

SUBJECT CODE : RIT-E22

DECODE[®]

AKTU - 2016 Syllabus
Semester-VI [CSE/CS&IT/IT]
Elective - II
Choice Based Credit System

DATA WAREHOUSING AND DATA MINING

Mrs. Sunanda Phadtare - Bhosale

M.Sc. (Elect.) SET, MCM, ADCSSAA
(HOD, Computer Science Department)
SNBP College, Pune

FEATURES

- ☒ Covers Entire Syllabus
- ☒ Question Answer Format
- ☒ Exact Answers and Solutions

SOLVED UNIVERSITY QUESTION PAPERS

- | | | |
|-------------|-------------|-------------|
| • 2010 - 11 | • 2011 - 12 | • 2012 - 13 |
| • 2013 - 14 | • 2014 - 15 | • 2015 - 16 |
| • 2016 - 17 | • 2017 - 18 | |

CONTENTS

Chapter - 1 Data Warehousing	(1 - 1) to (1 - 24)
Chapter - 2 Data Warehouse Process & Technology	(2 - 1) to (2 - 8)
Chapter - 3 Data Mining	(3 - 1) to (3 - 30)
Chapter - 4 Classification	(4 - 1) to (4 - 59)
Chapter - 5 Data Visualization & Overall Perspective	(5 - 1) to (5 - 24)
Solved Question Papers of AKTU	(S - 1) to (S - 30)
2010-11 to 2017-18	

© Copyright with Technical Publications

All publishing rights (printed and ebook version) reserved with Technical Publications. No part of this book should be reproduced in any form, Electronic, Mechanical, Photocopy or any information storage and retrieval system without prior permission in writing, from Technical Publications, Pune.

Published by :



Amit Residency, Office No.1, 412, Shaniwar Peth, Pune - 411030, M.S. INDIA
Ph. : +91-020-24495496/97, Telefax : +91-020-24495497
Email : sales@technicalpubilcations.org Website : www.technicalpubilcations.org

Printer :

Yogiraj Printers & Binders, Works : Sr.No. 10\1A, Ghule Industrial Estate, Nanded Village Road,
Tal-Haveli, Dist-Pune - 411041.

First Edition : 2019 (As per CBCS Scheme)

Price : ₹ 85/-

ISBN 978-93-332-1765-1



9 789333 217651

AKTU 16

IMPORTANT NOTE FROM PUBLISHER

Dear Students,

It gives us immense pleasure to present you this Decode guide prepared according to the latest syllabus & question paper pattern of AKTU.

We know that in semester pattern students don't get time to study all subjects thoroughly from text books which are very big in size.

Decode contains whole subject in question-answer format.

It contains model answers & solutions to all the questions & numerical problems asked in last 8 years university question papers. There are also additional questions & numerical problems with answers so as to cover every concept of the subject.

This decode guide contains complete syllabus of the subject in simple & lucid language. Answers, solutions have been illustrated using neat, correct & easily reproducible diagrams.

Solved University Question Papers are given as per new paper pattern of AKTU.

All the latest syllabus papers are solved and included at the end of the book.

This systematic approach will help you to study & revise the subject completely from the examination point of view.

We assure that this decode guide will help you to secure a high percentage of marks in your exam.

Best of Luck

SYLLABUS

DATA WAREHOUSING & DATA MINING [RIT-E22]

UNIT - I

Data Warehousing : Overview, Definition, Data Warehousing Components, Building a Data Warehouse, Warehouse Database, Mapping the Data Warehouse to a Multiprocessor Architecture, Difference between Database System and Data Warehouse, Multi Dimensional Data Model, Data Cubes, Stars, Snow Flakes, Fact Constellations, Concept (**Chapter - 1**)

UNIT - II

Data Warehouse Process and Technology : Warehousing Strategy, Warehouse /management and Support Processes, Warehouse Planning and Implementation, Hardware and Operating Systems for Data Warehousing, Client/Server Computing Model & Data Warehousing. Parallel Processors & Cluster Systems, Distributed DBMS implementations, Warehousing Software, Warehouse Schema Design, (**Chapter - 2**)

UNIT - III

Data Mining : Overview, Motivation, Definition & Functionalities, Data Processing, Form of Data Pre-processing, Data Cleaning: Missing Values, Noisy Data, (Binning, Clustering, Regression, Computer and Human inspection), Inconsistent Data, Data Integration and Transformation. Data Reduction:-Data Cube Aggregation, Dimensionality reduction, Data Compression, Numerosity Reduction, Discretization and Concept hierarchy generation, Decision Tree. (**Chapter - 3**)

UNIT - IV

Classification : Definition, Data Generalization, Analytical Characterization, Analysis of attribute relevance, Mining Class comparisons, Statistical measures in large Databases, Statistical-Based Algorithms, Distance-Based Algorithms, Decision Tree-Based Algorithms. Clustering : Introduction, Similarity and Distance Measures, Hierarchical and Partitional Algorithms. Hierarchical Clustering- CURE and Chameleon. Density Based Methods-DBSCAN, OPTICS. Grid Based Methods- STING, CLIQUE. Model Based Method –Statistical Approach, Association rules: Introduction, Large Item sets, Basic Algorithms, Parallel and Distributed Algorithms, Neural Network approach. (**Chapter - 4**)

UNIT - V

Data Visualization and Overall Perspective : Aggregation, Historical information, Query Facility, OLAP function and Tools. OLAP Servers, ROLAP, MOLAP, HOLAP, Data Mining interface, Security, Backup and Recovery, Tuning Data Warehouse, Testing Data Warehouse. Warehousing applications and Recent Trends : Types of Warehousing Applications, Web Mining, Spatial Mining and Temporal Mining (**Chapter - 5**)