbolic and the more recent statistical approaches.

Course Description: NLP tasks in syntax, semantics, and pragmatics. Applications such as **information** extraction, question answering, and machine translation. The problem of ambiguity. The role of machine learning.

Course Contents:

Module I Introduction: Knowledge in speech and language processing, Ambiguity, Models and Algorithms, Brief History Regular Expressions and Automata: Regular Expressions, Finite-State Automata, Regular Languages and FSA Morphology and Transducers: Inflectional and derivational morphology, finite state morphological parsing, Combining FST Lexicon and rules. Lexicon free FST: Porter Stemmer N-grams: Counting Words in Corpora, SIMPLE (UNSMOOTHED) N-GRAMS, Smoothing, Entropy HMM and Speech Recognition: Speech Recognition Architecture, Overview of HMM, A* decoding .

Module II Word Classes and Part-of-Speech Tagging: English word classes, Targets for English, Part of speech Tagging, Rule Based part of speech Tagging, Transformation Based Tagging. Context Free Grammars for English: Constituency, Context Free rules and Trees, Sentence level construction, The Noun Phrase, Coordination, Agreement, The verb phrase and sub-categorization. Spoken Language Syntax, Grammar Equivalence and Normal form, Finite state context free grammars, Grammar and human processing. Parsing with context free grammars: Parsing as Search, Basic Top down Parser, Problems with basic top-down-parsers, the early Algorithm, Finite state parsing method Features and Unifications: Feature structures, Unification of Features Structures, Features Structures in the grammar, Implementing Unification. Lexicalized and probabilistic parsing: Probabilistic context free grammars, problems with probabilistic context free grammars, probabilistic lexicalized GFG.

Module III Semantics Representing Meaning: Computational Desiderata for representation, Meaning structure of language, First order predicate calculus, linguistically relevant concept, Related Re-presentational approaches, Alternative approaches to meaning. Semantic Analysis: Syntax driven semantic analysis, Attachment of Fragment of English, Integrating semantic analysis with early parser. Robust Semantic Analysis. Lexical Semantics: Relation among lexemes and their senses, Internal Structure of words.

Module IV Pragmatics Discourse: Reference resolution, Text Coherence, Discourse Structure, Psycholinguistics Studies of reference and coherence. Natural Language generation: Introduction to language generation, Architecture for generation, Surface realization, Discourse planning, Macro planning, Lexical selection, evaluating generation systems, generating speech

Course learning outcomes:

After successful completion of this course, student will be able to

1. Understand approaches to syntax and semantics in NLP.
2. Understand approaches to discourse, generation, dialogue and summarization within NLP.
3. Understand current methods for statistical approaches to machine translation.
4. Understand machine learning techniques used in NLP, including hidden Markov models and probabilistic context-free grammars, clustering and unsupervised methods, log-linear and discriminative models, and the EM algorithm as applied within NLP

Text books: _____

1. Speech and Language processing An introduction to Natural Language Processing, Computational Linguistics and speech Recognition by Daniel Jurafsky and James H. Martin (ISBN13: 978-0131873216)
2. 2. Natural Language Processing with Python by Steven Bird, Ewan Klein, Edward Lopper (ISBN13:978-0596516499)

Reference book:

1. Handbook of Natural Language Processing, Second Edition—Nitin Indurkha, Fred J. Damerau, Fred J. Damerau (ISBN13: 978-1420085921)

Online links for study & reference materials:

1. NPTEL

Assessment method :(Continuous Internal Assessment = 40% , Final Examination = 60%)

Assessment -1	- 05%
Assessment-2	- 05%
Assessment-3(Midexam)	- 20%
Assessment-3	- 05%
Assessment-4	- 05%
Total Internal Assessment	- 40%

Course Code: PEC-CS-D 703

Course Name: Data Analytics

Course Credit Hour: 3hr

Total Contact Hour: 42hr

Course Objective:

- To provide an overview of an exciting growing field of big data analytics.
- To introduce the tools required to manage and analyze big data like Hadoop, NoSqlMapReduce.
- To teach the fundamental techniques and principles in achieving big data analytics with scalability and streaming capability.
- To enable students to have skills that will help them to solve complex real-world problems in for decision support.

Course Description:

- This course prepares students to gather, describe, and analyze data, and use advanced statistical tools to make decisions on operations, risk management, finance, marketing, etc.
- Analysis is done targeting economic and financial decisions in complex systems that involve multiple partners. Topics include probability, statistics, hypothesis testing, regression, clustering, decision trees, and forecasting.

Course Contents:

Module 1: Big Data and its Importance– Four V’s of BigData– DriversforBigData–IntroductiontoBigDataAnalytics–BigDataAnalyticsapplications, Hadoop’s Parallel World–Data discovery Opensourcetechnology for Big Data Analytics–cloud and Big Data–PredictiveAnalytics–MobileBusinessIntelligenceandBigData–CrowdSourcing Analytics–Inter-andTrans-FirewallAnalytics-InformationManagement.

Module2: Integratingdisparatedatastores-Mappingdatatotheprogrammingframework- Connecting and extracting data from storage -Transforming data for processing – SubdividingdatainpreparationforHadoopMapReduce, Hadoop Map Reduce-Creating the componentsofHadoop.

Module3: MapReducejobs-Distributingdataprocessingacrossserverfarms-Executing HadoopMapReducejobs-Monitoringtheprogressofjobflows-TheBuildingBlocksof Hadoop Map Reduce -Distinguishing Hadoop daemons-Investigating the HadoopDistributedFileSystem Selectingappropriateexecutionmodes:local,pseudo-distributed,Fullydistributed.

Module 4:Real-TimeArchitecture–OrchestrationandSynthesisUsingAnalyticsEngines– Discovery using Data atRest– Implementation of Big Data Analytics – Big Data Convergence–AnalyticsBusinessMaturityModel,InstallingandRunningPig– Comparison with Databases–Pig Latin UserDefineFunctions–Data Processing Operators–Installing andRunningHive–HiveQL– Tables–QueryingData–User-DefinedFunctions–Oracle BigData.

Course learning outcomes:

- Understand the key issues in big data management and its associated applications in intelligent business and scientific computing.
- Acquire fundamental enabling techniques and scalable algorithms like Hadoop, Map Reduce and NO SQL in big data analytics.
- Interpret business models and scientific computing paradigms, and apply software tools for big data analytics.
- Achieve adequate perspectives of big data analytics in various applications like recommender systems, social media applications etc.

Suggested books:

1. Data Mining and Business Analytics with R, by Johannes Ledolter; Publisher: Wiley (2013), ISBN-13: 978-1118447147;
2. An Introduction to Statistical Learning with Application in R, by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani;Publisher: Springer (2013); ISBN-13: 978-1461471370;

Suggested reference books:

- 1.MichaelMinelli,Michehe Chambers,“Big Data,Big Analytics: Emerging Business IntelligenceandAnalyticTrendsforToday’sBusiness”,1stEdition,AmbigaDhiraj,WielyCIO Series,2013.
- 2.ArvindSathi,“BigDataAnalytics:DisruptiveTechnologiesforChangingtheGame”, 1stEdition,IBMCorporation,2012.
- 3.BillFranks,“TamingtheBigDataTidalWave:Finding Opportunities in Huge Data Streams withAdvancedAnalytics”,1stEdition,WileyandSASBusinessSeries,2012.
- 4.Tom White,“Hadoop:TheDefinitiveGuide”,3rdEdition,O’reilly,2012.

Online links for study & reference materials:

1. https://catalyst.library.jhu.edu/catalog/bib_6591386
2. https://catalyst.library.jhu.edu/catalog/bib_4637122

Assessment method :(Continuous Internal Assessment = 40% , Final Examination = 60%)

Assessment -1	- 05%
Assessment-2	- 05%
Assessment-3(Midexam)	- 20%
Assessment-3	- 05%
Assessment-4	- 05%
Total Internal Assessment	- 40%
