PH 532.1 RELATIONAL DATABASE MANAGEMENT SYSTEMS

Total No. of Lectures : 48	**Total Marks : 100**	**[L – T – P – S]**
No. of Lectures / Week : 4	**Credits : 4**	**[3 – 1 - 0 - 2]**

Course Objectives: This course helps the students.

- Introduction to the basic terminologies related to database management systems. Understanding various data models, relational models and relational algebra operations.
- Understanding the functional dependencies, schema refinement and Normalization for the design of Database.
- Knowledge on Transaction, concurrency control and Lock management for database design.
- Knowledge on Storage and Indexing.
- Preparing students for PL/SQL basic syntax, operators, functions and triggers.

 Course Outcomes: On successful completion of the course students will be able:

- Have good understanding about data and database systems. Describe the fundamental elements of relational database management systems.
- Understand the design of relational databases through the use of Entity-Relationship Diagrams and Normalization procedures and Develop basic skills in the use of SQL in defining and creating a database, inserting and modifying entries in a table.
- Gain Knowledge about Transaction, concurrency control and Lock management for database design.
- Have awareness about how data is stored in different storage media and how data is indexed.
- Prepare the students to understand the power of Query languages and also write PL/SQL transactions and to create different data objects.

Unit - I

Database System Applications: Purpose of Database Systems, View of Data, Data Abstraction, Instances and Schemas, Data Models – The ER Model, Relational Model, Other Models; Database Languages – DDL, DML, Data Storage and Querying, Database Architecture, ER Diagrams.

Relational Model: Introduction to the Relational Model – Integrity Constraints over Relations, Enforcing Integrity Constraints, Querying relational data, Logical data, Logical database design. Introduction to Views, Altering tables and views, Relational Algebra, Basic SQL Queries, Nested Queries, Complex Integrity Constraints in SQL, Triggers.

(**10 hours)**

Unit – II

Introduction to Schema Refinement: Problems caused by redundancy, Decompositions, Problem related to decomposition, Functional dependencies – Reasoning about FDS, Normal Forms – FIRST, SECOND, THIRD Normal Forms – BCNF – Properties of Decompositions – Lossless – Join decomposition, Dependency preserving Decomposition, Schema refinement in Database design – Multi valued Dependencies –FOURTH Normal Form, Join Dependencies, FIFTH Normal Form.

(10 hours)

Unit - III

Transaction Management: The ACID properties, Transactions and Schedules, Concurrent Execution of Transactions; Lock based concurrency control, Deadlocks – Performance of Locking, Transaction Support in SQL. **ConcurrencyControl:**Serializability and recoverability; Introduction to Lock Management – Lock Conversions, dealing with Deadlocks, Specialized Locking Techniques; Concurrency Control without Locking., Introduction to Crash recovery.

(10 hours)

Unit – IV

Overview of Storage and Indexing: Data on External Storage, File Organization and Indexing – Clustered Indexes, Primary and Secondary Indexes, Index Data Structures – Hash Based Indexing, Tree Based Indexing. **Storing data:** Disks and Files – The Memory Hierarchy – Redundant arrays of Independent disks. Tree Structured indexing: Intuitions for Tree indexes, Indexed Sequential Access Methods (ISAM) B+ Trees: A Dynamic Index Structure, Search, Insert, and Delete.

(9 hours)

Unit - V

Introduction to PL/SQL: Overview, Environment, Basic Syntax, Data Types, Variables, Constants and Literals, Operators, Conditions, Loops, Strings, Arrays, Procedures, Functions, Cursors, Records, Exceptions, Triggers, Packages, Collections, Transactions, Date & Time, DBMS Output, PL/SQL.

(9 hours)

Text Book:

[1]. Abraham Silberschatz, Heny F Korth, S. Sudarshan, "Database System Concepts", 6th Edition,McGraw Hill Publishers, India.
[2]. Kevin Loney, "Oracle Database 11i The Complete Reference", 1st Edition, 2014, McGraw Hill International,India

Reference Books:

[1].RamezElmasri, ShamkantNavathe, "Database Systems – Models, Languages, Designand Application Programming', 6th Edition, 2011, Pearson Asia.

[2].Thomas Conolly, Carolyn Begg, "Database Systems: A Practical Approach to Design,Implementation & Management", 5th Edition, 2012, Pearson Asia.

[3].Raghu Ramakrishnan, Johannes Gehrke, "Database Management Systems",3rd Edition, 2009, McGraw Hill Publishers International, India.

[4].Hector Garcia Molina, Jeffrey D Ullman, Jennifer Widom, "Database Systems : Thecomplete book", 2nd Edition, 2011, Pearson Education Asia.

[5].Catherine M Ricardo, "Databases Illuminated", 2nd Ed, 2010, Jones &BartletPublishers.

[6].Jeffrey A Hoffer, V Ramesh, HeikkiTopi, "Modern Database Management",11th Edition, 2013, Pearson Education, Asia.

[7].Jeffrey Ullman, Jennifer Widom, "First Course in Database Systems", 3rd Ed, 2009,Pearson Asia.

[8].David M Kroenke, David Auer, "Database Concepts", 6th Edition, 2009, Pearson Asia.

[9].Carlo Zaniolo, Stefano Ceri, Christos Faloutsos, Richard T Snodgrass, V S Subramanian, Roberto Zicari, "Advanced Database Systems", 6th Edition, 2012, Morgan Kaufmann.

*****

PH 533.1 OBJECT ORIENTED PROGRAMMING WITH JAVA

Total No. of Lectures: 48   **Total Marks : 100**   **[L – T – P – S]**

No. of Lectures / Week : 4   **Credits   :   4**   **[3 – 1 - 0 - 2]**

Course Objectives: The course objective is to:

- Introduce the principles and concepts of Object oriented programming followed the basic constructs of Java language.
- Explore the OOP concepts like constructors, method overloading, objects as parameters, nested classes and concept of Arrays in Java
- Understand String Handling in Java, concept of Inheritance and its importance and purpose of Interfaces
- Learn Exception Handling, File Handling and Collection Frameworks in Java
- Understand the concept of Multithreaded Programming and the importance of Generic programming

Course outcomes: After completing the course, the student should develop:

- An ability to understand the Object Oriented Concepts well and relate it with real world problems, develop solutions with programming constructs
- An understanding on classes, objects, methods, attributes, constructors and arrays and also write efficient programs using these concepts
- An ability to do string manipulation, understand and apply reusability using inheritance and also use Interfaces for efficient programming
- An understanding and clear knowledge about Exceptions and Exception handling, File I/O streams and also collection frameworks
- An ability   to develop and understand multithreaded applications with synchronization and apply generic programming concepts wherever required

Unit-I

Principles of Object Oriented Programming: A look at procedure oriented programming – characteristics, Object Oriented programming paradigm, Basic concepts of OOP, Benefits and application of OOP.

Java Programming Fundamentals : The Java Language, The Java Development Kit, The Java Keywords, Identifies in Java, The Java Class Libraries, Java's Primitive Types, Literals, The Scope and Lifetime of Variables, operators, Type conversion, Expressions, Input characters from the Keyword, Control structures – if, if – else and nested ifs, The if-else-if Ladder, The Switch Statement, Looping structures - for Loop, while Loop, The do-while Loop, Use of break and continue statements, Nested Loops and labelled loops.

(10 hrs)

Unit-II

Classes, Objects and Methods: Class Fundamentals, Creating Objects – the new operator, Reference Variables and Assignment, Methods, Returning from a Method, Returning Value, Using Parameters – this Keyword.

Constructors – default, Parameterized Constructors, Garbage Collection and Finalize method. Controlling Access to Class Members, Passing Objects to Methods, Returning Objects, and Method Overloading, Overloaded Constructors, Static members, Introducing Nested and Inner Classes.

Arrays: Arrays in Java, Multidimensional Arrays, Using the Length Member, the For- Each Style for Loop.

(10 hrs)

Unit-III

String Handling: String class Fundamentals, The String Constructors, String class methods, StringBuffer and StringBuilder classes.

Inheritance: Inheritance Basics, Member Access and Inheritance, Constructors and Inheritance, Using super under different contexts, Different types of inheritance, Method Overriding, Dynamic Method Dispatch, Using Abstract Classes, Using final keyword.

Interfaces: Interface Fundamentals, Creating an Interface, Implementing an Interface, Using Interface References, Implementing Multiple Interfaces.

(9 hrs)

Unit-IV

Exception Handling: Exception Handling Fundamentals, Try… throw…catch block, using Multiple catch clauses, Nested try blocks, Throwing an Exception, Throwable class, using finally, creating user defined exceptions and its use.

File Handling : Files and Streams, Operations using Reader and Writer classes

Collection Frameworks: Introduction to Java Collections Framework, Benefits of Java Collections Framework, Java Collections Interfaces – Iterator Interface, List Interface, Java Collection Classes – Hash Map Class

(10hrs)

Unit-V

Multithreaded Programming: Multithreading fundamentals, Thread life cycle, Creating a Thread - Thread Class and Runnable Interface, Creating Multiple Threads, Thread Priorities, Synchronization.

Generics: Introduction, Defining a Generic class – general form, Bounded types, Using wildcard arguments, creating generic method and constructors, generic interfaces and

Generic class hierarchies – use, generic subclass, casting and overriding methods ingeneric class.

(9 hrs)

Text Books:

[1].   Herbert Schildt, "Java :The Complete Reference", 8th Edition, 2014, McGraw- Hill. [2].   Paul J Deitel, Harvey M Deitel, "Java for Programmers", 2nd Edition, 2013, Pearson.

Reference Books:

[1].   Cay S Horstmann, "Java Concepts Compatible with Java 5, 6, 7 th Edition", 2013,Wiley.

[2].   Julie Anderson, Harvey J, "Java Illuminated", 3rd Ed, 2013, Jones & Bartlett.

[3].   C Xavier, "Java Programming: A Practical Approach", 1st Edition, 2011, McGrawHill.

[4].   Cay Horstman, Gary Cornel, "Core Java : Volume – 1 Fundamentals", 8th Edition,2010, Pearson Asia.

[5].   Edward Currie, "Fundamentals of Programming using Java", 2nd Edition, 2009,Cengage.

[6].   Anthony J Dos Reis, "Introduction to Programming using Java", 1st Ed, 2012, J&B Int [7].   Jana, Debashish, "Java and Object Oriented Programming Paradigm", 2nd Edition, 2010, PHI Learning,

[8].   Joyce Farrell, "Java Programming", 2nd Edition, 2008, Cengage Learning India[9].   Steve Holzner, "Java 2 Black Book", 1st edition, 2009, Wiley India

----

PH 534.1 : WEB DESIGN WITH PHP and MYSQL

Total No. of Lectures : 48	**Total Marks: 100**	**[L – T – P – S]**
No. of Lectures / Week: 4	**Credits : 4**	**[3 – 1 - 0 - 2]**

Course Objectives: The objective of this course is to make the students:

- Understand the principles of creating an effective web page, including an in-depth consideration of information architecture using HTML and CSS.

- To develop basic programming skills using JavaScript and jQuery

- To become familiar with jQuery and its use in efficiently building Ajax applications.

- To provide the necessary knowledge to design and develop dynamic, database-driven web applications using PHP.

- Will learn how to connect to any database, and to create database-driven HTML forms and reports.

Course Outcomes: The Candidate will be able

- To use knowledge of HTML and CSS code and an HTML editor to create personal and/or business websites following current professional and/or industry standards. Use critical thinking skills to design and create websites.
- To create effective scripts using JavaScript.
- To enhance the end user experience using JQuery.
- Students can be employed on entry-level jobs of PHP based web development in software industry
- To develop interactive and dynamic website using PHP and database connectivity.

Unit I

HTML 5: Difference between HTML 4 & 5, Page Structure of HTML5. Section Tag & Article Tag; HTML5 Forms Tags, search, tel, url and email; date/time input types, colornumber, range, min, max, and step attributes. Form Attributes, auto complete, no validate. Form Field Attributes, required, placeholder.**CSS:** Introduction to CSS, CSS Rules, Pseudo classes and pseudo elements, Selector Forms, The Box Model, Conflict Resolution; Cascading; Levels of Style Sheet, Style Specification Formats; Borders, and backgrounds. **(10 hrs)**

Unit II

JavaScript: The Basic of JavaScript, Primitives Operations and Expressions, Value, Data types , Variables, Control Statements, Arrays, Functions, Constructors,Object Creation and Object Model, Events and Listeners, Applications ,JavaScript Libraries.

(9 hrs)

Unit III

JQuery : Overview of jQuery; Structure of jQuery, Using jQuery and including .js file to HTML, Type of Selectors, Handling Events with jQuery, Filters;

AJAX: Introduction to AJAX, AJAX Load, Change AJAX data type, Status Codes, JSON, Accessing and consuming remote JSON Data.

JQuery UI: Introduction to JQuery UI; JQuery UI Interactions, JQuery UI Widgets, JQuery UI Effects.**(10 hrs)**

Unit IV

PHP Basics : Introduction to PHP; Embedding PHP code; Working with Types and operators; Integer division, Generating Random numbers; Comparing values with the Spaceship operator; Condition and Looping statements; Arrays, Numeric Associative andMulti-dimensional arrays; Strings and String functions; Regular Expressions; Date and Time function; Mathematical functions; User-Defined Functions; Return type declarationswith compound Types; Scalar Type Hinting; Scalar Parameter and Return Type declarations; Working with web forms; PHP POST & GET form elements; Validating form data; Understanding magic quotes; Setting default values in forms; Handling Errors,Throwing and Catching Exceptions. Logging exceptions.

(9hrs)

Unit V

PHP Intermediate: Classes and Objects Constructors and Destructors Access Specifies, Static and final modifiers Inheritance and its types Polymorphism, Abstract and Final classes. Anonymous Classes, Generator Return expressions, Generator Delegation. Using MySQL; Database Basics, Configuring PHP for Database Support, Managing Database Connections, PHP's Database & SQL APIs, Performing Queries, Processing Result Sets, PDO. PHP cookie handling; PHP session handling; Reading & writing; files with PHP; Processing uploaded files; Retrieving uploaded files; Mail functions.

(10 hrs)

Text Books:

[1]. Robin Nixon, "Learning PHP, MySQL & JavaScript: With jQuery, CSS & HTML5", 5th Ed, 2018, O'Reilly Publishers.

[2]. Christopher Murphey, R Clark, "Beginning HTML 5 and CSS 3", 1st Ed, 2015, Wiley Apress.

Reference Books:

[1].Steve Prettyman, "Learn PHP 7: Object Oriented Modular Programming using HTML5,CSS3, JavaScript, XML, JSON, and MySQL", 2016, Apress.

[2]. Simon Sarris,"HTML5 Unleashed", 2014, Pearson Education.[3]. Piotr Sikora, "Professional CSS3", 2016, Packt Publishing.

[4]. Joel Murach, Ray Harris , "Murach's PHP and MySQL", 2nd Ed, 2014, Mike Murach& Associates

[5]. Mary Delamater, Zak Ruvalcaba, "Murach's JavaScript and JQuery", 3rd Ed, 2017,Mike Murach& Associates

[6]. Eric Sarrion, "jQuery UI", 2014, Oreilly Publishers.

[7]. Richard Blum, "PHP, MySQL & JavaScript", 2018, John Wiley & Sons.

[8]. Luke Welling, Laura Thomson, "PHP and MySQL Web Development", 5th Ed, 2018,Pearson

Education.

PH 532.2 MOBILE APPLICATION DEVELOPMENT WITH ANDROID

Total No. of Lectures : 48   **Total Marks: 100**   **[L – T – P – S]**
No. of Lectures / Week: 4   **Credits   :   4**   **[3 – 1 - 0 - 2]**

Course Objectives: The objective of this course is to make the students:

- To understand the development environments, Architectures and programming paradigms of Android devices.
- To design and develop an Android App with various controls that provide interactivity and handle user inputs.
- To understand the various building blocks of Android and implement Broadcast Receiving, Services, Content Providers etc.
- To understand how various types of data is stored in an Android Based device and also perform read and write operations on the same.
- Implement and understand how various advanced features of Android such as Sensors, Accelerometer, GPS, etc.

Course Outcomes: Students must be able to

- Understand the architecture, working and environmental setup of Android
- Design and Implement simple GUI based Android Apps that handle user input and provide information
- Implement Android apps that are able to receive broadcasted messages, act as content provider or receiver and run background services.
- Create Android Apps that can manipulate data from various data stores such as internal, external memory and also SQLite as a Database.
- Design and Work with advanced sensors of the phone and manipulate Telephony and SMS in an Android Phone.

Unit - I

Introduction to Android - Architecture - Features of Android SDK - Introducing the Development Framework, Developing for Android- Android Development Tools- The Android Virtual Device Manager - Android SDK Manager - The Android Emulator - The Dalvik Debug Monitor Service - The Android Debug Bridge. Installing and Configuring Android SDK, ADT and AVD - Android Software Stack, Dalvik Virtual Machine.

(10hrs)

Unit – II

Creating Android Applications: Working of Android Application: Android Application Life Cycle; Building User Interfaces - Introducing Activities, Various Layouts, Fragments, and Adapters. Intents, Intent Filters, Controls, Dialogs, Toasts and Notifications: Displaying Pictures and Menus with Views; Adding Interactivity and Handling UI events.

<div align="right">(10 hrs)</div>

Unit – III

Building Blocks of Android: Activities: Life Cycle and Working; Broadcast Receivers: Creating, Registering and unregistering a Broadcast Receiver. Content Providers: Concept of Content Resolver, Creating and using Content Providers. Services: Bound Services and Unbound Services, Life Cycle of services. **(9 hrs)**

Unit - IV

Data Access in Android: Creating, Saving and Retrieving Shared Preferences; Introducing the Preference Framework and the Preference Activity; Including Static Files as Resources; Working with the File System.Introducing Android Databases; Introducing SQLite; Content Values and Cursors; Working with SQLite Databases; Creating and Using Content Providers; Using Native Android Content Providers. **(10 hrs)**

Unit – V

Advanced Android:Introducing Services, Using Background Threads, Using Alarms; Controlling Device Vibration; Introducing Android Text-to-Speech; Using Sensors and the Sensor Manager; Using Accelerometer, Compass and GPS; Audio, Video and Using the Camera; Using Telephony and SMS.

<div align="right">(9 hrs)</div>

Text Books:

[1]. Reto Meier, Ian Lake, "Professional Android", 4th edition, 2018,Wiley Wrox Publications.

[2]. Dawn Griffiths, David Griffiths, "Head First Android Development", 2nd Edition, 2017, O'Reilly Media.

Reference Books:

[1]. Bill Phillips, Chris Stewart, Kristin Marsicano, "Android Programming: The Big Nerd Ranch Guide", 3rd edition, 2017, Big Nerd Ranch Guides.

[2]. John Horton, "Android Programming for Beginners", 2nd edition, 2018, Ingram short title.

[3]. Paul Dietel, Harvey Dietel, "Android 6 for Programmers : An App Driven approach", 2016, Pearson

[4]. Reito Meier, "Professional Android Application Development", 1st Ed, 2012, Wiley Press.

[5]. Wei- Meng Lee, "Beginning Android 4 Application Development", 1st Ed, 2012, Wiley Press.

[6]. Mednieks, Laird Dornin, Blake Meike, "Programming Android", 2nd Ed,2012, O'Reilly Media.

PS 535.2 E2 MACHINE LEARNING AND DEEP LEARNING

Total No. of Lectures: 42   **Total Marks: 100**   **[L –T – P – S]**
No. of Lectures / Week: 4   **Credits: 3**   **[3 – 1 - 0 - 2]**

Course Objectives: The objective of this course is to make the students:
- To understand the theoretical and practical aspects of machine learning, learning the types of machine learning, application of bayes rule in machine learning, learning the imporatnace of dimensionality reduction
- To appreciate supervised and unsupervised learning algorithms and their applications, understand the method of clustering, Support Vector Machine and decision tree for developing machine learning models, importance of gradient descent in ML
- To learn the basic knowledge of HMM, MLP & reinforcement learning and their application in creating classification model, learning to evaluate the performance metric of a machine learning model
- To understand the concept of deep learning and importance of applied mathematics in machine learning, learn the basic idea of neural network, knowing the basic concept of CNN and RNN
- To gain knowledge of encoders and decoders, importance of transfer learning, understand the concept of structured probabilistic model

Course Outcomes: By the end of the module, the student should be able
- To implement Machine Learning with Bayes algorithm, to work out the concept of dimensionality reduction using PCA & LDA
- To implement Machine Learning with SVM, Decision tree and clustering methods
- To use MLP, HMM for classification and also to measure the performance of the classification algorithm, to design models using reinforcement learning
- To implement CNN and RNN for Deep Learning models by applying all the methods for creating optimal model
- To implement Transfer learning and Auto encoders for Deep Learning models

Unit - I

Introduction: Machine Learning Foundations, Overview, Applications, Types of Machine Learning,Basic Concepts in Machine Learning, Examples of Machine Learning, Applications.

Supervised Learning: Probably Approximately Correct (PAC) Learning, Learning Multiple Classes, Regression, Model Selection and Generalization, Dimensions of a Supervised Machine Learning Algorithm.

Dimensionality Reduction: Introduction, Subset Selection, Principal Components Analysis, Linear Discriminant Analysis.

Bayesian Decision Theory: Introduction**,** Classification, Discriminant Functions, Bayesian Networks, Association Rules.

(9 hrs)

Unit - II

Decision Trees: Introduction, Univariate Trees, Pruning, Rule Extraction from Trees, Multivariate Trees.

Linear Discrimination: Introduction, Generalizing the Linear Model, Geometry of the Linear Discriminant, Pairwise Separation, Gradient Descent, Logistic Discrimination, Support Vector Machines.

Clustering: Introduction, Mixture Densities, k-Means Clustering, Expectation-Maximization Algorithm, Hierarchical Clustering, Choosing the Number of Clusters.
Nonparametric Methods: Introduction, Nonparametric Density Estimation, Nonparametric Classification, Nonparametric Regression: Smoothing Models.

(9 hrs)

Unit - III

Multilayer Perceptrons: Introduction, Training a Perceptron, Back propagation Algorithm, Training Procedures, Tuning the Network Size.
Hidden Markov Models: Introduction, Discrete Markov Processes, Three Basic Problems of HMMs, Evaluation Problem, Learning Model Parameters, Model Selection in HMM.
Accessing and Comparing Classification Algorithms: Introduction, Cross-Validation and Resampling Methods, Measuring Error, Assessing a Classification Algorithm's Performance, Comparing Two Classification Algorithms, Comparing Multiple Classification Algorithms: Analysis of Variance.
Reinforcement Learning: Introduction, Elements of Reinforcement Learning, Model-Based Learning, Temporal Difference Learning, Generalization.

(8hrs)

Unit - IV

Introduction: Historical trends in Deep learning, Deep learning: Overview of Methods.
Applied Math for Machine Learning: Linear Algebra, Probability and Information Theory, Numerical Computation.
Deep Networks: Deep Feedforward Networks – Feed Forward networks, Backpropogation; Regularization- Overview, Parameter Penalties, Data Augmentation, Multi Task learning, Bagging, Dropout; Optimization for Training Deep Models – Optimization vs training, Basic Algorithms, Adaptive learning Rates; Convolution Networks – The Convolution operation and CNNs, Convolution Networks, Pooling; Sequence Modeling: Recurrent and recursive Nets – Sequence Modeling, Unfolding Graphs, Recurrent Neural networks, Bidirectional RNNs, Deep Recurrent Networks.

(8hrs)

Unit - V

Deep Learning research: Linear Factor Models- PCA and factor Analysis, ICA; Autoencoders – Stochastic Encoders and Decoders, DenoisingAutoencoders, Applications; Representation Learning- Greedy Layer wise Unsupervised Pretraining, Transfer Learning and Domain Adaptation, Semi-supervised Distangling of causal factors, Distributed Representation; Structured Probabilistic Models for Deep Learning – Using Graphs to describe model structure, Boltzmann Machines, Deep Belief Networks, Directive Generative nets.

(8hrs)

PH 531.2 PROGRAMMING WITH PYTHON

Total No. of Lectures : 48	Total Marks : 100	[L – T – P – S]
No. of Lectures / Week : 4	Credits : 4	[3 – 0 - 2 - 2]

Course Objectives: The objective of this course is to make the students:
- Describe the core syntax and semantics of Python programming language.
- Discover the need for working with the strings and functions.
- Illustrate the process of structuring the data using lists, dictionaries, tuples and sets.
- Indicate the use of regular expressions and built-in functions to navigate the file system.
- Infer the Object-oriented Programming concepts in Python.

Course Outcomes: At the end of this course students would have learned
- To design and program Python applications, use lists, tuples, and dictionaries in Python programs.
- To identify Python object types, use indexing and slicing to access data in Python programs.
- To build and package Python modules for reusability and to read and write files in Python.
- Tto design object oriented programs with Python classes and use class inheritance in Python for reusability.
- To use exception handling in Python applications for error handling.

Unit I

Introduction to Python: Installation and Working with Python, Understanding Python variables, basic operators, blocks. Declaring and using Numeric data types: int, float, complex, using string data type and string operations, defining list and list slicing, Use of Tuple data type.

(10 hrs)

Unit II

Program Flow Control & Functions: Conditional blocks using if, else and elif, for loops in, for loop using ranges, use of while loops in python, Loop manipulation using pass, continue, break and else.Organizing python codes using functions, organizing python projects into modules, importing own module as well as external modules, Understanding Packages.

(10 hrs)

Unit III

String, List, Dictionary &File handling: Understanding in-built string methods, List manipulation, Dictionary manipulation. Programming using string, list and dictionary in-built

functions. Reading files in python, writing files in python, Understanding read and write functions, manipulating file pointer using seek, Programming using file operations.

(10 hrs)

Unit IV

OO Programming & Regular Expressions: Concept of class, object and instances, constructor, class attributes and destructors, Inheritance, overlapping and overloading operators, Adding andretrieving dynamic attributes of classes. Pattern matching and searching, Power of pattern searching using regex in python, Pattern finding programs using regular expression.

(9 hrs)

Unit V

Exception Handling & Database Interaction: Avoiding code break using exception handling, Safe guarding file operation using exception handling, Handling and helping developer with errorcode. SQL Database connection using python, Creating and searching tables, Reading, storing andupdating information on database.

(9 hrs)

Text Books:

[1]. R. Nageswara Rao, "Core Python Programming", 2nd Edition, 2018, Dreamtech Press.

[2]. Jason Rees, "Python Programming: A Practical Introduction To Python Programming For Total Beginners Paperback", 2019, Independent Publication.

Reference Books:

[1]. Eric Matthes, "Python Crash Course, 2nd Edition: A Hands-On, Project-Based Introduction to Programming", 2019, No Starch Press.

[2]. Zed A. Shaw, "Learn Python 3 the Hard Way", 2017, Addison Wesley.

[3]. John M Zelle, "Python Programming: An Introduction to Computer Science", 2016, Ingram short title; Third edition.

[4]. David Beazley, Brian K. Jones, "Python Cookbook: Recipes for Mastering Python 3, 2013, Oreilly 3rd Edition.

PS 534.3 E1 INTERNET OF THINGS and APPLICATIONS

Total No. of Lectures: 42	**Total Marks: 100**	**[L – T – P – S]**
No. of Lectures / Week : 4	**Credits : 3**	**[3 – 1 - 0 - 2]**

Course Objectives: As part of this course, students will study,

- Learn the basic concepts of internet of things, network and communication, security privacy trust
- Learn the concepts of IoT reference architecture, Identity Management Models, trust management in IoT
- Learn different IoT protocols in different layers of OSI Layers. Also learn IoT analytics for cloud
- Learn how IoT data is explored and how data is displayed using visualization techniques
- Learn how machine learning techniques are used to extract the useful data and how the data is organized for data analytics

Course Outcomes – upon successful completion of this course, the participant will be able to:

- Understand why IoT is used and how it is implemented and how networks and communication is used to implement IoT
- Understand how identity management models are used in IoT, also understand why trust management is important for IoT environment
- Understand the use of protocols which are used in different layers and how it is combined with other protocols down the layers to carry out the communication
- Understand how data is stored in cloud and how it is represented using different application to carry out or execute different data analytics tools
- Understand the concepts of data science for IoT analytics, how to organize data for analytics, and how to get benefits from IoT analytical tools.

Unit I

IoT Web Technology: The Internet of Things Today, Time for Convergence, Towards the IoT Universe, Internet of Things Vision, IoT Strategic Research and Innovation Directions, IoT Applications, Future Internet Technologies, Infrastructure, Networks and Communication, Processes, Data Management, Security, Privacy & Trust, Device Level Energy Issues, IoT Related Standardisation, Recommendations on Research Topics.

Internet of Things Privacy, Security and Governance: Introduction, Overview of Governance, Privacy and Security Issues, Contribution from FP7 Projects, Security,

Privacy and Trust in IoT-Data-Platforms for Smart Cities, First Steps Towards a Secure Platform, Smartie Approach. Data Aggregation for the IoT in Smart Cities, Security.

(9 hrs)

Unit II

Architectural Approach for IoT Empowerment : Introduction, Defining a Common Architectural Ground, IoT Standardisation, M2M Service Layer Standardisation, OGC Sensor Web for IoT, IEEE, IETF and ITU-T standardization activities, Interoperability Challenges, Physical vs Virtual, Solve the Basic First, Data Interoperability, Semantic Interoperability, Organizational Interoperability, Eternal Interoperability, Importance of Standardisation, Plan for validation and testing, Important Economic Dimension,Research Roadmap for IoT Testing Methodologies. Semantic as an Interoperability Enabler and related work.

Identity Management Models in IoT : Introduction, Vulnerabilities of IoT, Security requirements, Challenges for a secure Internet of Things, identity management, Identity portrayal, Different identity management model: Local identity, Network identity, Federated identity, Global web identity, Identity management in Internet of Things, User-centric identity management, Device-centric identity management, Hybrid identity management.

Trust Management in IoT: Introduction, Trust management life cycle, Identity and trust, Third party approach, Public key infrastructure, Attribute certificates, Web of trust models, Web services security, SAML approach, Fuzzy approach for Trust, Access control in IoT, Different access control schemes, Authentication and Access control policies modelling.   **(9 hrs)**

Unit - III

Defining IOT Analytics and challenges: The situation; Defining IoT analytics; IoT analytics challenges
Business value concerns.

IoT Devices and Networking Protocols: IoT devices, Networking basics, IoT networking connectivity protocols, IoT networking data messaging protocols, Analyzing data to infer protocol and device characteristics.

IoT Analytics for Cloud: Building elastic analytics; Elastic analytics concepts; Designing for scale Cloud security and analytics; The AWS overview; Microsoft Azure overview; The Thing Worx overview;   **(8 hrs)**

Unit - IV

Exploring IoT Data: Exploring and visualizing data, attributes that might have predictive value R. Adding internal datasets; Adding external datasets

Visualizing and Dashboarding : Common mistakes when designing visuals; The Hierarchy of Questions method; Designing visual analysis for IoT data; Creating a dashboard with
Tableau; Creating and visualizing alerts.   **(8 hrs)**

Unit -V

Data Science for IoT Analytics: Machine learning (ML); Anomaly detection using R; Forecasting using ARIMA; Deep learning

Strategies to Organize Data for Analytics: Linked Analytical Datasets; Managing data lakes; The data retention strategy.

Economics of IoT Analytics: The economics of cloud computing and open source, Cost considerations for IoT analytics, revenue opportunities, predictive maintenance example

                                                                                    (8 hrs)

Text Books:

[1].   Andrew Minteer, "Analytics for the Internet of Things (IoT)", 2nd Ed, 2017, PACKT

[2].   Arsheep Bhaga, Vijay Madasetti, "Internet of Things : A Hands on Approach", 1st Edition, 2016, University Press India Ltd

[3].   Donald Norris, "The Internet of Things: Do-It-Yourself at Home Projects for Arduino, Raspberry Pi and BeagleBone Black", 1st Edition, 2015, McGraw Hill Publishers

Reference Books:

[1].   Hwaiyu Geng, "Internet of Things and Data Analytics Handbook", 1st Edition, 2016, Wiley

[2].   Peter Waher, "Mastering Internet of Things", 1st Edition, 2018, PACKT Publishers

[3].   Perry Lea, "Internet of Things for Architects", 2nd Edition, 2017, PACKT Publishers

[4].   Qusay F. Hassan, Atta ur Rehman Khan, Sajjad A. Madani, "Internet of Things: Challenges, Advances, and Applications", 2nd Edition, 2014, CRC Press

[5].   Pethuru Raj, Anupama C. Raman, "The Internet of Things: Enabling Technologies, Platforms, and Use Cases", 2nd Edition, 2014, CRC Press.

[6].   Olivier Hersent, David Boswarthick, Omar Elloumi, "The Internet of Things: Key Applications and Protocols", 2nd Edition, 2015, Wiley International

[7].   Adrian McEwen, Hakim Cassimally, "Designing the Internet of Things", 1st Ed, 2013, Wiley

[8].   Mile Lukodes, John Bruner, "What is the Internat of Things", 1st Ed, 2015, O'Reilly Publishers

[9].   Hakima Chaouchi, "The Internet of Things: Connecting Objects", 1st Edition, 2014, Wiley

[10].   Fawzi Behmann, Kwok Wu, "Collaborative Internet of Things (C-IoT): for Future Smart Connected Life and Business", 1st Edition, 2015, Wiley International

[11].   Fei Hu, "Security and Privacy in Internet of Things (IoTs): Models, Algorithms,", 2016, CRC Press

[12]. Weber, Rolph H, Romana, "Internet of Things : Legal Perspectives", 1st Ed, 2015, Springer

PS 535.3 E2: BIG DATA ANALYTICS with MAP REDUCE & HADOOP

Total No. of Lectures : 42	**Total Marks : 100**	**[L – T – P – S]**
No. of Lectures / Week : 4	**Credits : 3**	**[3 – 1 - 0 - 2]**

Course Objectives: The course aims to cover the following objectives

- Understanding the Big Data Fundamentals, including the characteristics of Big Data, the sources Big Data (such as social media, sensor data, geospatial data etc) and Big Data Analytics.

- Introducing to No-SQL databases and different properties and characteristics associated with it.

- Learning about Hadoop, the need for it, Distributed processing, and using it for analyzing huge amount of data.

- Understanding Map Reduce programming and learning to write programs using mapper and reducer.

- Exploring Hadoop Eco system, familiarizing with Hive and Pig.

Course Outcomes : Upon Completion of the course, the students will be able to

- Identify and distinguish big data analytics applications from other applications and the use of Big Data.

- Describe No SQL databases and understanding different concepts related to No SQL and its applications using MongoDB.

- Understanding Hadoop and its advantage over the traditional database applications in solving practical problems

- Writing programs using mapper and reducer.

- Using Hive and Pig for analyzing and querying data and knowing the advantages over the traditional Data handling solutions.

Unit- I

Introduction to big data: Data, Characteristics of data and Types of digital data:, Sources of data, Working with unstructured data, Evolution and Definition of big data, Characteristics and Need of big data, Challenges of big data.

Big data analytics: Overview of business intelligence, Data science and Analytics, Meaning and Characteristics of big data analytics, Need of big data analytics, Classification of analytics, Challenges to big data analytics, Importance of big data analytics, Basic terminologies in big data environment.

(8 hrs)

Unit- II

NoSQL Data Management : 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; Case studies using MongoDB.

(8 hrs)

Unit –III

Introduction to Hadoop: Introducing Hadoop, need of Hadoop, limitations of RDBMS, RDBMS versus Hadoop, Distributed Computing Challenges, History of Hadoop , Hadoop Overview, Use Case of Hadoop, Hadoop Distributors, HDFS (Hadoop Distributed File System) , Processing Data with Hadoop, Managing Resources and Applications with Hadoop YARN (Yet another Resource Negotiator), Interacting with Hadoop Ecosystem **.**

(**8 hrs)**

UNIT –IV

Introduction to MAPREDUCE Programming: Introduction, Mapper, Reducer, Combiner, Partitioner , Searching, Sorting , Compression, Real time applications using MapReduce, Data serialization and Working with common serialization formats, Big data serialization formats.

(**8hrs)**

UNIT –V

Introduction to Hive: Introduction to Hive, Hive Architecture , Hive Data Types, Hive File Format, Hive Query Language (HQL), User-Defined Function (UDF) in Hive.

Introduction to Pig: Introduction to Pig, The Anatomy of Pig , Pig on Hadoop , Pig Philosophy , Use Case for Pig: ETL Processing , Pig Latin Overview , Data Types in Pig , Running Pig , Execution Modes of Pig, HDFS Commands, Relational Operators, Piggy Bank , Word Count Example using Pig , Pig at Yahoo!, Pig versus Hive.

(9 hrs)

Text Book :

[1]   Seema Acharya, SubhashiniChellappan "Big Data and Analytics", 1st Edition, 2015,

Wiley International

[2]   S. Mohanthy, MadhuJagadish, Harsh Srivatsa, "Big Data Imperatives: Enterprise Big Data Warehouse, BI Implementations and Analytics", 1st Edition, 2015, Wiley Apress

Reference Books :

1. JeyLiebowitz, "Big Data and Business Analytics", 1st Edition, 2014, CRC Press

2. Tom White, "Hadoop – Definitive Guide", 1st Edition, 2015, O'Reilly Publishers

3. Boris Lublinsky, Kevin t. Smith, Alexey Yakubovich,

   "Professional Hadoop Solutions",1st Edition, 2015, Wiley India

   Publications.

4. E Caprilo, Dean Wampler, "Hive Programming", 1st Edition, 2015, O'Reilly Publishers

5. SherifSakr, "Large Scale and Big Data: Processing and Management", 2014, CRC Press.

PH531.3 CLOUD COMPUTING WITH AMAZON WEB SERVICES

Total No. of Lectures : 48   **Total Marks : 100**   **[L – T – P – S]**

No. of Lectures / Week : 4   **Credits   :   4**   **[3 – 1 - 0 - 2]**

Course Objectives:
- To introduce the broad perceptive of cloud architecture and model
- To understand the design of cloud Services and the trusted cloud Computing system
- To understand applications of cloud
- To know how to design cloud applications
- To know how to use AWS services in cloud applications

Course Outcomes: By the end of the module, the student should be able
- Describe the key technologies, architecture, strengths, limitations and applications of cloud computing

- Explain the types and service models of cloud and Understand security implications in cloud computing

- Design Cloud Services and Set a private cloud

- Create and automate infrastructure to design cost-effective, highly available applications

- Integrate AWS services with your application to meet and exceed non-functional requirements

Unit I

Introduction ,Cloud Computing at a Glance, The Vision of Cloud Computing, 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, ComputingPlatforms and Technologies. 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.

(10 hrs)

Unit II

Cloud Computing Architecture, Introduction, Cloud Reference Model, 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, Principles of Parallel and Distributed Computing- Elements of parallel computing, Elements of distributed computing, Technologies for distributed computing.

(10 hrs)

Unit III

Energy efficiency in clouds . Energy-efficient and green cloud computing architecture, Market-based management of clouds, Market-oriented cloud computing, Reference model, A reference model for MOCC , Technologies and initiatives supporting MOCC , Federated clouds/Inter Cloud- Characterization, Cloud federation stack, Aspects of interest, Technologies for cloud federations, Third-party cloud services. Cloud Applications- Scientific applications (Healthcare, Biology, Geoscience) Business and consumer applications.

(10 hrs)

Unit IV

Designing Cloud Applications: Introducing cloud-based multitier architecture , Designing for multi-tenancy , Understanding cloud applications design principles Understanding emerging cloud-based application architectures , Estimating your cloud computing costs , A typical e-commerce web application, AWS Components, Managing costs on AWS Cloud, Application development environments

(9 hrs)

Unit V

Scalability: Objectives, Designing scalable application architecture, Leveraging AWS infrastructure services for scalability, architecture –increasing loads, Event handling at scale .High Availability: Objectives, Nature of failures, AWS High availability architecture, Security – Security objectives, Best practices in implementing AWS security.

(9 hrs)

Text Books:
[1]. Rajkumar Buyya, Christian Vecchiola, S ThamaraiSelvi, "Mastering Cloud Computing", 1st Edition, 2015, Tata Mcgraw Hill Publishing Co Ltd.
[2]. Aurobindo Sarkar, Amith Shah "Learning AWS" Second edition 2018,Packt Publishing Ltd.

Reference Books:
[1]. Atul V. Mistry ,"Expert AWS Development", First edition 2018, Packt Publishing Ltd.
[2]. Kris Jamsa, "Clooud Computing", 1st Edition, 2013, Jones & Bartlett Publishers.
[3]. Anthony T Velte, "Cloud Computing: A practical Approach", 1st Ed, 2010, Tata McGraw Hill.
[4]. Fern Halper, Kaufman, Bloor Robin, Hurwit, "Cloud Computing for Dummies", 1st Ed, 2010, Wiley Press.
[5]. Joshy Joseph , Craig Fellenstein, "Grid Computing", 4th Edition, 2013, Pearson Asia.
[6]. Ahmar Abbas, "Grid Computing: A Practical Guide to technology and Applications", 2010, CRC Press.
[7]. Kaittwang Geoffrey C.Fox, Jack J Dongrra,"Distributed and Cloud Computing",1st Ed, 2012 , Elsevier.

PH 532.3:   WEB TECHNOLOGIES and .NET FRAMEWORK

Total No. of Lectures: 48	**Total Marks: 100**	**[L –T–P– S]**
No. of Lectures / Week: 4	**Credits : 4**	**[3 – 1 –0 –2]**

Course Objectives: The objective of this course is to make the students:
- To study the elements of the .NET Framework platform, features of ASP.NET IDE and to learn the features of C# language.
- To develop Dynamic Web Applications using ASP.NET and to implement web 2.0 features using AJAX.
- To understand the architecture and main classes of ADO.NET and to use LINQ and Entity Framework in Applications.
- To Develop ASP.NET MVC application and Create distributed applications by means of Web Services.
- To learn ASP.NET Core 3.0 architecture, principles and benefits.

Course Outcomes: By the end of the module, the student should be able

- Learn to develop correct, well documented programs using C# programming language.
- Create visually rich and attractive Web applications with ASP.NET controls and controls in the AJAX Control Toolkit
- Display dynamic data from a data source by using Microsoft ADO.NET, LINQ and EF.
- Create MVC Models and write code that implements business logic within Model methods, properties, and events.Dynamic web applications, create and consume web services,understand the Microsoft Web Technologies stack.
- Write an application that can create, edit, and view data from a database using ASP.Net Core, and create Single Page Applications (SPAs)and Navigation, Routing, State Management, Security.

Unit -I

.NET Framework and C#: Architecture of .NET Framework. -.NET Compliant Languages - The Common Language Runtime (CLR); Common Type System (CTS); Common Language Specification (CLS); Common Intermediate Language(CIL/MSIL);Types of JIT Compilers; Garbage Collection; Base Classes Framework; Assemblies and Modules.

C#: Data types – Operators and Expressions - Statements – Decision Making – Loops – Arrays and Collections – Strings ; Object Oriented Programming - Encapsulation, Polymorphism and Inheritance in C#.

(10 hrs)

Unit – II

ASP.NET: ASP.NET Architecture;Programming with ASP.NET Web pages, ASP.NET Controls, Web Parts; Master pages; Applying Themes and Styles to Controls; ASP.NET State Management; Caching;

ASP.NET AJAX Components of Microsoft's AJAX Support; Partial Page Update; Script Manager; Client-Side Page Lifecycle; Update Panel Control; Update Modes; Timer.

(10 hrs)

Unit – III

Data Access using ADO.NET, LINQ, Entity Framework: The ADO.NET Object Model; Managed Providers; Working with Connected and Disconnected Architectures.

LINQ; Using Standard 'LINQ to' Technologies namely: LINQ to Object, LINQ to Dataset, LINQ to SQL, and LINQ to XML.

+Entity Framework: Entity Data Model, Navigating the EF, Querying Data, CRUD Operations in EF. **(9hrs)**

Unit - IV

ASP.NET MVC: MVC Overview; Advantages of using MVC Model; ASP.NET MVC application Structure; Creating the Model; Creating a Controller; Adding a View; Razor View Engine; Razor Syntax, Variables, Loops and Logic; Specifying a Partial View; Models, Layout; Scaffolding , Databases; Custom Routing ASP.NET MVC and Security.

Web Services: Overview; Working of a Web Service; Creating a Web Service using ASP.NET; Consuming WebService. **(10hrs)**

Unit – V

ASP.NET Core 3.: Introduction; Routing; Controllers and Actions, API actions, Maintaining State; Views, Areas, Razor Pages; Forms and Models, Helpers, Model Binding,Minifications, Validations and Data Annotations; Security; Caching; Module Development,Filters.

(9hrs)

Text Books:

[1]. Andrew Trolesen, "Professional C# 6.0 and the .NET 4.6 Framework", 7th Editon, 2015, Wiley press.
[2]. Adam Freeman, "Pro ASP.Net MVC 5th Edition,2014, Wiley press.
[3]. Kenneth Yamikani Fukizi, Jason De Oliveira, "Learn ASP.NET Core 3",2nd Edition, 2019, Packt Publishing.

Reference Books:

[1].   John Sharp, "Microsoft Visual C# Step by Step", 9th Edition, 2018, Microsoft press.[2]. Matthew MacDonald, "Asp.Net: The Complete Reference", 2016, Tata McGraw-Hill Education.

[3].   William Penberthy "Beginning ASP.NET for Visual Studio, 2016, Wiley Wrox press.

[4].   Adam Freeman, "Pro ASP.Net MVC 2", 2017, Wiley press.

[5].   Jon Galloway, Brad Wilson, K. Scott Allen, David Matson "Professional ASP.NET MVC 5",2014, Wiley Wrox press.

[6].   Bruce Johnson, "Professional Visual Studio - AJAX", 2015, Wiley Wrox press.[7]. Adam Freeman, "Pro ASP.NET Core 3", 2020, Apress.

PS 534.3 E2 NATURAL LANGUAGE PROCESSING

Total No. of Lectures: 42   **Total Marks:100**   **[L – T – P – S]**
No. of Lectures / Week: 4   **Credits   : 3**   **[3 – 1 - 0 - 2]**

Course Objectives: As part of this course, students will study,
- To understand about NLP. To know its origin. To know about different Indian languages and grammar of languages in general. Knowledge about regular expressions and finite state automata. Understand about morphology and parts-of-speech tagging
- To learn about Context-Free Grammar (CFG). Understand about syntactic and semantic analysis
- To understand Cohesion, Reference resolution, Discourse Coherence and structure.
- To understand the architectures of NLG systems, generation tasks and representations, applications of NLG
- To understand some of the NLP applications like MT, IR, automatic text summarization, question-answering system

Course Outcomes: upon successful completion of this course, the participant will be able to:
- Ability to create morphemes and perform morphological analysis. Construct simple DFA. Perform POS tagging
- Ability to construct parse trees for sentences when CFG is given. Perform leftmost and rightmost derivations. Perform top-down and bottom-up parsing. Perform ambiguity analysis and word sense disambiguation.
- Perform reference resolution on sentences. Differentiate Cohesion and Coherence.
- Differentiate pipelined, interleaved and integrated architecture of NLG.
- Compare direct MT system with transfer system. Implement a simple MT system.

Unit - I

Introduction: What is Natural Language Processing? Origins of NLP, Language and Knowledge, Language and Grammar, Processing Indian Languages, Early NLP systems.

Word Level Analysis: Regular Expressions, Finite-State Automata, Morphological parsing, Words and Word Classes, Part-of-Speech Tagging

(9 hrs)

Unit - II

Syntactic Analysis: Context-Free Grammar, Phases of Syntactic Analysis, Constituency, Parsing – Top-down, Bottom-up parsing.

Semantic Analysis: Meaning representation, Lexical semantics, ambiguity, word sense disambiguation

(9 hrs)

Unit - III

Discourse and Reference Analysis: Reference Resolution, the structure of coherent discourse, types of reference phenomenon, Syntactic Constraints on Reference - Number agreement, Gender agreement, grammatical role etc.,

(8 hrs)

Unit - IV

Natural Language Generation: Introduction, Architectures of NLG systems, Generation tasks and representations, Applications of NLG.

(8hrs)

Unit - V

NLP Applications: Machine translation – characteristics of Indian languages, MT types, Information retrieval, automatic text summarization, question-answering system.

(8hrs)

Text Books:

[1]. Tanveer Siddiqui, U.S. Tiwary, "Natural Language Processing and Information Retrieval", Oxford Higher Education, 2008, Oxford University Press

[2]. Ela Kumar, "Natural Language Processing", 2013, I.K. International Publishing House Pvt. Ltd.

Reference Books:

[1].Daniel Jurafsky and James H. Martin, "Speech and Language Processing An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition", 2000, Pearson Education Publication.

[2].Michael W. Berry, "Survey of Text Mining: Clustering, Classification and Retrieval", 2007, Springer International Edition

[3]. Daniel M Bikel, ImedZitouni, "Multilingual Natural Language Processing Applications", 2013, Pearson Education Publication.

[4]. Steven Bird, Ewan Klein, and Edward Loper, "Natural Language Processing with Python" 2009, O'Reilly Media.

PS535.3 E1 BIG DATA ANALYTICS WITH SCALA AND SPARK

Total No. of Lectures : 42	**Total Marks :**	**100**	**[L – T – P – S]**
No. of Lectures / Week : 4	**Credits :**	**3**	**[3 – 1 - 0 - 2]**

Course Objectives: This course introduces students to the following objectives

- Understanding the concepts of Functional programming and introducing to Scala and object oriented Scala.
- Elaborating on the Spark's programming model in detail, being careful to understand how and when it differs from familiar programming models, like shared-memory parallel collections or sequential Scala collections.
- Understanding Spark SQL and exploring the use of Spark SQL for Data Exploration, Data Munging and Data streaming.
- Using Spark in Machine Learning for classification using Bayes and Naïve Bayes Algorithm.
- Learn to perform Testing in Spark and introducing PySparrk and Sparkr.

Course Outcomes: Upon successful completion of this course, the student will be able to:

- Understand what Functional programming is and will know why classical data analysis techniques are no longer adequate
- Understand the benefits that Spark and Spark SQL offers for processing structured and unstructured data.
- Understand conceptually how Spark SQL is used for Data Exploration, Data Munging and Data Streaming.
- Understand how Spark can be used for Machine Learning.
- Understand the use of PySparrk and Sparkr

Unit- I

Scala: Purposes of Scala, Platforms and editors, Installing and setting up Scala, Scala: the scalable language, Scala for Java programmers, Scala for the beginners

Object Oriented Scala: Variables in Scala, Methods, classes, and objects in Scala, Packages and package objects, Java interoperability, Pattern matching, Implicit in Scala, Generic in Scala, SBT and other build systems.

Functional Programming Concepts: Introduction to functional programming, Functional Scala for the data scientists, FP and Scala for learning Spark; Pure functions and higher-order functions; Using higher-order functions; Error handling in functional Scala; Functional programming and data mutability

Collection of APIs : Scala collection APIs, Types and hierarchies, Performance characteristics, Java interoperability, Using Scala implicits

(9 hrs)

Unit- II

Spark : Introduction to data analytics, Introduction to big data, Distributed computing using Apache Hadoop, Apache Spark.

Spark – REPL & RDDS: Dig deeper into Apache Spark, Apache Spark installation, Introduction to RDDs, Using the Spark shell, Actions and Transformations, Caching, Loading and saving data.

Spark SQL: Introducing Spark Session, Understanding Spark SQL concepts, Using Spark SQL in streaming applications; Spark SQL and DataFrames, DataFrame API and SQL API, Aggregations, Joins.

Spark SQL for Processing Structured and Unstructured Data: data sources in Spark applications, Spark with relational databases, Spark with MongoDB (NoSQL database), Spark with JSON data, Spark with Avro files, Spark with Parquet files, Defining and using custom data sources in Spark

(9 hrs)

Unit- III

Spark SQL for Data Exploration: Exploratory Data Analysis (EDA), Spark SQL for basic data analysis, Visualizing data with Apache Zeppelin, Sampling data with Spark SQL APIs, Spark SQL for creating pivot tables

Spark SQL for Data Munging: Data munging, Exploring data munging techniques, Munging textual data, Munging time series data, Dealing with variable length records, Preparing data for machine learning

Spark Streaming: Spark Streaming, Discretized streams, Stateful /stateless transformations Check pointing, Interoperability with streaming platforms (Apache Kafka), Structured streaming.

(8 hrs)

Unit- IV

Spark MLLIB & ML: Introduction to machine learning, Spark machine learning APIs, Feature extraction and transformation, Creating a simple pipeline, Unsupervised machine learning, Binary and multiclass classification

Bayes and Nave Bayes: Multinomial classification, Bayesian inference, Naive Bayes, The decision trees

Text Analytics using Spark ML : Understanding text analytics, Transformers and Estimators, Tokenization, StopWordsRemover, NGrams, TF-IDF, Word2Vec, CountVectorizer, Topic modeling using LDA, Implementing text classification.

(8 hrs)

Unit- V

Spark Tuning: Monitoring Spark jobs, Spark configuration, Common mistakes in Spark app development, Optimization techniques;

Testing and Debugging in Spark: Testing in a distributed environment, Testing Spark applications, Debugging Spark applications.

PySparrk&Sparkr: Introduction to PySpark, Installation and configuration, Introducing SparkR, the SparkR architecture, SparkRDataFrames, SparkR for EDA and data munging tasks, SparkR for computing summary statistics, SparkR for data visualization, SparkR for machine learning. **(8 hrs)**

Text Book :

[1]   Md. Rezaul Karim, Sridhar Alla, "Scala and Spark for Big Data Analytics - Harness the power of Scala to program Spark and analyze tonnes of data in the blink of an eye", 1st Ed, 2017, PACKT

[2]   Romeo Kienzler, "Mastering Apache Spark 2.x", 2nd Edition, 2016, PACKT Publishers.

Reference Books :

[1]   Ilya Ganelin, EmaOrhian, Kai Sasaki, Brennon York, "Spark: Big Data Cluster Computing in Production", 1st Edition, 2016,Wiley

[2]   Rishi Yadav, "Apache Spark 2.x Cookbook", 2nd Edition, 2016, PACKT Publishers

[3]   Dean Wampler, Alex Payne, "Programming Scala, Scalability = Functional Programming + Objects", 2nd Edition, 2016, O'Reilly Publishers

[4]   Jason Swartz, "Learning Scala Practical Functional Programming for the JVM", 2016, O'Reilly

[5]   Guller, Mohammed, "Big Data Analytics with Spark A Practitioner's Guide to Using Spark for Large Scale Data Analysis", 2nd Edition, 2016, Apress

[6]   Luu, Hien, "Beginning Apache Spark 2 With Resilient Distributed Datasets, Spark SQL, Structured Streaming and Spark Machine Learning library", 2nd Ed, 2016, Apress Edition

[7]   Nabi, Zubair, "Pro Spark Streaming The Zen of Real-Time Analytics Using Apache Spark", 2nd Edition, 2017, Apress Edition.

[8]   MateiZaharia, Holden Karau, Andy Konwinski, Patrick Wendell, "Learning Spark Lightning-Fast Big Data Analysis", 2nd Edition, 2017, O'Reilly Publishers

[9]   MateiZaharia, Bill Chambers, "Spark: The Definitive Guide Big Data Processing Made Simple", 3rd Edition, 2106, O'Reilly Publications

[10]   Paul Chiusano, "Functional Programming in Scala", 2nd Edition, 2017, Wiley

[11]   JanekBogucki, Alessandro Lacava, AliaksandrBedrytski, Matthew de Detrich, Benjamin Neil, "Professional Scala", 2nd Edition, 2016, Wiley Wrox