DSEU

Delhi Skill and Entrepreneurship University

B.Sc.
(Data Analytics)

Syllabus Document

Effective from Academic Year 2021-22

Program Information

Introduction

In today's dynamic business landscape, the ability to drive corporate business decisions through data-driven insights leads to market leadership and gives organisations a competitive edge. Increasingly, the integration of data analytics and science is proving to be a game-changer for businesses across industries. Organisations are seeking to gain revenue-generating insights from data and build future strategies for business growth. We strongly believe that in the coming days no business will survive without Data and Business Analytics, making knowledge of analytics an indispensable and universally desired skill set.

Program Objective

Delhi Skill and Entrepreneurship University (DSEU) Professional Certificate Programme in Data Analytics (*exit after first year*), Diploma Program in Data Analytics (*exiting after second year*) and B. Sc. in Data Analytics (*on completion all 3 years*) graduate program provides strong and sustainable human resources to corporates to meet their needs on data management, wrangling, storage, exploratory data analysis, automation, predictive and prescriptive analytics, and the use of machine learning and artificial Intelligence algorithms. Along with a very strong business problem-solving skills and technical understanding, the programme will help professionals enhance their proficiency in data science and gain in-depth skills and robust knowledge of machine learning and artificial Intelligence algorithms and techniques supported by Python, Excel, Tableau, Power BI, IoT and AI-ML using cloud resource including GPUs and TPUs.

Since most of the data available today is in semi structure, unstructured, streaming image and video format, and is ever growing at an exponential rate, joining this program will train professionals to gain skills and techniques such as text mining and social media analytics that are vital for maximising business growth and transformation.

Pedagogy & Teaching methodology

Additionally, this program's effective pedagogy (*developed by who's who of Data Science and Analytics experts from academia from IITs, IIM, ISB and various prestigious IT industries like CISCO, Cognizant, and Honeywell*) and focus on real-world examples and hands-on projects from corporates, case studies, and practical sessions will assist in identifying data insights and making high-output business decisions. Apart from technical skills, the program focuses on deep integration of soft skills, or as we call them 'Face the work skills (FTW)', across all semesters. These FTW skills encompass communication skills, digital literacy, professional development, workplace behaviour and career development, just to name a few, and prepare you to enter the corporate world as well-groomed professionals.

Student Outcomes (SOs)

The Program Educational Objectives of B. Sc Data Analytics Programs is to produce graduates who would:

- Establish themselves as Business Analysts, Data Analysts, Artificial Intelligence and Machine Learning scientists and subject matter experts in various private and public sectors that are

involved in the design, creation, maintenance and use of industrial and organization data and help nation building.

- Solve real world business problems by applying knowledge ethically that will benefit organizations and society at large.
- Adapt to changing trends in Data Science, Business Analytics, Artificial Intelligence and Machine Learning and become lifelong learners.

Placement and Internship

With strong industry partnership since the inception of this course, design and development of curriculum of this prestigious program called B. Sc Data Analytics focusing mostly on industry desired skill development, students have a very high chance to get much required industry internship experience and build their strong career for sure. After completing this program, you will join as a data analyst with compensation ranging from INR 20,000 to 25,000 in MNCs such as Infosys, TCS, Wipro, IBM, etc. One can further specialize in niche domains by enrolling in a master's program either in India or abroad.

Credit scheme

Semester V							
Sl No.	Course Code	Course Name	Hours/week				Total Credits
			L	T	P	Total	
1	DTA-DC501	Model Deployment and Cloud Management	3	0	2	5	4
2	DTA-DC502	Time Series and Forecasting	3	0	2	5	4
3	DTA-DC503	Cyber Security	3	0	0	3	3
4	DTA-ECX	Elective 1	3	0	2	5	4
5	DTA-ECXX	Elective 2	3	1	0	4	4
6	AE-025	English for success (Business English Communication skills)	2	0	0	0	2
7	DTA-DC504	Major Project	0	0	4	2	2
Total			**17**	**1**	**10**	**29**	**23**

**Students to select one course from each category of electives*

List of Electives (For Semester V)	
Course Code	**Course Titles**
Elective-I - ECX	
DTA-EC501	Advanced Data Structure & Algorithms
DTA-EC502	JAVA Programming
DTA-EC503	Artificial Intelligence
DTA-EC504	Audio Visual Analytics
DTA-EC505	Deep Learning
DTA-EC506	Recommendation System
Elective-II- ECXX	
DTA-EC507	Graph Theory
DTA-EC508	Ambient Intelligence
DTA-EC509	Data Warehouse and Data Mining
DTA-EC510	BFSI Analytics
DTA-EC511	Operations and Supply Chain Analytics

DSEU

SEMESTER V

DTA-DC501 | MODEL DEPLOYMENT AND CLOUD MANAGEMENT

Teaching Scheme			
Lecture Hours per week	**Tutorial hours per week**	**Practical hours per week**	**Credit**
3	**0**	**2**	**4**

Prerequisite:

Basic of Machine Learning, Introduction to Python Programming

Objectives:

The entire course goes over several machine learning-based methods which are trained on the input data. Deployment of the model is an important activity once the model is developed, and this course would introduce the students to the steps involved in model deployment - both on a standalone machine or a cloud instance.

Learning Outcomes:

The learner of this course would be able to conduct the following activities:

1. Distinguish between development and production environment
2. Understand the overall system architecture
3. Deploy models in production environment
4. Deploy models in cloud environment

Syllabus:

Unit-I Introduction to Model Deployment (6 hrs)

Difference from other ML steps, Importance and usage of ML models, Research and development

Testing and Validation, Production, Hardware and workload requirements for each environment Machine Learning System Architecture, Feature Engineering, Training and Prediction

Unit-II Machine Learning Pipeline & Components (7 hrs)

Data Ingestion, Data Preparation, Data Segregation - Training vs. Testing, Model Development, Model Evaluation and Model Deployment. Model Deployment with Open-Source Tools: Working with *tox* package, Working with CI/CD pipelines

Unit-III Packaging model for production (7 hrs)

Feature engineering as pre-processing, Package configurations, Building the package, Model prediction in run time. Deployment Server and API: API Architecture, Working with FastAPI, Web Server environment, Deployment of the application

Unit-IV Containers and Docker: Value of Container and Docker (8 hrs)

Container Deployment Process, Performance Monitoring, Time evaluation for run time model , Time evaluation with multiple concurrent users, Optimization for performance

Unit-V Deployment on Cloud environment **(8 hrs)**

Cloud environment Introduction, Cost, instance and other aspects of cloud, Logging onto cloud environment, Deployment on the cloud environment

Unit- VI Management of the production cloud infrastructure **(6 hrs)**

Managing multiple instances on the cloud, Cloud management techniques

Practical/Laboratory Content:

1. Introduction to multiple tools that address ML experiment management and collaboration for data scientists, such as Neptune.ai, Weights & Biases, Comet.ml, and mlflow.
2. Work through different phases of MLOps Operations lifecycle to put machine learning models into production by using a sample real-time cloud dataset.
3. Start an ML project and learning how to go about ML Development and Create a directory. Make use of Azure.
4. Implementation of steps for Model Building and Training, Training Operationalization and Model Versioning.
5. Learn to deploy to Azure using terraform. To test databricks+mlflow simply run the CLI command terraform init && terraform apply. Writing scripts to execute.
6. Learning to work on CI/CD pipleline for End-to-End Pipeline Development, Deployment, and MLOps.
7. Practical hands-on to build and deploy a web application by working through the process of Train and develop a machine learning pipeline for deployment (simple linear regression model).
8. Build a web app using the Flask framework. It will use the trained ML pipeline to generate predictions on new data points in real-time (front-end code is not the focus of this tutorial).
9. Create a Docker image and container. Publish the container onto the Azure Container Registry (ACR). Deploy the web app in the container by publishing onto ACR. Once deployed, it will become publicly available and can be accessed via a Web URL.
10. Students must recognise the difference in workflow of MLOps from DevOps.

Lab resources & reading material:

1. Building Machine Learning Powered Applications: Going from Idea to Product - Emmanuel Ameisen -February 11, 2020

2. "Cloud Native DevOps with Kubernetes: Building, Deploying, and Scaling Modern Applications in the Cloud" by John Arundel - 22 March 2019

3. Practical Data Science with Docker: Build, Ship, Run, and Scale Successful Data Science Applications with Docker" by Joshua Cook -19 December 2017

DTA-DC502 | TIME SERIES AND FORECASTING

Teaching Scheme			
Lecture hours per week	**Tutorial hours per week**	**Practical hours per week**	**Credit**
3	**0**	**2**	**4**

Prerequisite:

Advanced Data Analytics, Excel, and Programming

Objectives:

In the Time Series and Forecasting course, we look at data sets that represent sequential information, such as stock prices, annual rainfall, sunspot activity, the price of agricultural products, and more. We look at several mathematical models that might be used to describe the processes which generate these types of data. We also look at graphical representations that provide insights into our data. Finally, we also learn how to make forecasts that say intelligent things about what we might expect in the future. The language for the course is Python, a free implementation language. It is a professional environment and fairly easy to learn. You can discuss material from the course with your fellow learners. Time Series and Forecasting can take effort to learn - we have tried to present those ideas that are "mission critical" in a way where you understand enough of the math to feel satisfied while also being immediately productive.

Learning Outcomes:

After completion of the course students will be able to:

1. Visualize the time series data including ACF and PACF graphs along with trend, seasonality and will be able to distinguish stationarity and nonstationary.
2. Understand Random walk / basics of Stochastic process.
3. Should be able to simulate and generate time series data.
4. Should be able to decompose all the time series components
5. Should be able to model simple AR and MA model (with ARMA model)
6. Understand difference and back shift operators to get an Integrated process with ARMA
7. Finally, should be able to use SARIMA process and should be able to forecast time series,

Syllabus:

Unit I: Basic Statistics (prerequisites) and Introducing TSA: (**6 hrs)**

Descriptive Statistics, time series data visualization, Hypothesis testing related to time series data. Introducing Time Series Analysis (TSA), Significance and Types of Time Series Analysis.

Unit II: Exponential smoothing (or first, second and third order) models (**7 hrs)**

Components of Time Series Analysis, Models- Simple Exponential Smoothing, Double Exponential Smoothing, Triple Exponential Smoothing. Level, trend and seasonality understanding with respect to exponential smoothing and forecasting. Hyper parameters for smoothing

Unit III: Visualizing Time Series and Beginning to Model Time Series: (**7 hrs)**

Exploring and visualizing time series on acquired data sets with respect to Decomposition models, Moving average, autoregressive order ARIMA

Unit IV: Demonstration of various time series data (**8 hrs)**

Demonstration of various time series data (from time series data library - TSDL) and their visualization, Data Types of Time Series, Methods to Check Stationarity, Converting Non-Stationary into Stationary, Case studies - Explore a simple time series data on excel and understand stationarity, Moving Average component, Seasonality component and forecasting taking a product sales for 4 years with some seasonality and forecast 5th year's sales.

Unit V: Time Series Models and Stationarity **(8 hrs)**

Stationarity, MA(q) model building in detail, AR(p) processes and model building, Stationarity, Backward shift operator, Invertibility, and Duality, Autoregressive processes and Yule-Walker equations, Akaike Information Criterion (AIC), Mixed Models & Integrated Models such as such as ARMA, ARIMA and model few real-world datasets.

Unit VI: Seasonality, SARIMA, Forecasting: **(6 hrs)**

SARIMA, Fit SARIMA models to various datasets and start forecasting, Project on time series

Practical/Laboratory Content:

1. Time series data visualization from TSDL (time series data library)
2. Work on a forecasting case studies using excel to understand all the time series components
3. Simulate simple MA model with intuition and simulate simple AR model with intuition
4. Combine both above model to get ARMA model with intuition
5. Comnite integrated / difference component and its intuition
6. Case study to work on 2 to 3 time series datasets with complete ARIMA models
7. Understand SARIMA model and its application with few mode cases of or above already used time series datasets.
8. Few assignments need to be tried by students (with fresh data from tsdl)
9. A presentation by groups of students (max 5 in each group) to be arranged to understand their knowledge about this subject.

Required Readings:

1. “Time Series Analysis and its application, by Robert H. Shumway and David S. Stoffer, 4th edition, Springer publication, 2018

Recommended Readings:

1. Introduction to Time Series and Forecasting, by Peter J. Brockwell and Richard A Davis, Springer publication, 3rd edition-2016
2. A course in Time Series Analysis, by Suhasini Subba Rao, Email: suhasini.subbarao@stat.tamu.edu, January 17, 2021
3. An Introductory Study on Time Series Modeling and Forecasting, Ratnadip Adhikari, R. K. Agrawal 2013

Lab resources & reading material:

1. Python programming environment with astsa, forecast, tsdl packages
2. Alteryx software (free download for a month)
3. Excel and Tableau also can be used for visualization.
4. Excel and Tableau also can be used for visualization.

DTA-DC503| CYBER SECURITY

Teaching Scheme			
Lecture hours per week	**Tutorial hours per week**	**Practical hours per week**	**Credit**
3	**0**	**0**	**3**

Prerequisite:

Basics of Computing and foundational knowledge of programming .

Objectives:

The objective of this course is to teach the students about different types of security attacks and how to handle them.

Learning Outcomes:

At the end of this course, the student should be able to:

1. Understand how to protect a Microsoft word document.
2. Be familiar with how to password protect Microsoft word documents in different types of operating system.
3. Understand the steps of operation how to remove password from Microsoft Word 2007.
4. Know how to hack a simple or a strong password.
5. Know the different types of hacking process and type of applications

Syllabus:

UNIT 1: Introduction to information systems **8 HRS**

Types of information Systems, Development of Information Systems, Introduction to information security, Need for Information security, Threats to Information Systems, Information Assurance, Cyber Security, and Security Risk Analysis, TCP/IP Protocol Suite.

UNIT 2: Application Security & it's technology **12 HRS**

Application security (Database, E-mail and Internet), Data Security Considerations-Backups, Archival Storage and Disposal of Data, Security Technology-Firewall and VPNs, Intrusion Detection, Access Control.

Security Threats -Viruses, Worms, Trojan Horse, Bombs, Trapdoors, Spoofs, E-mail viruses, Macro viruses, Malicious Software, Network and Denial of Services Attack, Security Threats to E-Commerce-Electronic Payment System, e- Cash, Credit/Debit Cards. Digital Signature, public Key Cryptography.

12 HRS

UNIT 3: Development and Design of Security Issues

Developing Secure Information Systems, Application Development Security, Information Security Governance & Risk Management, Security Architecture & Design Security Issues in Hardware, Data Storage & Downloadable Devices, Physical Security of IT Assets, Access Control, CCTV and intrusion Detection Systems, Backup Security Measures.

10 HRS

UNIT 4: Security Policies

Why Policies should be developed, WWW policies, Email Security policies, Policy Review Process-Corporate policies-Sample Security Policies, Publishing and Notification Requirement of the Policies.

Information Security Standards-ISO, IT Act, Copyright Act, Patent Law, IPR. Cyber Laws in India; IT Act 2000 Provisions, Intellectual Property Law: Copy Right Law, Software License, Semiconductor Law and Patent Law. **12 HRS**

Required Readings:

Textbooks:

TB 1: Analyzing Computer Security by Charles P. Pfleeger, Shari Lawerance Pfleeger – 25 August 2011

TB2: V.K. Pachghare, "Cryptography and information Security", PHI Learning Private Limited, Delhi India - 3rd Edition.

Reference Books:

RB 1: Dr. Surya Prakash Tripathi, Ritendra Goyal, Praveen Kumar Shukla," Introduction to Information Security and Cyber Law" Willey Dreamtech Press. - 1 January 2014

RB 2: Schou, Shoemaker, "Information Assurance for the Enterprise", Tata McGraw Hill. - 2006 Edition

RB 3. RB 3. CHANDER, HARISH," Cyber Laws And It Protection " , PHI Learning Private Limited ,Delhi ,India - 2012

Recommended Online Resources:

1. https://heimdalsecurity.com/pdf/cyber_security_for_beginners_ebook.pdf
2. http://larose.staff.ub.ac.id/files/2011/12/Cyber-Criminology-Exploring-Internet-Crimes-and-Criminal-Behavior.pdf
3. http://docshare04.docshare.tips/files/21900/219006870.pdf
4. http://index-of.co.uk/Hacking-Coleccion/Insider%20Attack%20&%20Cyber%20Security%20-%20Beyond%20the%20Hacker.pdf
5. http://www.uou.ac.in/sites/default/files/slm/FCS.pdf
6. https://cyber-cops.com/book_detail

DTA-EC501 | ADVANCED DATA STRUCTURE & ALGORITHMS

Teaching Scheme			
Lecture hours per week	**Theory hours per week**	**Practical hours per week**	**Credit**
3	**0**	**2**	**4**

Prerequisite:

Basics of Machine Learning

Objectives:

The Advanced Data Structures are the most important course which helps the student in competitive programming where time complexity of an algorithm is the key criteria to get best possible Data Structure for a particular algorithm with least time complexity. This course nurtures the Data Scientist to implement the model in an efficient way. This course intuits the student for complex problem-solving techniques by giving hands on practice of different variants of Data Structures like link Lists, Generic Trees, AVL Trees, Red Black Trees, Augmented Trees, Graphs, Backtracking, Selection Algorithms, String Matching Algorithms, priority queues etc

Learning Outcomes:

1. The student will be able to understand the different problem-solving techniques i.e., Recursion and Backtracking.
2. The student will be able to understand Complex Link List Data Structures and will be able to implement various applications using different variants of Link List.
3. Students have a thorough hands-on practice of different types of tree implementation of data storing like Generic Trees, Expression Trees, XOR Trees, Binary Search Trees, AVL Trees, Red Black Trees, B Trees, Augmented Trees and their implementation for real world applications.
4. Students will have a deep understanding of graphs and graph related algorithms.
5. Students will have a thorough understanding of Disjoint Sets, Priority keys and binary heaps.
6. Students will have hands-on practice of different selection algorithms, string algorithms.
7. The advanced data structure will help the students to not only understand different data structure algorithms but to also make them able to analyze the data structure and algorithms on the basis of time complexity and they will be able to optimum codes of different real-world applications.

Syllabus:

Unit I: Data Structures and Problem-Solving techniques **(7 hrs)**

Asymptotic Analysis, Recursion and Backtracking, Introduction to problem solving approaches: Greedy Algorithms, Dynamic Programming approaches.

Unit II: Linear Data Structures **(9 hrs)**

Circular Link List, Doubly Link List, Comparison of Link List with array and dynamic array, Priority Queues. String algorithms, Brute Force Method, Rabin-Karp String Matching Algorithms, String matching with Finite automata, KMP Algorithm, Boyer – Moore Algorithm.

Unit III: Trees and Heap Data Structures **(9 hrs)**

Generic Trees (N-ary Trees), Expression Trees, XOR Trees, Binary Search Trees, Balance Binary Search Trees, AVL (Adelson – Velskii and Landis) Trees, Red Black Trees, B – Trees, Augmented Trees, Segment Trees, Heaps and Binary Heaps.

Unit-IV: Graph Data Structure **(10 hrs)**

Graphs – Graph Traversals, Depth for search and breadth for search, shortest path algorithm, Shortest path in Unweighted graph, Shortest path in weighted graphs (Dijkstra's), Minimal Spanning Trees, Prim's Algorithm, Kruskal's Algorithm.

Unit-V: Sorting, Searching and Hash Tables **(7 hrs)**

Selection Algorithms – Selection by sorting, Partition based selection, Linear selection algorithm – Median of Medians Algorithm. Searching- Linear and Hash Tables, Data Structure for storing strings: Hashing tables

Practical/Laboratory Content:

1. Write both recursive and non-recursive functions for implementing the following searching methods: (a) Linear search (b) Binary search
2. Write a program to implement the following a) Stack ADT b) Queue ADT
3. Write a program that reads an infix expression and converts the expression to postfix form. (Use stack ADT).
4. Write a program that uses both a stack and a queue to test whether the given string is a palindrome or not.
5. Write a program to implement the following using a singly linked list. a) Stack ADT b) Queue ADT
6. Write a program to perform the following operations:
   a) Construct a binary search tree of elements.
   b) Search for a key element in the above binary search tree.
   c) Delete an element from the above binary search tree.
7. Write a program to implement all the functions of a dictionary (ADT) using Hashing.
8. Write programs that use recursive and non-recursive functions to traverse the given binary tree in a)Preorder b) Inorder c) Postorder
9. Write programs for the implementation of bfs and dfs for a given graph.
10. Write programs for implementing the following sorting methods: a) Bubble sort b) Insertion sort c) Quick sort d) Merge sort e) Heap sort f) Radix sort g) Binary tree sort
11. Write programs to perform the following operations: a) Insertion into a B-tree b) Searching in a B-tree

Required Readings:

1. Introduction to Algorithms – Thomas H Cormen (The MIT Press). 4th edition ,5apr 2022
2. An Introduction to Data Structures with Applications. by Jean-Paul Tremblay & Paul G. Sorenson Publisher-Tata McGraw Hill. 2nd edition,1984.
3. Fundamentals of Computer Algorithms by Horowitz, Sahni,Galgotia Pub. 2008 2ed.

DTA-EC502| JAVA PROGRAMMING

Teaching Scheme			
Lecture hours per week	**Theory hours per week**	**Practical hours per week**	**Credit**
3	**0**	**2**	**4**

Pre-requisites:

Knowledge of basic programming.

Objectives:

The objective of this course is to introduce basic concepts of object oriented and platform independent programming language and to demonstrate skills in writing programs using concepts like exception handling techniques and multithreading.

Learning Outcomes:

At the end of this course, the student should be able to:

1. Ability to apply object-oriented programming concepts.
2. Understand syntax and semantics of java programming language.
3. Able to use packages, multithreading, exception handling to solve the problems.
4. Able to understand the database connectivity, web applications using java.
5. Able to design event driven GUI and web related applications which mimic the real word scenario.

Syllabus:

Unit 1: **10 hrs**

Introduction: Overview and characteristics of Java, Java program Compilation and Execution Process, Organization of the Java Virtual Machine, Security Promises of the JVM.

Java Language Fundamentals: Data Types & Literals Variables, dynamic initialization, scope and life time, Wrapper Classes, Arithmetic Operators, Logical Operators, Control of Flow.

Classes and Objects: Concepts, methods, constructors, usage of static, access control, this key word, garbage collection, method overloading, parameter passing mechanisms, nested classes and inner classes.

Unit 2: **10 hrs**

Inheritance: Basic concepts, access specifiers, usage of super key word, method overriding, final methods and classes, abstract classes, dynamic method dispatch, Object class.

Interfaces: Differences between classes and interfaces, defining an interface, implementing interface, variables in interface and extending interfaces.

Packages: Creating a Package, setting CLASSPATH, Access control protection, importing packages.

Unit 3: **12 hrs**

Exception Handling: Concepts of Exception handling, types of exceptions, usage of try, catch, throw, throws and finally keywords, Built-in exceptions, creating own exception sub classes.

Strings And Arrays: Exploring the String class, String buffer class, Command-line arguments. Arrays in Java.

Multithreading: Concepts of Multithreading, differences between process and thread, thread life cycle, Thread class, Runnable interface, creating multiple threads, Synchronization, thread priorities, inter thread communication, daemon threads, deadlocks.

I/O Streams: Streams, Byte streams, Character streams, File class, Filestreams.

Unit 4: **10 hrs**

Applets: Concepts of Applets, life cycle of an applet, creating applets, passing parameters to applets, accessing remote applet, Color class and Graphics

Event Handling: Events, Event sources, Event classes, Event Listeners, Delegation event model, handling events.

Required Readings:

Textbooks:

TB 1: Schidt, Herbertz (2022) . Java Complete Reference, 12/e, McGraw Hills Publication.

Reference Books:

RB 1: Balaguruswamy, E. (2019). Programming with Java , 6/e, McGraw Hill Publication.

RB 2: Horstmann, C. S. (1999). Computing Concepts with Java 2 Essentials. John Willey and Sons Publication.

Practical/Laboratory Contents :

Lab Assignment-1 (Basic Java Programming)

1. WAP to find factorial of a given number
2. WAP to print Fibbonacci series for a given number
3. WAP to print grades for given input percentage from command line using nested if-else.
4. WAP to find maximum of three numbers.
5. WAP to check whether a number is prime.
6. WAP to print first n prime nymbers
7. WAP to print series of prime numbers upto n.
8. WAP to print reverse of digits of a given number (for example 123 becomes 321)
9. WAP to find whether a given char is vowel using switch case.
10. WAP to print

```
        1
    1   2   1
1   2   3   2   1
```

Lab Assignment-2 (Classes)

1. WAP to create a Simple Box class that defines three instance variables width, height and depth. Add a method “void volume ()” to compute volume of box. Create two instances and compute their volume.
2. Rewrite program 1 to modify volume method containing return statement.
3. Create a class to compute area of square, rectangle and triangle (use method overloading concept)
4. Add constructor to your box class
5. Show constructor overloading using Box class.
6. Write a program to show use of this keyword.
7. Write a class arithmetic for calculation of addition, subtraction, multiplication and division of two numbers.
8. WAP to demonstrate call by value and call by reference.
9. Write a program to show use of static members and static block.

Lab Assignment-3 (Inheritance and polymorphism)

1. create a base class shape. It contains 2 methods get() and print () to accept and display parameters of shape respectively. Create a subclass Rectangle. It contains a method to display length and breadth of rectangle (Method overrifing)
2. Use Box class. Create a subclass ColoredBox with one parameter String color. Override print method of Box class.
3. Write a program to show Dynamic method dispatch concept
4. WAP to implement stack using arrays in java
5. Write an abtract class Employee with three variables name, sal and Grosssal , suitable constructors, print method and two abstract methods calculategrosssalary() and annualincrement(). Create a manager sub class of employee with Hra as member variable and write implimentation of the abstract method. Also create a subclass of manager as sales manger with commision as member variable and override the calulategrosssalary() method.
6. Write a program to show the use of Interfaces (show multipe inheritence as discussed in class)

Lab Assignement-4 (Exception Handling)

1. Demonstrate Exception handling using single try catch block
2. Show use of multiple catch block
3. Show use of finally block
4. Write a program to show use of throw and trhows keyword. Create your own exception class AgeException. Also create two subclassess TooYoungException and InvalidAgeException of AgeExcep-tion as discussed in class.

Lab Assignment 5 (Packages and Multithreading)

1. Write a program to display the current thread.
2. Write a program to create multiple threads , use Thread class.
3. Write a program to create multiple threads, implementing Runnable Interface.
4. Write a program to create multiple threads by making use of thread priorities.
5. Create a package named as MyPackage with class names as Calculate. The class should contain three methods with the following specifications:

a. Volume: accepts three double type arguments i.e. width, height, depth. Calculate volume and return double type value.
b. Add: which accepts two integer type values, adds them and returns the value.
c. Divide:Accepts two integer type values, divides them and returns results.

6. mport package created in previous program into a file named as PackageDemo and call the above three methods to add, divide and find the volume.

Lab Exercise 6 (Applets ,AWT Controls and Event handling)

1. WAP to create an applet showing different AWT controls like buttons,check box,list box,text box, choice,scroll bar,labels
2. WAP to demonstrate events to buttons, checkbox,checkbox group, text box, Labels.
3. WAP for key Event Demonstration
4. WAP to show Mouse Evewnts using MouseListeners and MouseMotion Listerners
5. WAP to demonstrate a calculator functioning as addition,multiplication,division and subtraction using Button and Label controls and Button Events.Program to demonstrate

Lab Exercise 7 (Layouts, Graphics and Menus)

1. WAP to Design an menu driven program which draw different shapes
2. WAP to show use of different layouts
    a. Border
    b. Card
    c. Flow
    d. Gridlayout
3. Programs to create and display an image
6. Program to demonstrate menu containing a list of movie channels, a list of car names and a list of prices of different books.(Movie Chanels Car names books)
7. Program to draw rectangle,line,polygon,circle,ellipse,arc also fill them with colour.
8. Program for copying contents of one file into another file.

Lab Exercise 8 (I.O. & JDBC)

1. Program or console input output.
2. Program to show JDBC connectivity

DTA-EC503 | ARTIFICIAL INTELLIGENCE

Teaching Scheme			
Lecture hours per week	**Theory hours per week**	**Practical hours per week**	**Credit**
3	**0**	**2**	**4**

Prerequisite:

Python Programming, Data Structure and Algorithm

Objectives:

AI is transforming how we live, work, and play. By enabling new technologies like self-driving cars and recommendation systems or improving old ones like medical diagnostics and search engines, the demand for expertise in AI and machine learning is growing rapidly. This course will enable you to take the first step toward solving important real-world problems and future-proofing your career. Artificial Intelligence explores the concepts and algorithms at the foundation of modern artificial intelligence, diving into the ideas that give rise to technologies like game-playing engines, handwriting recognition, and machine translation. Through hands-on projects, students gain exposure to the theory behind graph search algorithms, classification, optimization, reinforcement learning, and other topics in artificial intelligence and machine learning as they incorporate them into their own Python programs. By course's end, students emerge with experience in libraries for machine learning as well as knowledge of artificial intelligence principles that enable them to design intelligent systems of their own. You will learn the theoretical frameworks that enable these new technologies while gaining practical experience in how to apply these powerful techniques in your work.

Learning Outcomes:

After completion of the course, students would be able to develop an understanding of learning agent, programming for breadth first and depth first search for visiting all possible solution and picking up the most optimum solution, information gathering from data for decision making, graph search algorithms, adversarial search, knowledge representation, logical inference, constraint satisfaction, Bayesian networks, Markov models, machine learning, neural networks, basic natural language processing, basic reinforcement learning.

Syllabus:

Unit-I Fundamentals of Artificial Intelligence (10 hrs)

Introduction: Why AI, History, Relevance to business; Intelligent Systems - Intelligence and its types, Agents and Environments, Brute-Force Search Strategies- Depth first search, Breadth first search. Heuristic Search Strategies

Unit-II Fuzzy Logic Systems (10 hrs)

Introducing Fuzzy Logic, Fuzzy Logic Systems Architecture, Membership Function, Logic Building and conclusion; Uncertainty: Uncertainty and risk in decision making, Application Areas of Fuzzy Logic.

Unit-III Optimization Techniques (12 hrs)

Optimization: Different optimization technique, usage of available library in Python and their usage; Learning: How machine will learn (incremental learning) for decision making; Neural Networks: Linking Neural networks topic taught to be used in AI

Unit-IV Neural Networks in AI (10 hrs)

Neural Networks: Linking Neural networks topic taught to be used in AI; Language: How do you construct language, Backpropagation and Feed forward networks, Types of Backpropagations: Static Back-propagation and Recurrent Backpropagation

Practical/Laboratory Content:

1. Search
   a. Write a program that determines how many “degrees of separation” apart two actors are.
   b. Using Minimax, implement an AI to play Tic-Tac-Toe optimally
2. Knowledge
   a. Write a program to solve logic puzzles - Knights
   b. Minesweeper - Write an AI to play Minesweeper.
3. Uncertainty
   a. PageRank - Write an AI to rank web pages by importance
   b. Heredity - Write an AI to assess the likelihood that a person will have a particular genetic trait.
4. Optimization - Crossword - Write an AI to generate crossword puzzles
5. Learning
   a. Write an AI to predict whether online shopping customers will complete a purchase.
   b. Write an AI that teaches itself to play Nim through reinforcement learning.
6. Neural Network - Write an AI to identify which traffic sign appears in a photograph - Traffic
7. Language
   a. Parser - Write an AI to parse sentences and extract noun phrases
   b. Questions - Write an AI to answer questions

Required Readings:

1. Artificial Intelligence A Modern Approach - Stuart J. Russell , Peter Norvig, Pearson Education, 4th Edition, 2020

2. Ivan Bratko, Prolog Programming for Artificial Intelligence, 4th Ed., Addison-Wesley, 2018.

Recommended Readings:

1. Nils Nilsson, Artificial Intelligence: A New Synthesis, Morgan Kaufmann, 2010.

2. David Poole, Alan Mackworth, Artificial Intelligence: Foundations for Computational Agents, Cambridge Univ. Press, 2018.

3. Artificial Intelligence, Structures and Strategies for Complex Problem Solving, George F Luger, Pearson Education 2019
4. Ronald Brachman, Knowledge Representation and Reasoning, Morgan Kaufmann, 2014.

Lab resources & reading material:

1. CS50: Introduction to Artificial Intelligence with Python - Harvard University (https://learning.edx.org/course/course-v1:HarvardX+CS50AI+1T2020/home)
2. http://aima.cs.berkeley.edu/global-index.html

DTA-EC504 | AUDIO VISUAL ANALYTICS

Teaching Scheme			
Lectures hours per week	**Theory hours per week**	**Practical hours per week**	**Credit**
3	**0**	**2**	**4**

Prerequisite:

Python programming , Mathematics of functions and trigonometric functions, matrices

Objectives:

In today's scenario the social media is full of audio, visual data. This multimedia data can be quantized to extract useful information. The major area of application of Data analytics is audio, video or image data. This course is designed to give the understanding of the basic building blocks of these types of data, that is signal processing using sampling theorem, Discrete Fourier transform, convolution, filters etc. These topics give an insight into how the audio, video, images can be processed to extract the information, features which can be used for classification, segmentation models using CNN, RNN, GANs etc.

Learning Outcomes:

1. The students will understand the mathematics of signal processing using sampling theorem, DFT, FFT, DCT, Convolution theorem.
2. The students should build a thorough understanding to process and synthesize the signals.
3. The students learn the feature extraction techniques. SIFT, SURF, HoG, LBP etc. The students apply the feature extraction on open-source audio, video and image data.
4. Students will be able to understand the Visual matching algorithms and implement these on open-source data.
5. Implement CNN, RNN on audio/video/image data, hands on practice different designs of CNN in Tensorflow.

Syllabus:

Unit-I: Sampling and Transformation techniques: **(9 hrs)**

Revising Reinforcement Learning in comparison to Supervised and Unsupervised models. Sampling Theorem, Discrete Time Fourier Transforms, Noise signals and natural sources of noise, Auto-correlation function for estimating pitch, Discrete Cosine Transform, Fast Fourier Transforms.

Unit-II: Filters and Stochastic Models: **(9 hrs)**

Basic structure for Infinite Impulse Response Filter, Low Pass filter, High Pass Filter and Band Pass filter, Convolution Theorem, Phase unwrapping and Zero Padding. Sinusoidal Synthesis, Stochastic Model, Stochastic Representation of Sound, Spectral based audio features.

Unit-III: Images and Visual Analysis: **(8 hrs)**

Introduction to image formation, Visual Features and representations: Edge, Blobs, Corner Detection; Scale Space and Scale Selection; SIFT, SURF; HoG, LBP, etc. Visual Matching: Bag-of-words, VLAD; RANSAC, Hough transform; Pyramid Matching; Optical Flow

Unit-IV: Convolution Neural Networks: **(6 hrs)**

Introduction to CNNs; Evolution of CNN Architectures: AlexNet, ZFNet, VGG, InceptionNets, ResNets, DenseNets. Introduction to Attention Models in Vision;

Unit-V: Vision and Language **(10 hrs)**

Image Captioning, Visual QA, Visual Dialog; Spatial Transformers; Transformer Networks. Deep Generative Models: Review of (Popular) Deep Generative Models: GANs, VAEs; Other Generative Models: PixelRNNs, NADE, Normalizing Flows, etc. Variants and applications of Generative Models in Vision: Applications: Image Editing, Inpainting, Superresolution, 3D Object Generation, Security; Variants: CycleGANs, Progressive GANs, StackGANs, Pix2Pix etc. Recent Trends: Zero-shot, One-shot, Few-shot Learning.

Practical/Laboratory Content:

1. Exploring Python audio processing and visualizing the spectrogram libraries and extracting features SciPy, pydub, libROSA, pyAudio Analysis
2. Simple audio classification model
3. Exploring Scikit-image, OpenCV, SimpleITK for image and video processing and feature extraction Generating classification model for images and videos.
4. Exploring Tensor flow for CNN, RNN GAN for audio and image classification.

Required Readings:

References for deep learning:

1. Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning, 2016
2. Michael Nielsen, Neural Networks and Deep Learning, 2016
3. Yoshua Bengio, Learning Deep Architectures for AI, 2009

References for computer vision:

1. Richard Szeliski, Computer Vision: Algorithms and Applications, 2010.
2. Simon Prince, Computer Vision: Models, Learning, and Inference, 2012.
3. David Forsyth, Jean Ponce, Computer Vision: A Modern Approach,2nd edition -2012.

Recommended Readings:

1. Deep Learning for Computer Vision https://onlinecourses.nptel.ac.in/noc21_cs93/preview

DTA-EC505 | DEEP LEARNING

Teaching Scheme			
Lecture hours per week	**Theory hours per week**	**Practical hours per week**	**Credit**
3	**0**	**2**	**4**

Prerequisite:

Mathematics for Data Analytics, Probability and Statistics for Data Analytics, Python Programming, Neural Network

Objectives:

Deep Learning presents a simplified explanation of some of the hottest topics in data science today: What is Deep Learning? Why should we learn Deep learning? The course offers learning of following topics:

- Convolutional neural networks (CNN)
- Recurrence Neural Network OR Sequential Neural Network (RNN)
- Autoencoder and decoder
- Boltzmann Machines
- Why is deep learning so powerful and what can it be used for?

Learning Outcomes:

After completion of the course students will be able to:

1. Define what is deep learning and how it is different from Neural Network and statistical learning (Linear regression and Logistic regression)
2. Develop an intuition for why deep learning is the solution to many recent problems
3. Learn about the different deep networks that are available
4. Learn about the different deep learning platforms that are available

Syllabus:

Unit-I: Introduction to Deep Learning **(9 hrs)**

Need of Deep Learning, Introducing a neural network, Basic Neural Network, Biological Neurons Vs Artificial Neurons, Single Layer Perceptron, Multi-Layer Perceptron, Types of propagation networks: Forward and backward propagation and Feed-forward neural networks, Introduction to Activation Function, Types of Activation Function, Introducing Pytorch, Keras and TensorFlow

Unit-II: Artificial Neural Network (ANN) **(9 hrs)**

Cost function in neural networks, How does Gradient Descent work, Vanishing or Exploding Gradients Problems,

Choose the optimal number of epochs, Batch Normalization in Deep Learning, Solving Classification problems with TensorFlow and PyTorch, Solving Regression (Linear) problems with PyTorch and Tensorflow, Fine Tuning & Hyperparameters

Unit-III: Convolution Neural Network (CNN) **(8 hrs)**

Introducing Convolutional Neural Network, CNN for Image Recognition and Image classification, Layers in a Convolutional Neural Network: Convolution layer, ReLU layer, Pooling layer and Fully connected layer, Use case implementation using CNN on CIFAR-10 dataset.

Unit-IV: Recurrence Neural Network (RNN) **(8 hrs)**

Recurrent Neural Network architecture, Sentiment Analysis using RNN, Time Series forecasting using RNN, Short-Term Memory problem in RNN, Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU), Application with Text Generation.

Unit-V: Deep Learning Models **(8 hrs)**

Generative Learning models with AutoEncoder, How Autoencoders works, Types of AutoEncoder, Boltzmann Machines, Deep Belief Nets, Basics of Generative Adversarial Networks (GANs), Thompson Sampling Reinforcement Learning, Introducing Q-Learning

Practical/Laboratory Content:

1. Classification: Simple dataset on cats and dogs to apply and try out Deep Learning technique
2. On the same datasets apply CNN method
3. Use MNIST dataset to try out CNN method
4. Take Indian Stock market dataset to try out RNN
5. Use GRU and LSTM methods
6. Apply Boltzmann Machine methods Images
7. Apply Auto Encoder and Decoder methods to image and document for dimension reduction and reconstruction of original image or doc
8. Apply other methods and learn limitation of these methods
9. Major project

Required Readings:

Dive into Deep Learning, Release 0.16.6, Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola, 2021

Recommended Readings:

1. Deep Learning Adaptive Computation and Machine Learning series, 2016, by Ian Goodfellow, Yoshua Bengio, Aaron Courville
2. Deep Learning with Python, 2017, by François Chollet, Publisher Manning.
3. Coursera course on Deep Learning by Andrew NG (https://www.coursera.org/specializations/deep-learning)

Lab resources & reading material:

- https://d2l.ai/d2l-en.pdf
- http://faculty.neu.edu.cn/yury/AAI/Textbook/DeepLearningBook.pdf
- https://www.cs.tau.ac.il/~dcor/Graphics/pdf.slides/YY-Deep%20Learning.pdf
- http://cs229.stanford.edu/materials/CS229-DeepLearning.pdf
- http://astro.dur.ac.uk/~cmb/Durham.pdf

DTA-EC506 | RECOMMENDATION SYSTEM

Teaching Scheme			
Lecture hours per week	**Theory hours per week**	**Practical hours per week**	**Credit**
3	**0**	**2**	**4**

Prerequisite:

Basic Machine Learning, Fundamentals of Data Analytics

Objectives:

Any online or internet business involving people uses some form of recommendation system. We receive these recommendations during online shopping, viewing movies, browsing books online, etc. This course will expose the learners to the underlying technologies and algorithms used by these recommendation systems.

Learning Outcomes:

Student would be able to perform the following tasks after taking this course:

1. Understand the internal working of recommendation systems
2. Understand the concepts of collaborative filtering and matrix factorization
3. Develop simple recommendation system

Syllabus:

Unit-1: Introduction to Recommendation System **(10 hrs)**

Simple Recommendation System, Profiling customers and products, Aspects of recommendation-Personal attribute based and Content based. Types of Content based Recommendation -Browsing history, Product catalog, Product inventory, etc.

Unit-2: Market Basket Algorithms **(10 hrs)**

Apriori algorithm, Recency, Frequency and Monetary (RFM) attributes and analysis, Classification approaches for recommendation

Unit-3: Collaborative Filtering **(12 hrs)**

Nearest Neighbor based, User Based Collaborative Filtering, Item Based Collaborative Filtering, Evaluation and Metrics for Recommendation Systems, Calculate term-frequency and inverse-document frequency. Create user-item matrix for collaborative filtering using Explicit and Implicit feedbacks.

Unit-4: Developing a recommendation system **(10 hrs)**

Matrix Factorization Methods, Working with real-time Datasets, Algorithmic Fairness, Rating Predictions, Hands-on Recommendation system for business case study.

Practical/Laboratory Content:

1. Practical on simple recommendation system using Python
2. Practical on Market Basket Analysis using Python
3. Practical on RFM Analysis
4. Practical on using Classification for Recommendation
5. Practical exercise on matrix factorization
6. Practical exercise on Collaborative Filtering
7. Practical on working with real world use-cases

Required Readings:

1. Recommender System with Machine Learning and Artificial Intelligence: Practical Tools and Applications in Medical, Agricultural and Other Industries, United States: Wiley, 2020
2. Banik, Rounak. Hands-On Recommendation Systems with Python: Start Building Powerful and Personalized, Recommendation Engines with Python. United Kingdom, Packt Publishing, 2018.

Recommended Readings:

1. Recommender Systems: Algorithms and Applications. United States, CRC Press, 2021.
2. Recommender Systems Handbook. Germany, Springer US, 2015.

Lab resources & reading material:

1. https://researchoutput.csu.edu.au/ws/portalfiles/portal/56945122/37059927_Published_article.pdf
2. https://core.ac.uk/download/pdf/227486364.pdf
3. https://dl.acm.org/doi/pdf/10.1145/2988450.2988454

DTA-EC507 | GRAPH THEORY

Teaching Scheme			
Lecture hours per week	**Theory hours per week**	**Practical hours per week**	**Credit**
3	**1**	**0**	**4**

Prerequisite:

Foundation of Data Analytics

Objectives:

1. To explain basic concepts in combinatorial graph theory.
2. To define how graphs, serve as models for many standard problems.

Learning Outcomes:

1. Solve problems using basic graph theory.
2. Identify induced subgraphs, matchings, and covers in graphs.
3. Determine whether graphs are Hamiltonian and/or Eulerian.
4. Solve problems involving vertex and edge connectivity, planarity and crossing numbers
5. Solve problems involving vertex and edge coloring.
6. Model real world problems using graph theory.

Syllabus:

Unit I: Introduction **12 hrs**

Applications of Graphs; Finite and Infinite Graphs; Incidence and Degree; Isolated and Pendant Vertex; Isomorphism; Sub Graph, Walks, Paths and Circuits; Connected and Disconnected Graphs; Components of A Graph; Euler Graphs; Hamiltonian Paths and Circuits; The Traveling Salesman Problem.

Unit II: Trees, Circuits and Cut-sets **12 hrs**

Properties of Trees; Pendant Vertices in A Tree; Center of a Tree; Rooted and Binary Trees; Spanning Tree, Spanning Trees in A Weighted Graph, Algorithm for Shortest Spanning Tree, Fundamental Circuits, Cut-sets and Cut Vertices; Fundamental Cut-sets, Connectivity and Separability.

Unit III: Planar Graphs **12 hrs**

Combinatorial Vs Geometric Graphs; Planar Graph; Kuratowski's Graphs; Detection of Planarity; Geometric Dual; Thickness and Crossings.

Unit IV: Matrix representation and coloring **12 hrs**

Path Matrix, Cut- Set Matrix, Circuit Matrix, Incidence Matrix, Adjacency Matrix and Their Properties. Chromatic Number, Chromatic Polynomial, Chromatic Partitioning, Matchings, Covering and Four-Color Problem.

Unit V: Directed Tree & Graphs **12 hrs**

Arborescence; Paired Comparison and Tournaments; Counting Labeled and Unlabeled Trees. Digraphs and Binary Relations; Directed Path and Connectedness; Adjacency Matrix of Digraph Shortest Path, Minimum Spanning Tree, Connectedness and Components, Fundamental Circuits, Cut-vertices and Separability, Isomorphism.

Required Readings:

1. N. Deo. Graph Theory with Applications to Engineering and Computer Science, PHI
2. Richard J. Trudeau, Introduction to Graph Theory, Dover Publications Inc

Recommended Readings:

1. An Introduction to Social Network Data Analytics-2011
2. Mathematical Foundations and Aspects of Discrete Mathematics, by Jean Gallier and Jocelyn Quaintance, Second Edition In Progress, 2020

DTA-EC508 | AMBIENT INTELLIGENCE

Teaching Scheme			
Lecture hours per week	Theory hours per week	Practical hours per week	Credit
3	1	0	4

Prerequisite:

Python Fundamentals, Basics of Computing

Objectives:

Connected devices and the Internet of Things (IoT) are becoming very popular these days, due to the advent of technologies like 5G and digitalization of various devices. These connected devices generate huge volumes of data every moment - how to deal with and manage these huge volumes of data? This course would provide an overview on how the data generated can be used for practical decision making purposes.

Learning Outcomes:

The students will get an overview of the following elements:

1. To introduce the concept of Internet of things and connected devices
2. To enable students to understand scope of Internet of things in Industry
3. To develop and apply Advance method for Implementation of Internet of Things
4. Get knowhow of the hardware and software for managing IoT data
5. Take simple actions based on the incoming data from the IoT devices.

Syllabus:

Unit-1: Introducing Internet of Things **(10 hrs)**

Introduction of IoT, Architecture of IoT, Introduction Industry 4.0, Need of IoT for Industry 4.0, Architecture of IoT, Block Diagrams of IoT System, Virtual Private server and IoT Cloud, Application Programming Interface (API)

Unit-2: Internet and Communication Protocols **(10 hrs)**

Introduction of Internet Protocol, Internet Layer: IP, Transport layer-TCP, UDP, Application Layer- HTTP, MQTT, FTP,CoAP, SPDY, IoT Protocols and types, Network: LORA, NRF, Xbee, IoT Gateway.

Unit-3: IOT Hardware and Software **(12 hrs)**

Hardware used in IoT (Arduino UNO, Nano, Nodecmu and Ethernet Shield), Software for IoT (Arduino IDE), Development of Things using Arduino Platform: Introduction of IoT Node with Sensor and Actuator, Interface sensors & devices, NodeMCU and ESP 32 wifi Microcontroller, IoT Platform and Application:

Customized IoT Platform using Virtual Private Server, Amazon Alexa, Google API, Blynk, Cayenne, Thingsboard, Thigspeak.

Unit-4: Applications of IOT **(10 hrs)**

Working with arduino language - structure, syntax, etc. , Applications and case studies of IoT, Working with Sensors, Ethernet boards, WiFi boards, Smart Devices and home automation, Advanced Use-case of IoT

Required Readings:

1. Samuel Greengard, The Internet of Things by Samuel Greengard -2015
2. Klaus Schwab, "The Fourth Industrial Revolution" by Klaus Schwab -2017
3. Getting started with Internet of Things by Cuno P Fister -2011

Recommended Readings:

1. Cuno Pfister, Getting Started with the Internet of Things: Connecting- 2011
2. Sensors and Microcontrollers to the Cloud (Make: Projects) 2018
3. Adrian McEwen, Designing the Internet of Things Kindle Edition- 2013
4. IoT and Smart Building Data – by Senseware. (Now Attune)
5. Sky Hook, Everything You Need to Know About LPWAN Location

DTA-EC509: Data Warehousing and Data Mining

Teaching Scheme			
Lecture hours per week	Theory hours per week	Practical hours per week	Credit
3	1	0	4

Pre-requisites:

NIL

Objectives:

Provide the student with an understanding of the concepts of data warehousing and data mining.

Learning Outcomes:

At the end of this course, the student should be able to:

1.Understand the functionality of the various data mining and data warehousing components.
2 Study the dimensional modelling technique for designing a data warehouse.
3. Explain the knowledge discovery process.
4. Describe the data mining tasks and study their well-known techniques.
5. Compare different approaches of data warehousing and data mining with various technologies.

Syllabus:

Unit - 1 **10hrs**

Introduction to Data Warehousing

Need of Data Warehousing, Evolution of Data Warehousing, Differences between Operational Database Systems and Data Warehouses, Data Warehousing concepts, Benefits of Data Warehousing, Data Ware House characteristics, Data Ware House Architecture and its components, Building a Data warehouse, Multidimensional Data Model, Data Cubes, Stars, Snow Flakes, Fact Constellations.

Unit - 2

Data Warehousing Tools and Technology **10hrs**

Tools and Technologies: Extraction, cleaning and Transformation tools, Data Warehouse DBMS, Data Warehouse Meta-Data, Rreporting and Query tools and Applications.

OLAP Operations: Drill-down and roll-up, slice-and-dice , pivot or rotation, OLAP models, overview of variations, the MOLAP model, the ROLAP model, the DOLAP model, ROLAP versus MOLAP,

Unit -3 **10hrs**

Introduction to Data Mining

What is Data Mining, Definition, KDD, Challenges, Data Mining Tasks, Data Pre-processing, Data Cleaning, Missing Data, Classification of Data Mining Systems, Association Rule Mining: - Efficient and

Scalable Frequent Item set Mining Methods, Mining Various Kinds of Association Rules, Association Mining to Correlation Analysis

UNIT 4 **12hrs**

Classification & Cluster Analysis: Classification by Decision Tree Introduction, Bayesian Classification, Rule Based Classification, Types of Data in Cluster Analysis, A Categorization of Major Clustering Methods, Partitioning Methods, Hierarchical methods, Density-Based Methods, Grid-Based Methods, Model-Based Clustering Methods, Outlier Analysis.

Required Readings:

1. A. Berson, S. J. Smith (2017), Data Warehousing, Data-Mining & OLAP, Tata McGraw Hill, Publications.
2. J. Han and M. Kamber (2012), Data Mining: Concepts and Technique, Third Edition, Elsevier.

Recommended Readings:

1. M. Humphries, M. W. Hawkins and M. C. Dy (1998), Data Warehousing: Architecture and Implementation, Pearson.
2. I. Singh (2014), Data Mining and Warehousing, Khanna Publishing House.
3. M. H. Dunham, S. Sridhar (2006), Data Mining: Introductory and Advanced Topics, Pearson Education.
4. A. K. Pujari (2016), Data Mining Techniques, Universities Press
5. **Pang-Ning Tan, Michael Steinbach** (2021), Introduction to Data Mining, 2nd edition, , Pearson Education.

DTA-EC510 | BFSI ANALYTICS

Teaching Scheme			
Lecture hours per week	**Theory hours per week**	**Practical hours per week**	**Credit**
3	**1**	**0**	**4**

Prerequisite:

Basics of Machine Learning

Objectives:

India is currently undergoing a fintech revolution and one of the key competitive tools in the hands of these fintech companies are Data Science and Artificial Intelligence. As aspects of the business in the fields of Banking, Finance or Insurance can use Data science for efficiency, cost saving, etc. This course would expose the students to some of the leading applications in the BFSI domain.

Learning Outcomes:

The students would be able to do the following:

1. Apply Data Analytics to the BFSI functions of the company
2. Understand possible areas for applying Data Analytics in BFSI

Syllabus:

Unit-I: The Changing BFSI Landscape **(10 hrs)**

Data Analytics help in BFSI Domain, Implementing Data Analytics at Enterprise Level from BFSI perspective, Infrastructure, Challenges, Business Impact

Unit-II: Data Warehouse or Data Lakes for BFSI domain **(10 hrs)**

Ingesting Data from various sources, Merging data across sources, Master Data Management , Deduplication of records, Analysing BFSI Data, Exploratory Data Analysis, Insights from data

Unit-III: BFSI with Machine Learning perspective **(10 hrs)**

Loan Default application using Machine Learning, Credit Risk Assessment using Regression methods, Insurance claims, detecting fraud using Predictive Analytics, Time Series data - Stock Market Prediction

Unit-IV: Working with Speech, Image and Business Data **(12 hrs)**

Working with Natural Language Processing, Analysis of Customer Complaints, Automated responses to customer requests, Chatbots for helping customers, Customer Segmentation, For Marketing efforts, Speech and Image Recognition, Use-cases for business process automation, Handwriting recognition.

Recommended Readings:

1. Machine Learning in Finance, Springer Publishing, 2021, Mathew F. Dixon

2. The Enterprises Big Data Lake, Alex Gorelik, O'reilly, 2021 e

DTA-EC511 | OPERATIONS AND SUPPLY CHAIN ANALYTICS

Teaching Scheme			
Lecture hours per week	**Theory hours per week**	**Practical hours per week**	**Credit**
3	**1**	**0**	**4**

Prerequisite:

Statistics

Objectives:

In this introductory Operations and Supply Chain Logistics course, we cover the three major building blocks of logistics networks: transportation, warehousing, and inventory. After completing this course, you will be able to differentiate the advantages and disadvantages of different modes of transportation. You will understand what goes into designing and setting up a warehousing facility. Finally, you will be able to select the options that enable you to develop logistics networks, that minimize costs and deliver top customer service. This is an introductory course designed to provide you with a start on your learning journey in logistics. You do not need to have any background in logistics, but it would be beneficial if you had a basic understanding of business concepts. Join me and we will learn together about logistics!

Learning Outcomes:

After completion of the course students will be able to do the followings related to supply Chain and logistics:

1. Warehouse Management and Logistics Planning and operations
2. Supply Chain, Inventory, and Lean Six Sigma
3. Supply Chain and Six Sigma
4. Forecasting, Planning, and Demand Forecasting

Syllabus:

Unit-I Transportation, Warehouse Management and Logistics (9 hrs)

a. Readings: Objectives, Summary, and Further Information about Moving Freight over the Road, Further Information about Motor Carriers, Further Information about Flying Freight, Further Information about Air Carriers, Further Information on Containers on a Train, Further Information on Express Delivery, Express Delivery Firms Background, Comparison of Transportation Modes

b. Faculty should explain the concepts like the first mile, middle mile, and last mile. Concentrate more on the last mile. Technical advancements in last mile – routing tools, tools used by last mile transportation companies (may be casu studies can be used to explain the concepts)

c. Why do companies need warehouses? How should we design a warehouse?, How much inventory do we need?, When to order inventory?

d. Reading: Warehousing Best Practices

e. ABC analysis, based on cost, in-fill rates

f. Introduction to new age analytical and asset management software

g. Facilities: How many and where? Factors Influencing Logistics Networks, Striving for Logistics Customer Service

h. Reading: The Square Root Law

Unit-II Lean Operations and Lean Inventory (9 hrs)

a. What are Supply Chain Operations? The Goals of Operations: Speed, Flexibility, Quality, and Cost, Lean Operations, Theory of Constraints, Linear Programming

b. Readings: "Lean Thinking", "The Goal"

c. Lean Inventory, Screencast - Economic Order Quantity, Screencast - Safety Stock Calculation

d. Readings: Economic Order Quantity Terms and Notation, The EOQ Calculation, Safety Stock Terms and Calculation, The Safety Stock Calculation, Lean Inventory Quiz Instructions

Unit-III Six Sigma and Lean Six Sigma (8 hrs)

a. From "Quality is Free" to "Six Sigma", DMAIC and the Define Stage, Screencast of DMAIC and the Measure Stage, Screencast of DMAIC and the Analyze Stage, Screencast of DMAIC and the Improve Stage, Screencast of DMAIC and the Control Stage

b. Readings: Summary of the DMAIC Example, Some more Six Sigma Examples

c. Explaining students why to do these six sigmas from Data Analytics perspective

d. The Lean Six Sigma Toolkit

e. Readings: The Road to Lean Six Sigma, Selecting a Great Lean Six Sigma Project

Unit-IV Simple Forecasting Methods, Naive Forecast and Cumulative Mean (8 hrs)

a. What is Supply Chain Planning? So, you want to forecast demand?, The Naive Forecast, Naive Forecast Screencast, The Cumulative Mean, Cumulative Mean Screencast,

b. Readings: Naive Forecast Example, Naive Forecast Solution, Cumulative Mean Example, Cumulative Mean Solution, Naive Method and Cumulative Mean Quiz Instructions

c. Forecast Accuracy Measures, Forecast Accuracy Measures Screencast, Moving Average, Screencast on Moving Average

d. Readings: Notes on Forecast Accuracy Measures, Forecast Accuracy Solution, Forecast Accuracy Quiz Instructions, Moving Average Notes, About Rounding, Moving Average Solution, Moving Average Quiz Instructions

Unit-V Exponential Smoothing, Forecast Selection and Supply Chain Planning (8 hrs)

a. Exponential Smoothing, Exponential Smoothing Screencast, Selecting the Best Forecast
b. Readings: Exponential Smoothing Notes, Exponential Smoothing Solution, Exponential Smoothing Quiz Instructions, Optimizing Exponential Smoothing, Forecasting Best Practices
c. Supply, Manufacturing and Distribution Planning
d. Reading: Sales and Operations Planning
e. Full procure to pay cycle

Required Readings:

1. A Practical Application of Supply Chain Management Principles. by Thomas I Schoenfeldt, ASQ Quality Press, 2008

Recommended Readings:

1. Supply Chain Analytics, A Wiley Brand, by Mark Morley, 2017