Course Code	Course Name	Group	Teaching Scheme	Credits
			Lectures	
RJSPGDSAI103	**Database and Warehousing**	DSC	2	2

Course Outcome:

- To understand the principles of Data warehousing and Database
- To be familiar with the Data warehouse architecture and its Implementation.
- To know the Architecture of a Database system.
- To understand the various Data preprocessing Methods.
- To perform classification and prediction of data.

Learning Outcome:

1. Students will be able to describe architecture and methods for storage and provision of enterprise data.
2. Students will develop competency in query development and essential business intelligence reporting.
3. Students will demonstrate competency in data modeling, including dimensional modeling

UNIT	TOPICS
UNIT - I	**Database Concepts:** Why Databases?, Data versus Information, Introducing the Database, Why Database Design Is Important, Overview of MongoDB, Introduction of MongoDB, Nosql Database, Advantage over RDBMS, MongoDB Data Types, Install MongoDB, MongoDB Data Modeling, MongoDB Operators, Query & Projection Operator, MongoDB Update Operator, Aggregation Pipeline Stages , MongoDB limit(), MongoDB sort(), Query Modifiers
UNIT - II	**Database Commands:** Aggregation Commands, **Database:** Create Database, Drop Database Collection, Create Collection, Drop Collection **CRUD :** Documents, Inset Documents, Update Documents, Delete Documents, Query Documents, SQL to MongoDB Mapping , MongoDB text search, Partial Updates & Document Limits, Connectivity, Java MongoDB, PHP MongoDB, Python MongoDB
UNIT - III	**Data Warehousing and Business Analysis**: - Data Warehousing Components, building a Data warehouse, Data Warehouse Architecture, DBMS Schemas for Decision Support, Data Extraction, Cleanup, and Transformation Tools –Metadata – reporting, Query tools and Applications, Online Analytical Processing (OLAP), OLAP and Multidimensional Data Analysis.

	Data Mining: - Data Mining Functionalities, Data Pre-processing, Data Cleaning, Data Integration and Transformation, Data Reduction, Data Discretization and Concept Hierarchy Generation, Architecture of Typical Data Mining Systems, Classification Of Data Mining Systems. **Association Rule Mining:** - Efficient and Scalable Frequent Itemset Mining Methods, Mining Various Kinds of Association Rules, Association Mining to Correlation Analysis, Constraint-Based Association Mining.
UNIT - IV	**Classification and Prediction:** - Issues Regarding Classification and Prediction, Classification by Decision Tree Introduction, Bayesian Classification, Rule Based Classification, Classification by Back propagation, Support Vector Machines, Associative Classification, Lazy Learners, Other Classification Methods, Prediction, Accuracy and Error Measures, Evaluating the Accuracy of a Classifier or Predictor, Ensemble Methods, Model Section. **Cluster Analysis**: 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, Clustering High, Dimensional Data, Constraint-Based Cluster Analysis, Outlier Analysis.

Text Book

Jiawei Han, Micheline Kamber and Jian Pei"Data Mining Concepts and Techniques", Third Edition, Elsevier, 2011.

Reference Books

1. Alex Berson and Stephen J. Smith "Data Warehousing, Data Mining & OLAP", Tata McGraw – Hill Edition, Tenth Reprint 2007.
2. K.P. Soman, Shyam Diwakar and V. Ajay "Insight into Data mining Theory and Practice", Easter Economy Edition, Prentice Hall of India, 2006.
3. G. K. Gupta "Introduction to Data Mining with Case Studies", Easter Economy Edition, Prentice Hall of India, 2006.
4. Pang-Ning Tan, Michael Steinbach and Vipin Kumar "Introduction to Data Mining", Pearson Education, 2007.
5. Shannon Bradshaw, Eoin Brazil, Kristina Chodorow "MongoDB: The Definitive Guide, 3rd Edition", O'Reilly, 2019.
6. Kyle Banker Peter Bakkum Shaun Verch Douglas Garrett Tim Hawkins, "MongoDB in Action", MANNING, 2nd Edition.

Course Code	Course Name	Group	Teaching Scheme	Credits
			Lectures	
RJSPGDSAI103P	**Database and Warehousing**	DSC	4	2

Practical List:

1. Demonstrate MongoDB Basic Operations
2. Demonstrate MongoDB Aggregation Operations
3. Demonstrate MongoDB Sorting & limiting & skipping Operations
4. Demonstrate MongoDB Comparison Operators
5. Demonstrate MongoDB Logical Operators
6. Demonstrate MongoDB $abs, $floor, $ceil Operator
7. Demonstrate MongoDB $log, $mod, $divide, $multiply operator.
8. Demonstrate MongoDB $pow, $sqrt, $subtract
9. Demonstrate MongoDB $trunc, $round, $cmp operator
10. Demonstrate MongoDB $concat, $size, $rename operator
11. Implementation of Supervised Learning:
    A. Decision Tree
    B. Logistic
    C. KNN
    D. IDK
    E. SMO
    F. Naive Bayes
12. Implementation of UnSupervised Learning:
    A. Clustering
        a. EM
        b. Hierarchical
        c. Density-Based
        d. Simple K Means
        e. Association
    B. Apriori

Course Code	Course Name	Group	Teaching Scheme	Credits
			Lectures	
RJSPGDSAI203	**Time Series Analysis and Forecasting**	DSC	2	2

Course Outcome:

1. Forecast the trend pattern exhibited by the given data by using various methods
2. Run and interpret time series models and regression models for time series
3. Use the Box-Jenkins approach to model and forecast time series data empirically
4. Analyze and estimate the cyclic components using special processes

Learning Outcome:

1. Fit various growth curves, trend and to measure seasonal indices
2. Understand forecasting by different methods
3. Able to calculate variance of a random component

UNIT	TOPICS
UNIT - I	Introduction to Trend Introduction to times series data, application of time series from various fields, Components of a time series, Decomposition of time series. **Trend:** Estimation of trend by free hand curve method, method of semi averages, fitting a various mathematical curve and growth curves.
UNIT - II	Trend and Seasonal Component Method of moving averages, Detrending, Effect of elimination of trend on other components of the time series. **Seasonal Component:** Estimation of seasonal component by Method of simple averages, Ratio to Trend, Ratio to moving average and Link relatives
UNIT - III	**Forecasting Variate component method** - Stationary Time series: Weak stationary, autocorrelation function and correlogram of moving average Forecasting: Exponential smoothing methods, short term forecasting methods: Brown‟s discounted regression, Box-Jenkins Method.
UNIT - IV	**Cyclic Component Deseasonalization** - Cyclic Component: Harmonic Analysis. **Some Special Processes**: Moving-average (MA) process and Autoregressive (AR) process of orders one and two, Estimation of the parameters of AR (1) and AR (2) – YuleWalker equations.

References

1. Kendall, M. (1976) Time Series. 2nd Edition, Charles Griffin and Co Ltd., London and High Wycombe.

2. Chatfield C. (1980). The Analysis of Time Series –An Introduction, 6th Edition, Chapman & Hall.
3. Mukhopadhyay P. (2011). Applied Statistics, 2nd ed. Revised reprint, Books and Allied Shumway, R. H., and Stoffer, D. S. (2006). Time Series Analysis and Its Applications With R Examples, 2 ed. Springer, New York, NY
4. Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). Time Series Analysis: Forecasting and Control. Prentice - Hall, Inc., Upper Saddle River, NJ.
5. Yaffee, R. and McGee, M. (2000). Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS. Academic Press, Inc., San Diego, CA.

Course Code	Course Name	Group	Teaching Scheme	Credits
			Lectures	
RJSPGDSAI203P	**Time Series Analysis and Forecasting**	DSC	4	2

Practical List:

1. Decompose time series data to find trend, seasonality, cyclic and irregularity..
2. Data conversion of non-stationary to stationary.
3. Perform a duckey-fuller test to check stationarity of data.
4. Implementation of moving averages models.
5. Demonstration of autocorrelation functions and partial autocorrelation functions.
6. Implementation of Autoregressive models.
7. Implementation of ARIMA model.
8. Implementation of SARIMA model.
9. Time series forecasting using exponential smoothing.
10. Implementation of LSTM for time series forecasting.