DATA SCIENCE SEM 2

**19ECS702: SOFT COMPUTING**

Module I: Introduction to Soft Computing and Neural Networks

Evolution of Computing:

Soft Computing Constituents,

From Conventional AI to Computational Intelligence:

Machine Learning Basics

Module II: Basics of Fuzzy Logic

Fuzzy Logic:

Fuzzy Sets,

Operations on Fuzzy Sets,

Fuzzy Relations,

Membership Functions:

Fuzzy Rules and Fuzzy Reasoning,

Fuzzy Inference Systems,

Fuzzy Expert Systems,

Fuzzy Decision Making.

Module III: Introduction to Neural Networks

Neural Networks:

Machine Learning Using Neural Network,

Adaptive Networks,

Feed forward Networks,

Supervised Learning Neural Networks,

Radial Basis Function Networks:

Reinforcement Learning,

Unsupervised Learning Neural Networks,

Adaptive Resonance architectures

Module IV: Basics of Genetic Algorithms

Introduction to Genetic Algorithms (GA),

Applications of GA in Machine Learning:

Machine Learning Approach to Knowledge Acquisition

Module V: Introduction to Matlab /Python

Introduction to Matlab/Python,

Arrays and array operations,

Functions and Files,

Study of neural network toolbox and fuzzy logic toolbox,

Simple implementation of Artificial Neural Network and Fuzzy Logic

**19ECS773: Deep Learning**

Module I:

Introduction:

Feed forward neural networks (FFNN).

Gradient descent and the back propagation algorithm.

Unit saturation,

aka the vanishing gradient problem,

and ways to mitigate it.

Regularization techniques.

Module II:

Convolution Neural Network:

Architectures,

convolution / pooling layers,

Visualizing Convolution Networks,

Deep learning frameworks.

Module III:

Recurrent Neural Networks:

LSTM,

GRU,

Encoder Decoder architectures.

The Unreasonable Effectiveness of Recurrent Neural Networks.

Module IV:

Deep Unsupervised Learning:

Auto encoders (standard, sparse, de-noising, contractive, etc),

Variational Auto encoders,

Adversarial Generative Networks,

Auto encoder and DBM.

Module V:

Applications of Deep Learning:

Image segmentation,

object detection,

automatic image captioning,

Image generation with Generative adversarial networks.

**19ECS772: DATA SECURITY AND ACCESS CONTROL**

Module I:

Access Control Introduction to Access Control,

Purpose and fundamentals of access control,

brief history.

Policies of Access Control,

Models of Access Control,

and Mechanisms,

Discretionary Access Control (DAC),

Non- Discretionary Access Control,

Mandatory Access Control (MAC).

Capabilities and Limitations of Access Control Mechanisms:

Access Control List (ACL) and Limitations,

Capability List and Limitations.

Module II:

Role-Based Access Control (RBAC) and Limitations,

Core RBAC,

Hierarchical RBAC,

Statically Constrained RBAC,

Dynamically Constrained RBAC,

Limitations of RBAC.

Comparing RBAC to DAC and MAC Access control policy.

Module III:

Biba’s intrigity model,

Clark-Wilson model,

Domain type enforcement model,

mapping the enterprise view to the system view,

Role hierarchies- inheritance schemes,

hierarchy structures and inheritance forms,

using SoD in real system,

Temporal Constraints in RBAC,

MAC and DAC.

Integrating RBAC with enterprise IT infrastructures:

RBAC for WFMSs,

RBAC for UNIX and JAVA environments Case study:

Multi-line Insurance Company.

Module IV:

Smart Card based Information Security,

Smart card operating system-fundamentals,

design and implantation principles,

memory organization,

smart card files,

file management

atomic operation,

smart card data transmission ATR,

PPS Security techniques- user identification,

smart card security,

quality assurance and testing,

smart card life cycle-5 phases,

smart card terminals.

Module V:

Recent trends in Database security and access control mechanisms.

Case study of Role-Based Access Control (RBAC) systems,

Recent Trends related to data security management,

vulnerabilities in different DBMS

**19EOE746: OPERATIONS RESEARCH**

Module I:

Optimization Techniques,

Model Formulation,

models,

General L.R Formulation,

Simplex Techniques,

Sensitivity Analysis,

Inventory Control Models

Module II:

Formulation of a LPP

Graphical solution revised simplex method

duality theory

dual simplex method

sensitivity analysis

parametric programming

Module III:

Nonlinear programming problem

Kuhn-Tucker conditions min cost flow problem

max flow problem - CPM/PERT

Module IV:

Scheduling and sequencing –

single server and multiple server models

deterministic inventory model

Probabilistic inventory control models

Geometric Programming.

Module V:

Competitive Models,

Single and Multi-channel Problems,

Sequencing Models,

Dynamic programming,

Flow in Networks,

Elementary Graph Theory,

Game Theory Simulation

**19ECS767: BIG DATA ANALYTICS**

Module I:

Introduction to Bigdata

What is big data,

why big data,

convergence of key trends,

unstructured data,

industry examples of big data,

web analytics,

big data and marketing,

fraud and big data,

risk and big data,

credit risk management,

big data and algorithmic trading,

big data and healthcare,

big data in medicine,

advertising and big data,

big data technologies,

introduction to Hadoop,

open source technologies,

cloud and big data,

mobile business intelligence,

Crowd sourcing analytics,

inter and trans firewall analytics.

Module II: Introduction to NoSQL

Introduction to NoSQL,

aggregate data models,

aggregates,

key-value and document data models,

relationships,

graph databases,

schemaless databases,

materialized views,

distribution models,

sharding,

master-slave replication,

peer peer replication,

sharding and replication,

consistency,

relaxing consistency,

version stamps,

map-reduce,

partitioning and combining,

composing map-reduce calculations

Module III: Hadoop distributed file system

Data format,

analysing data with Hadoop,

scaling out,

Hadoop streaming,

Hadoop pipes,

design of Hadoop distributed file system (HDFS),

HDFS concepts,

Java interface,

data flow,

Hadoop I/O,

data integrity,

compression,

serialization,

Avro,

File-based data structures.

Module IV: Map-Reduce

MapReduce workflows,

Unit tests with MRUnit,

test data and local tests,

anatomy of MapReduce job run,

classic Map-reduce,

YARN,

failures in classic Map-reduce and YARN,

job scheduling,

shuffle and sort,

task execution,

MapReduce types,

input formats,

output formats

Module V:

HBase Data Model

HBase,

data model and implementations,

HBase clients,

HBase examples,

praxis.

Cassandra,

Cassandra data model,

Cassandra examples,

Cassandra clients,

Hadoop integration.

Hive,

data types and file formats,

HiveQL data definition,

HiveQL data manipulation,

HiveQL queries.