

| GANPAT UNIVERSITY | | | | | | | | | |
|-------------------------------------|-------------|---------------------|-----------------|----|----------------------------|-------------------------------------|----|----------------------|-----------|
| FACULTY OF ENGINEERING & TECHNOLOGY | | | | | | | | | |
| Programme | | Diploma Engineering | | | | Branch | | Computer Engineering | |
| Semester | | V | | | | Version | | 1.0.0.0 | |
| Effective from Academic Year | | | 2020-21 | | | Effective for the batch Admitted in | | | JULY 2018 |
| Subject code | | 1CE2504 | Subject Name | | | DATA MINING AND WAREHOUSING | | | |
| Teaching scheme | | | | | Examination scheme (Marks) | | | | |
| (Per week) | Lecture(DT) | | Practical(Lab.) | | Total | | CE | SEE | Total |
| | L | TU | P | TW | | | | | |
| Credit | 3 | 0 | 2 | 0 | 5 | Theory | 40 | 60 | 100 |
| Hours | 3 | 0 | 4 | 0 | 7 | Practical | 60 | 40 | 100 |

| Pre-requisites: |
|---|
| The knowledge of programming languages like Python, Perl, C/C++, SQL, and Java and basics of database |

| Course Learning Outcomes: |
|--|
| The course content should be taught and implemented with an aim to develop different skills leading to the achievement of the following competencies and course learning outcomes: |
| T1. To understand the concept of Data Mining & its attributes |
| T2. To apply the concept of data mining components and techniques in designing data mining systems |
| T3. To Learn how to pre-process data before applying data mining techniques |
| T4. Able to Choose the data-mining task (classification, regression, clustering, Association Rules etc.) |
| T5. Install and Configure WEKA Tool |
| The practical should be carried out in such a manner that students are able to acquire different learning outcomes in cognitive, psychomotor and affective domain to demonstrate course learning outcomes. |

| Course Content | | | | |
|---|---|---|-------|-----|
| Name of UNIT | Unit Content | Unit Learning Outcomes | Marks | Hrs |
| UNIT – 1 Fundamentals of data mining | 1.1 Basics of Data Mining 1.2 History, strategies, techniques, applications 1.3 challenges of data mining, Future of data mining 1.4 Types of Data 1.4.1 Database Data 1.4.2 Data Warehouses 1.4.3 Transactional Data | 1a. Describe the concept of Data Mining 1b. Describe types of Data | | 7 |

| | | | | |
|---|---|--|--|---|
| | 1.4.4 Other Kinds of Data | | | |
| UNIT – 2 Objects, Attributes & Statistical Description of Data | 2.1 Data Attribute 2.1.1 Nominal Attributes 2.1.2 Binary Attributes 2.1.3 Ordinal Attributes 2.1.4 Numeric Attributes 2.1.5 Discrete versus Continuous Attributes 2.2 Mean, Median, and Mode 2.3 Measuring the Dispersion of Data: Range, Quartiles, Variance, Standard Deviation, and Interquartile Range | 2a. Explain Mining techniques and Attribute Relation File Format (ARFF) 2b. Solve basic Statistical calculations on Data | | 8 |
| UNIT – 3 Data Preprocessing | 3.1 Preprocess the Data 3.2 Major Tasks in Data Preprocessing 3.3 Data Cleaning 3.3.1 Missing Values 3.3.2 Noisy Data 3.3.3 Data Cleaning as a Process 3.4 Data Integration | 3a. Describe the aspect of data preprocessing 3b. Explain the concept of Data Cleaning & Integration | | 9 |
| UNIT – 4 Classification | 4.1 Decision tree 4.2 Probability based solving 4.3 Concepts of Clustering | 4a. Explain decision Trees and clustering | | 9 |
| UNIT – 5 Data Warehouse | 5.1 Data Warehouse 5.2 Differences between Operational Database Systems and Data Warehouses | 5a. Apply the concept of Data Ware housing using WEKA solution | | 7 |
| UNIT – 6 Data Mining Tool: WEKA | 6.1 Basic of WEKA 6.2 Installing WEKA 6.3 WEKA data file format 6.4 Data visualization in WEKA 6.5 Data filtering 6.6 Using the concepts of data mining with WEKA | 6a. Install and Configure WEKA Tool | | 5 |

| List of Practical | | |
|-------------------|------|---|
| No. | Unit | Name of Practical |
| 1 | 1 | Perform practical about basic python programming. |
| 2 | 1 | Study and perform basic libraries like NumPy, Pandas, Matplotlib in python. |
| 3 | 2 | Demonstrate the use of ARFF files taking input and display the output of the files. |

| | | |
|----|---|---|
| 4 | 2 | Create your own excel file. Convert the excel file to .csv format and prepare it as ARFF files. |
| 5 | 3 | Perform Preprocessing, Classification techniques on any dataset. |
| 7 | 4 | Perform Clustering technique on dataset. |
| 8 | 4 | Classify the dataset using decision tree. |
| 9 | 5 | Perform Association technique on any dataset. |
| 10 | 6 | Install and Configure WEKA Tool |
| 11 | 6 | Demonstration of Weka Explorer, Mining techniques and Attribute Relation File Format (ARFF). |

| List of Instruments / Equipment / Trainer Board | |
|---|--|
| 1 | Latest computers in sufficient numbers |

| List of Text Books | | | |
|--------------------|-------------------------------------|---------------------------------|---------------------------|
| No | Title of Text Books | Authors | Publication |
| 1 | Data Mining Concepts and Techniques | Jiawei Han and Micheline Kamber | Kaufmann Publishers, 2011 |
| 2 | Data Mining Techniques | Arun K Pujari | Orient Longman Publishers |

| List of Reference Books | | | |
|-------------------------|------------------------------------|---|------------------|
| No | Title of Reference Books | Authors | Publication |
| 1 | Fundamentals of Data Warehouses | M.Jarke, M Lenzerni | |
| 2 | Principles of Data Mining | David Hand, Heikki Mannila, Padhraic Smyth, | PHI |
| 3 | Data Mining:Methods and Techniques | A B M Shawkat Ali, Saleh A, Wasimi | CENGAGE Learning |

| Link of Learning Web Resource | |
|-------------------------------|---|
| 1 | http://www.tutorialspoint.com/data_mining/ |
| 2 | https://www.guru99.com/data-mining-tutorial.html |
| 3 | https://www.javatpoint.com/data-mining |