GOVERNMENT POLYTECHNIC, NAGPUR.

(An Autonomous Institute of Govt. of Maharashtra)

COURSE CURRICULUM

PROGRAMME	:	**DIPLOMA IN INFORMATION TECHNOLOGY**
LEVEL NAME	:	**ELECTIVE COURSES**
COURSE CODE	:	**IT504E$^\$$**
COURSE TITLE	:	**DATA MINING AND DATA WAREHOUSING**
PREREQUISITE	:	**NIL**
TEACHING SCHEME	:	**TH: 03; TU: 00; PR: 02 (CLOCK HRs.)**
TOTAL CREDITS	:	**04 (1 TH/TU CREDIT = 1 CLOCK HR., 1 PR CREDIT = 2 CLOCK HR.)**
TH. TEE	:	**03 HRs**
PR. TEE	:	**02 HRs (External)**
PT.	:	**01 HR**

❖ **RATIONALE:**

This course introduces advanced aspects of data warehousing and data mining, encompassing the principles, research results and commercial application of the current technologies.

❖ **COURSE OUTCOMES:**

After completing this course students will be able to–

1. Identify the concepts of data warehousing and data mining.
2. Identify difference between DBMS and Data warehouse.
3. Appreciate the issues underlying database implementation.
4. Perform various operations using Data warehousing and data mining.
5. Perform query facilities to formulate queries and manipulate the database e.g. Structured Query Language (SQL).
6. Create Data Warehouse

❖ **COURSE DETAILS:**

A. THEORY :

Units	Specific Learning Outcomes (Cognitive Domain)	Topics and subtopics	Hrs.
1.Introduction to data warehouse and Data mining	1. Define the term Data warehouse and Data mining 2. List the functionalities Data warehouse 3. Describe the stages of KDD 4. Describe Issues in Data Warehouse and Data Mining 5. List Application of Data Warehouse and Data Mining.	1.1 Concepts of Data Warehouse and Data Mining 1.2 Functionalities 1.3 stages of Knowledge discovery in database(KDD) 1.4 Setting up a KDD environment 1.5 Issues in Data Warehouse and Data Mining, 1.6 Application of Data Warehouse and Data Mining	06
2.Datawarehouse and OLAP Technology	1. Differentiate between DBMS vs. Data Warehouse. 2. Design Multidimensional data model 3. Describe/Define the term related to Data, metadata 4. Define Data cubes 5. Design Data Warehouse Architecture 6. Describe Distributed and Virtual Data Warehouse 7. Define the term: OLTP, OLAP, 8. List types of OLAP	2.1 DBMS vs. Data Warehouse 2.2 Data marts 2.3 Metadata 2.4 Multidimensional data model, Data Cubes 2.5 Schemas for Multidimensional Database: Stars, Snowflakes and Fact Constellations 2.6 Data Warehouse Architecture 2.7 Distributed and Virtual Data Warehouse 2.8 Data Warehouse Manager, OLTP (online transaction processing), OLAP (online analytical processing) types of OLAP, Servers.	14
3.Datamining	1. Define the term Data mining 2. Distinguish between KDD versus Data mining 3. Describe the Data mining techniques, tools. 4. List application of Data mining. 5. Analyse the pattern presentation & visualization specification 6. Describe	3.1 Data Mining definition and Task 3.2 KDD versus Data Mining 3.3 Data Mining techniques, tools and application 3.4 Data mining query languages 3.5 Data specification, specifying knowledge, hierarchy specification 3.6 Pattern presentation & visualization specification 3.7 Data mining languages and	8

	standardization of data mining	standardization of data mining.	
4.Mining Association Rules in Large Databases	1. Define/Describe Association Rule Mining 2. Describe why Association Mining is necessary 3. Describe the characteristics of Pros and Cons of Association Rules 4. Write Apriori Algorithm	4.1 Association Rule Mining 4.2 Need of Association Mining 4.3 Pros and Cons of Association Rules 4.4 Apriori Algorithm	4
5.Classification and Prediction	1. Describe the Issues regarding classification and Prediction 2. Classify decision Tree Induction 3. Define Regression 4. State the types of Regression 5. Illustrate Issues regarding Classification & Prediction	5.1 Issues regarding 5.2 Classification & Prediction 5.3 Classification by Decision Tree Induction 5.4 Introduction to Regression 5.5 types of Regression	6
6.Clustering And Applications And Trends In Data Mining	1. Describe the term related to Cluster analysis 2. Select appropriate methods for categorization of cluster. 3. Illustrate k-mean method 4. Differentiate between different methods of clustering. 5. List application of Data mining.	6.1 Cluster Analysis 6.2 Types of Data 6.3 Categorization of Major Clustering Methods 6.4 K-means, Partitioning Methods 6.5 Hierarchical Methods 6.6 Density-Based Methods ,Grid Based Methods, Model-Based Clustering Methods 6.7 Data Mining Applications.	10
		Total Hrs.	**48**

B. LIST OF PRACTICALS/LABORATORY EXPERIENCES/ASSIGNMENTS:

Practi cals	Specific Learning Outcomes (Psychomotor Domain)	Units	Hrs.
1.	Identify the fundamental concepts of data warehouse and Data mining	Introduction to data warehouse and Data mining	2
2.	Create a simple data warehouse		2
3.	Perform OLAP operations such as Roll Up, Drill Down, Slice, Dice through SQL- Server	Data warehouse and OLAP Technology Data mining	2
4.	Perform preprocessing on dataset Weather. ARFF (Specify the name of the dataset chosen by each individual, instead of Weather) includes creating an ARFF file and reading it into WEKA using the WEKA Explorer.	Classificatio n and Prediction	2
5.	Implement Data Cleansing applying uppercase on first name and last name in C++.		4
6.	Perform Preprocessing, Classification and Visualization techniques on Agriculture dataset.	Clustering And Applications And Trends In Data Mining	4
7.	Perform Association rule based on (Apriori algorithm) or Clustering algorithm (Kmeans)		4
8.	Perform Clustering technique on Customer dataset		4
9.	Perform Association technique on Agriculture dataset.		4
		Skill Assessment	4
		Total Hrs	**32**

❖ **SPECIFICATION TABLE FOR THEORY PAPER:**

Unit No.	Units	Levels from Cognition Process Dimension			Total Marks
		R	**U**	**A**	
01	Introduction to data warehouse and Data mining	04(04)	04(04)	06(00)	**14(08)**
02	Data warehouse and OLAP Technology	04(00)	04(08)	06(00)	**14(08)**
03	Data mining	00(02)	04(04)	06(00)	**10(06)**
04	Mining Association Rules in Large Databases	00(02)	06(04)	06(00)	**12(06)**
05	Classification and Prediction	04(00)	04(00)	00(06)	**08(06)**
06	Clustering And Applications And Trends In Data Mining	02(00)	10(00)	00(06)	**12(06)**
	Total	**14(08)**	**32(20)**	**24 (12)**	**70 (40)**

R – Remember   U – Understand   A – Analyze / Apply

❖ **QUESTION PAPER PROFILE FOR THEORY PAPER**

Q. No	Bit 1			Bit 2			Bit 3			Bit 4			Bit 5			Bit 6			option
	T	L	M	T	L	M	T	L	M	T	L	M	T	L	M	T	L	M	
01	1	R	2	1	R	2	4	U	2	6	U	2	6	R	2	3	R	2	**5/7**
	4	R	2																
02	2	R	4	1	U	4	2	U	4	2	U	4	2	U	4				**3/5**
03	3	U	4	4	U	4	5	R	4	1	R	4	1	U	4				**3/5**
04	5	U	4	6	U	4	6	U	4	3	U	4	4	U	4				**3/5**
05	1	A	6	2	A	6	5	A	6										**2/3**
06	3	A	6	4	A	6	6	A	6										**2/3**

T= Unit/Topic Number   L= Level of Question   M= Marks

R-Remember   U-Understand   A-Analyze/ Apply

❖ **ASSESSMENT AND EVALUATION SCHEME:**

		What	To Whom	Frequency	Max Marks	Min Marks	Evidence Collected	Course Outcomes
Direct Assessment Theory	**CA (Continuous Assessment)**	Progressive Test (PT)	Students	Two PT (average of two tests will be computed)	20	--	Test Answer Sheets	1, 2, 3
		Assignments		Continuous	10	--	Assignment Book / Sheet	1, 2, 3
	TEE (Term End Examination)	End Exam	Students	End Of the Course	**70**	**28**	Theory Answer Sheets	1, 2, 3
				Total	**100**	**40**		
Direct Assessment Practical	**CA (Continuous Assessment)**	Skill Assessment	Students	Continuous	20	--	Rubrics & Assessment Sheets	4,5,6
		Journal Writing		Continuous	05	--	Journal	4,5,6
				TOTAL	**25**	**10**		
	TEE (Term End Examination)	End Exam	Students	End Of the Course	**50**	**20**	Rubrics & Practical Answer Sheets	4,5,6
Indirect Assessment		Student Feedback on course	Students	After First Progressive Test	Student Feedback Form			1, 2, 3, 4,5,6
		End Of Course		End Of The Course	Questionnaires			

❖ **SCHEME OF PRACTICAL EVALUATION:**

S.N.	Description	Max. Marks
1	Drawing diagram, selection of equipment's., writing procedure etc.	10
2	Performance	20
3	Knowledge	10
4	Viva voice	10
	TOTAL	**50**

❖ **MAPPING COURSE OUTCOMES WITH PROGRAM OUTCOMES:**

Course Outcomes (COs)	Program Outcomes (POs)										PSOs	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1**	**2**
1	-	3	3	3	-	-	-	-	-	3	3	3
2	-	3	3	3	-	-	-	-	-	3	3	3
3	-	3	3	3	-	-	-	-	-	3	-	3
4	-	3	3	3	-	-	-	3	3	3	3	3
5	-	3	3	3	-	-	-	3	3	3	-	3
6	-	3	3	3	-	-	-	3	3	3	3	3

1: Slight (Low)   2: Moderate (Medium)   3: Substantial (High)

❖ **REFERENCE & TEXT BOOKS:**

S.N.	Title	Author, Publisher, Edition and Year Of publication	ISBN Number
1.	Data Warehousing	Reema Thareja, Oxford,2009	13: 9780198096221
2.	Data Warehousing Fundamentals	Paulraj Ponnian, John Willey, 2001.	13:978047046207-2
3	Data Mining Techniques	Arun K pujari, Universities Press, Second Edition, 2010.	13:978 81 7371 672 0
4.	Introduction to Data Mining	Ping-Ning Tan, Vipin kumar, Steinbach, Pearson, 2006	13: 9780321321367
5.	Introduction to Data Mining with Case Studies	G.K.Gupta, PHI learning Pvt. Ltd.,2014	13:9788120343269

❖ **E-REFERENCES:**

- https://docs.oracle.com/cd/B28359_01/datamine.111/b28129/process.htm, assessed on 01 August 2016
- https://www.youtube.com/watch?v=dGDuD10U4-M, assessed on 01 August 2016

❖ **LIST OF MAJOR EQUIPMENTS/INSTRUMENTS WITH SPECIFICATION**

1. Operating System(Windows 98/2000/XP/7/8/8.1/10)
2. CPU Intel core i3/i5/i7
3. Printers(INK JET/LASER)
4. Softwares:
   - i. IBM - Intelligent Miner and more
   - ii. SAS - Enterprise Miner
   - iii. Silicon Graphics - MineSet
   - iv. Oracle - Thinking Machines - Darwin
   - v. Angoss – knowledge SEEKER

❖ **LIST OF EXPERTS & TEACHERS WHO CONTRIBUTED FOR THIS CURRICULUM:**

S.N.	Name	Designation	Institute / Industry
1.	Dr.A.R.Mahajan	Head of Information Technology	Government Polytechnic, Nagpur
2.	Mr.S.P.Lambade	Head of Computer Engg.	Government Polytechnic, Nagpur
3.	Ms.I.G.Lokhande	Lecturer in Information Technology	Government Polytechnic, Nagpur.
4.	Mr.R.L.Meshram	Lecturer in Information Technology	Government Polytechnic, Nagpur
5.	Prof. Manoj Jethawa	HOD, Computer Science	ShriDattaMeghe Polytechnic, Nagpur
6	Prof. N. V. Choudhari	Asst. Professor(CSE),	DBACOE, Wanadongari, Nagpur
7	Mr. Atul Upadhyay	CEO	Vista Computers, Ramnagar, Nagpur

_____

(Member Secretary PBOS)

_____

(Chairman PBOS)