**NATURAL LANGUAGE PROCESSING**

**[As per Choice Based Credit System (CBCS) scheme]**

**(Effective from the academic year 2016 -2017)**

**SEMESTER – VII**

Subject Code 15CS741 IA Marks 20

Number of Lecture Hours/Week 3 Exam Marks 80

Total Number of Lecture Hours 40 Exam Hours 03

**CREDITS – 03**

**Course objectives:** This course will enable students to

Learn the techniques in natural language processing.

Be familiar with the natural language generation.

Be exposed to Text Mining.

Understand the information retrieval techniques

**Module – 1 Teaching**

**Hours**

**Overview and language modeling:** Overview: Origins and challenges of NLP- **8 Hours**

Language and Grammar-Processing Indian Languages- NLP Applications-

Information Retrieval. Language Modeling: Various Grammar- based Language

Models-Statistical Language Model.

**Module – 2**

**Word level and syntactic analysis:** Word Level Analysis: Regular Expressions- **8 Hours**

Finite-State Automata-Morphological Parsing-Spelling Error Detection and

correction-Words and Word classes-Part-of Speech Tagging. Syntactic Analysis:

Context-free Grammar-Constituency- Parsing-Probabilistic Parsing.

**Module – 3**

**Extracting Relations from Text: From Word Sequences to Dependency 8 Hours**

**Paths:**

Introduction, Subsequence Kernels for Relation Extraction, A Dependency-Path

Kernel for Relation Extraction and Experimental Evaluation.

**Mining Diagnostic Text Reports by Learning to Annotate Knowledge Roles:** Introduction, Domain Knowledge and Knowledge Roles, Frame Semantics and Semantic Role Labeling, Learning to Annotate Cases with Knowledge Roles and Evaluations.

**A Case Study in Natural Language Based Web Search:** InFact System Overview, The GlobalSecurity.org Experience. **Module – 4**

**Evaluating Self-Explanations in iSTART: Word Matching, Latent Semantic 8 Hours**

**Analysis, and Topic Models:** Introduction, iSTART: Feedback Systems, iSTART: Evaluation of Feedback Systems,

**Textual Signatures: Identifying Text-Types Using Latent Semantic Analysis to Measure the Cohesion of Text Structures:** Introduction, Cohesion, Coh- Metrix, Approaches to Analyzing Texts, Latent Semantic Analysis, Predictions, Results of Experiments.

**Automatic Document Separation: A Combination of Probabilistic Classification and Finite-State Sequence Modeling:** Introduction, Related Work, Data Preparation, Document Separation as a Sequence Mapping Problem, Results.

**Evolving Explanatory Novel Patterns for Semantically-Based Text Mining:**

Related Work, A Semantically Guided Model for Effective Text Mining.

**Module – 5**

**INFORMATION RETRIEVAL AND LEXICAL RESOURCES:** Information **8 Hours**

Retrieval: Design features of Information Retrieval Systems-Classical, Non classical, Alternative Models of Information Retrieval – valuation Lexical Resources: World Net-Frame Net- Stemmers-POS Tagger- Research Corpora. **Course outcomes:** The students should be able to:

Analyze the natural language text.

Generate the natural language.

Do Text mining.

Apply information retrieval techniques.

**Question paper pattern:**

The question paper will have ten questions.

T here will be 2 questions from each module.

E ach question will have questions covering all the topics under a module.

The students will have to answer 5 full questions, selecting one full question from

each module.

**Text Books:**

1. Tanveer Siddiqui, U.S. Tiwary, “Natural Language Pr ocessing and Information

Retrieval”, Oxford University Press, 2008.

2. Anne Kao and Stephen R. Poteet (Eds), “Natural Lang uageProcessing and Text

Mining”, Springer-Verlag London Limited 2007.

**Reference Books:**

1. Daniel Jurafsky and James H Martin, “Speech and Lan guage Processing: Anintroduction to Natural Language Processing, Computational Linguistics and SpeechRecognition”, 2nd Edition, Prentice Hall, 200 8.

2. James Allen, “Natural Language Understanding”, 2nd edition,

Benjamin/Cummingspublishing company, 1995.

3. Gerald J. Kowalski and Mark.T. Maybury, “Informatio n Storage and Retrieval

systems”, Kluwer academic Publishers, 2000.

**CLOUD COMPUTING AND ITS APPLICATIONS [As per Choice Based Credit System (CBCS) scheme] (Effective from the academic year 2016 -2017) SEMESTER – VII**

Subject Code 15CS742 IA Marks 20

Number of Lecture Hours/Week 3 Exam Marks 80

Total Number of Lecture Hours 40 Exam Hours 03

**CREDITS – 03**

**Course objectives:** This course will enable students to

Explain the fundamentals of cloud computing

Illustrate the cloud application programming and aneka platform

Contrast different cloud platforms used in industry

**Module – 1 Teaching**

**Hours**

Introduction ,Cloud Computing at a Glance, The Vision of Cloud Computing, **8 Hours**

Defining a Cloud, A Closer Look, Cloud Computing Reference Model,

Characteristics and Benefits, Challenges Ahead, Historical Developments,

Distributed Systems, Virtualization, Web 2.0, Service-Oriented Computing, Utility-Oriented Computing, Building Cloud Computing Environments, Application Development, Infrastructure and System Development, Computing Platforms and Technologies, Amazon Web Services (AWS), Google AppEngine, Microsoft Azure, Hadoop, Force.com and Salesforce.com, Manjrasoft Aneka

Virtualization, Introduction, Characteristics of Virtualized, Environments Taxonomy of Virtualization Techniques, Execution Virtualization, Other Types of Virtualization, Virtualization and Cloud Computing, Pros and Cons of Virtualization, Technology Examples Xen: Paravirtualization, VMware: Full Virtualization, Microsoft Hyper-V

**Module – 2**

Cloud Computing Architecture, Introduction, Cloud Reference Model, **8 Hours** Architecture, Infrastructure / Hardware as a Service, Platform as a Service, Software as a Service, Types of Clouds, Public Clouds, Private Clouds, Hybrid Clouds, Community Clouds, Economics of the Cloud, Open Challenges, Cloud Definition, Cloud Interoperability and Standards Scalability and Fault Tolerance Security, Trust, and Privacy Organizational Aspects

Aneka: Cloud Application Platform, Framework Overview, Anatomy of the Aneka Container, From the Ground Up: Platform Abstraction Layer, Fabric Services, foundation Services, Application Services, Building Aneka Clouds, Infrastructure Organization, Logical Organization, Private Cloud Deployment Mode, Public Cloud Deployment Mode, Hybrid Cloud Deployment Mode, Cloud Programming and Management, Aneka SDK, Management Tools

**Module – 3**

Concurrent Computing: Thread Programming, Introducing Parallelism for Single **8 Hours**

Machine Computation, Programming Applications with Threads, What is a

Thread?, Thread APIs, Techniques for Parallel Computation with Threads, Multithreading with Aneka, Introducing the Thread Programming Model, Aneka Thread vs. Common Threads, Programming Applications with Aneka Threads, Aneka Threads Application Model, Domain Decomposition: Matrix

Multiplication, Functional Decomposition: Sine, Cosine, and Tangent. High- Throughput Computing: Task Programming, Task Computing, Characterizing a Task, Computing Categories, Frameworks for Task Computing, Task-based Application Models, Embarrassingly Parallel Applications, Parameter Sweep Applications, MPI Applications, Workflow Applications with Task Dependencies, Aneka Task-Based Programming, Task Programming Model, Developing Applications with the Task Model, Developing Parameter Sweep Application, Managing Workflows.

**Module – 4**

Data Intensive Computing: Map-Reduce Programming, What is Data-Intensive **8 Hours**

Computing?, Characterizing Data-Intensive Computations, Challenges Ahead, Historical Perspective, Technologies for Data-Intensive Computing, Storage

Systems, Programming Platforms, Aneka MapReduce Programming, Introducing

the MapReduce Programming Model, Example Application

**Module – 5**

Cloud Platforms in Industry, Amazon Web Services, Compute Services, Storage **8 Hours** Services, Communication Services, Additional Services, Google AppEngine, Architecture and Core Concepts, Application Life-Cycle, Cost Model,

Observations, Microsoft Azure, Azure Core Concepts, SQL Azure, Windows

Azure Platform Appliance.

Cloud Applications Scientific Applications, Healthcare: ECG Analysis in the

Cloud, Biology: Protein Structure Prediction, Biology: Gene Expression Data Analysis for Cancer Diagnosis, Geoscience: Satellite Image Processing, Business and Consumer Applications, CRM and ERP, Productivity, Social Networking, Media Applications, Multiplayer Online Gaming.

**Course outcomes:** The students should be able to:

Explain cloud computing, virtualization and classify services of cloud computing

Illustrate architecture and programming in cloud

Describe the platforms for development of cloud applications and List the application

of cloud.

**Question paper pattern:**

The question paper will have ten questions. There will be 2 questions from each module.

Each question will have questions covering all the topics under a module.

The students will have to answer 5 full questions, selecting one full question from each module.

**Text Books:**

1. Rajkumar Buyya, Christian Vecchiola, and Thamarai Selvi Mastering

Cloud. Computing McGraw Hill Education

**Reference Books:**

**1.** Dan C. Marinescu, Cloud Computing Theory and Practice, Morgan Kaufmann, Elsevier 2013.

**INFORMATION AND NETWORK SECURITY [As per Choice Based Credit System (CBCS) scheme] (Effective from the academic year 2016 -2017) SEMESTER – VII**

Subject Code 15CS743 IA Marks 20

Number of Lecture Hours/Week 3 Exam Marks 80

Total Number of Lecture Hours 40 Exam Hours 03

**CREDITS – 03**

**Course objectives:** This course will enable students to

Analyze the cryptographic processes.

Summarize the digital security process.

Indicate the location of a security process in the given system

**Module – 1 Teaching**

**Hours**

Introduction. How to Speak Crypto. Classic Crypto. Simple Substitution Cipher. **8 Hours** Cryptanalysis of a Simple Substitution. Definition of Secure. Double Transposition Cipher. One-time Pad. Project VENONA. Codebook Cipher.

Ciphers of the Election of 1876. Modern Crypto History. Taxonomy of

Cryptography. Taxonomy of Cryptanalysis.

**Module – 2.**

What is a Hash Function? The Birthday Problem.Non-cryptographic Hashes. **8 Hours**

Tiger Hash. HMAC. Uses of Hash Functions. Online Bids. Spam Reduction. Other Crypto-Related Topics. Secret Sharing. Key Escrow. Random Numbers.

Texas Hold 'em Poker. Generating Random Bits. Information Hiding.

**Module – 3**

Random number generation Providing freshness Fundamentals of entity **8 Hours**

authentication Passwords Dynamic password schemes Zero-knowledge

mechanisms Further reading Cryptographic Protocols Protocol basics From

objectives to a protocol Analysing a simple protocol Authentication and key establishment protocols

**Module – 4**

Key management fundamentals Key lengths and lifetimes Key generation Key **8 Hours**

establishment Key storage Key usage Governing key management Public-Key

Management Certification of public keys The certificate lifecycle Public-key

management models Alternative approaches

**Module – 5**

Cryptographic Applications Cryptography on the Internet Cryptography for **8 Hours**

wireless local area networks Cryptography for mobile telecommunications Cryptography for secure payment card transactions Cryptography for video broadcasting Cryptography for identity cards Cryptography for home users

**Course outcomes:** The students should be able to:

Analyze the Digitals security lapses

Illustrate the need of key management

**Question paper pattern:**

The question paper will have ten questions.

There will be 2 questions from each module.

Each question will have questions covering all the topics under a module.

The students will have to answer 5 full questions, selecting one full question from

each module.

**Text Books:**

1. Information Security: Principles and Practice, 2nd Edition by Mark Stamp Wiley

2. Everyday Cryptography: Fundamental Principles and Applications Keith M.

Martin Oxford Scholarship Online: December 2013

**Reference Books:**

1. Applied Cryptography Protocols, Algorithms, and Source Code in C by

Bruce Schneier

**UNIX SYSTEM PROGRAMMING**

**[As per Choice Based Credit System (CBCS) scheme]**

**(Effective from the academic year 2016 -2017)**

**SEMESTER – VII**

Subject Code 15CS744 IA Marks 20

Number of Lecture Hours/Week 3 Exam Marks 80

Total Number of Lecture Hours 40 Exam Hours 03

**CREDITS – 03**

**Course objectives:** This course will enable students to

Explain the fundamental design of the unix operating system

Familiarize with the systems calls provided in the unix environment

Design and build an application/service over the unix operating system

**Module – 1 Teaching**

**Hours**

Introduction: UNIX and ANSI Standards: The ANSI C Standard, The ANSI/ISO **8 Hours**

C++ Standards, Difference between ANSI C and C++, The POSIX Standards,

The POSIX.1 FIPS Standard, The X/Open Standards. UNIX and POSIX APIs:

The POSIX APIs, The UNIX and POSIX Development Environment, API Common Characteristics.

**Module – 2**

UNIX Files and APIs: File Types, The UNIX and POSIX File System, The **8 Hours**

UNIX and POSIX File Attributes, Inodes in UNIX System V, Application

Program Interface to Files, UNIX Kernel Support for Files, Relationship of C

Stream Pointers and File Descriptors, Directory Files, Hard and Symbolic Links. UNIX File APIs: General File APIs, File and Record Locking, Directory File APIs, Device File APIs, FIFO File APIs, Symbolic Link File APIs.

**Module – 3**

UNIX Processes and Process Control: The Environment of a UNIX Process: **8 Hours**

Introduction, main function, Process Termination, Command-Line Arguments,

Environment List, Memory Layout of a C Program, Shared Libraries, Memory

Allocation, Environment Variables, setjmp and longjmp Functions, getrlimit, setrlimit Functions, UNIX Kernel Support for Processes. Process Control: Introduction, Process Identifiers, fork, vfork, exit, wait, waitpid, wait3, wait4

Functions, Race Conditions, exec Functions, Changing User IDs and Group

IDs, Interpreter Files, system Function, Process Accounting, User Identification, Process Times, I/O Redirection. Process Relationships: Introduction, Terminal Logins, Network Logins, Process Groups, Sessions, Controlling Terminal, tcgetpgrp and tcsetpgrp Functions, Job Control, Shell Execution of Programs, Orphaned Process Groups.

**Module – 4**

Signals and Daemon Processes: Signals: The UNIX Kernel Support for Signals, **8 Hours**

signal, Signal Mask, sigaction, The SIGCHLD Signal and the waitpid Function, The sigsetjmp and siglongjmp Functions, Kill, Alarm, Interval Timers, POSIX.lb Timers. Daemon Processes: Introduction, Daemon Characteristics, Coding Rules, Error Logging, Client-Server Model.

**Module – 5**

Interprocess Communication : Overview of IPC Methods, Pipes, popen, pclose **8 Hours**

Functions, Coprocesses, FIFOs, System V IPC, Message Queues, Semaphores.

Shared Memory, Client-Server Properties, Stream Pipes, Passing File Descriptors, An Open Server-Version 1, Client-Server Connection Functions. **Course outcomes:** The students should be able to:

Ability to understand and reason out the working of Unix Systems

Build an application/service over a Unix system.

**Question paper pattern:**

The question paper will have ten questions. There will be 2 questions from each module.

Each question will have questions covering all the topics under a module.

The students will have to answer 5 full questions, selecting one full question from each module.

**Text Books:**

1. Unix System Programming Using C++ - Terrence Chan, PHI, 1999.

2. Advanced Programming in the UNIX Environment - W.Richard Stevens, Stephen A.

Rago, 3nd Edition, Pearson Education / PHI, 2005.

**Reference Books:**

1. Advanced Unix Programming- Marc J. Rochkind, 2nd Edition, Pearson Education,

2005.

2. The Design of the UNIX Operating System - Maurice.J.Bach, Pearson Education / PHI, 1987.

3. Unix Internals - Uresh Vahalia, Pearson Education, 2001.