3.3.3 Total number of books and chapters in edited volumes/books published and papers in national/ international conference proceedings year wise during last five years

Sr. No.	File Description	Page No.
1	Books by Authors	2

3.3.3 Total number of books and chapters in edited volumes/books published and papers in national/international conference proceedings year wise during last five years

Sr. No	Activity Name
1	Books by Author – Dr G.T. Thampi
2	Books by Author – Dr Arti Deshpande
3	Book by Author – Dr Arun Kulkarni
4	Book by Author - Dr Bhushan Jadhav
5	Book by Author- Shilpa Ingoley

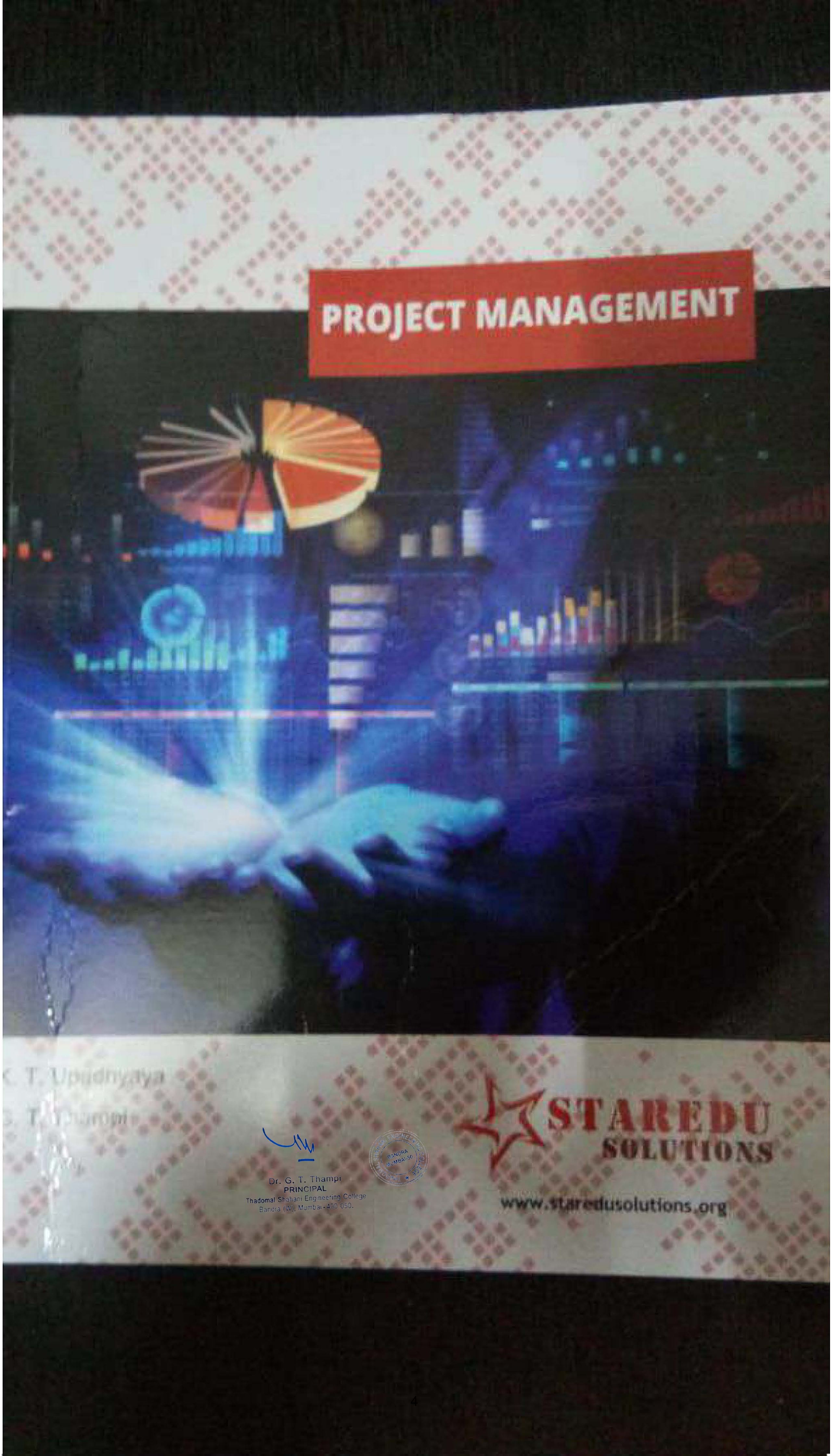
Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

Books by Author – Dr G.T. Thampi

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE ★ BANDRA MUMBAI-50.



PROJECT MANAGEMENT

K. T. Upadhyaya

G. T. Thampi

STAREDU SOLUTIONS

www.staredusolutions.org

Limits of Liability:

While the publisher and the author have used their best efforts in preparing this book, StarEdu Solutions and the author make no representation warranties with respect to the accuracy or completeness of the contents of this book, and specifically disclaim any implied warranties of merchantability or fitness for any particular purpose. There are no warranties which extend beyond the descriptions contained in this paragraph. No warranty may be created or extended by sales representatives or written sales materials.

Disclaimer:

The contents of this book have been checked for accuracy. Since deviations cannot be precluded entirely, StarEdu Solutions or its author cannot guarantee full agreement. As the book is intended for educational purpose, StarEdu Solutions or its author shall not be responsible for any errors, omissions or damages arising out of the use of the information contained in the book. This publication is designed to provide accurate and authoritative information with regard to the subject matter covered. Due credit for all contributions, images and reference are given in the courseware. In case, we missed on anything, we request you to bring it to our knowledge and we will address the same in next edition of the book. We believe in protecting the intellectual property. If you are copyright owner and if you believe that content constitutes an infringement of copyright, please bring to our knowledge and in case of justified issue, we will address the same in next edition of the book.

This book may contain reference to other resources related to topics. These references are for training purpose only. However, StarEdu Solutions do not recommend or promoting such references, solutions, software and products mentioned in reference or training content.

Images used in the book are for education purpose only and we are not promoting or advertising an product.

All Readers are requested to report errors, inaccuracies and omissions to StarEdu Solutions at **support@staredusolutions.org**, so that same can be addressed in the next edition of the the book.

ISBN: 978-93-86765-71-0

Edition: 2020

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50

Preface

Project Management is a significant concept in the modern times. The business organisations have been established to earn profits by creating deliverables and selling them in the market. These deliverables are developed through projects. Often these projects are carried out without any management and supervision. This leads to increased costs, delayed projects and unachieved aims. Hence, the Project Management is the need of the hour.

The project completion is achieved through the coordinated efforts of the members of project teams led by the project manager. A project leader is responsible for managing all the stages of the project and garnering support from the sponsors and the management, so as to ensure that the goals of the project are achieved. Therefore, the project managers should be familiar with all the project management techniques to control the execution of the projects and deliver the intended results on time. These techniques should be revisited, reorganised and reframed, depending on the ever-changing business scenarios. Keeping this view in mind, the Project Management book has been composed, which serves as a powerful tool for enabling the students to get acquainted with the innovative concepts of project management, its techniques and its applications.

The **Project Management book** is designed to enlighten the students about the importance of project management in business organisations. The book covers all the major aspects of project management, including project life cycle, initiation of projects, project planning and scheduling, project risk management, project execution and monitoring, and project termination. After studying this book, the students will be able to comprehend and apply the various concepts of project management in real-world business scenarios, so as to gain hands-on experience.

The **Project Management book** is developed with an updated content about the different aspects of the project management. It is written in a simple, lucid and easy-to-understand language. In this book, various real-life examples have been provided in all the chapters to enable the students to gain insights into the application of project management concepts. The examples will also guide them on how to deal with the similar real-life situations and take the necessary measures.

The content of the **Project Management book** has been developed after a wide-scale research done on the subject. The book has been designed in such a way so as to enhance the understanding of the students regarding the complex issues, often faced while handling project management related concepts. The book will enable the students to gain a clear understanding and confidence of using project management techniques and lead the multitude of projects in their real lives.

About the Authors

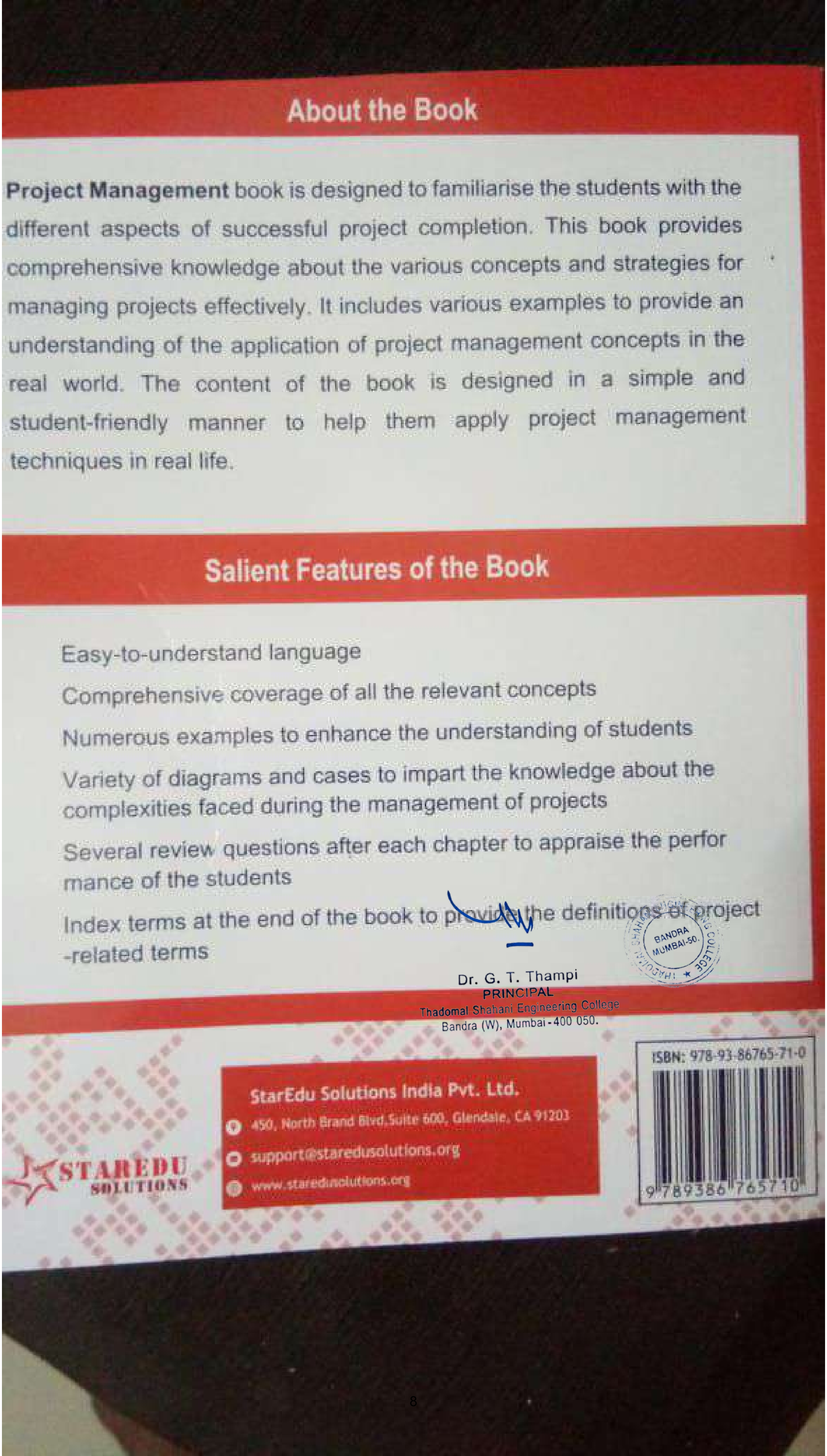
Dr. K. T. Upadhyaya has over 17 years experience in executing large scale projects initially in Cement, Petrol chemical, Food, and Dairy industry and currently in Software development, Data Analytics and Automation projects. Dr. K. T. Upadhyaya is engaged in training in Project Management, Risk Management, Data warehousing for past 16 years. He has trained over 10,000 participants from companies in different domains and students from top B-schools in India. He also trains on Data Analytics, Supply Chain Analytics, Business Process Management, Process Automation and AI.

Dr. K. T. Upadhyaya has keen interest in relating Indian Mythology, particularly Mahabharata, to management lessons. He is an avid reader of books relating to topics ranging from Data to Fiction and everything in between. Sachin Dev Burman and Sachin Tendulkar are his two favourites in their respective domains. Dr. K. T. Upadhyaya holds a Mechanical Engineering Degree from Sardar Patel College of Engineering, Mumbai and Masters and Ph. D. from BITS Pilani. He is a Certified Project Management Professional from PMI(R), Pennsylvania, USA.

Dr. G. T. Thampi is currently The Principal at Thadomal Shahani Engineering College, Bandra, Mumbai. He hold a Ph.D Degree in Technology and has more than 33 years of experience in renowned college. Dr. G.T. Thampi has been a part of some interesting researches and holds interests in Business Process and Re-Engineering in realm of Engineering Education. Apart from his own researches he has been a guide for multiple researches done in technology front. 17 research scholars awarded Ph.D. under his guidance. Dr. Thampi is also a co-author of more than 80 research publications and books

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE ★
BANDRA
MUMBAI-50.



About the Book

Project Management book is designed to familiarise the students with the different aspects of successful project completion. This book provides comprehensive knowledge about the various concepts and strategies for managing projects effectively. It includes various examples to provide an understanding of the application of project management concepts in the real world. The content of the book is designed in a simple and student-friendly manner to help them apply project management techniques in real life.

Salient Features of the Book

- Easy-to-understand language
- Comprehensive coverage of all the relevant concepts
- Numerous examples to enhance the understanding of students
- Variety of diagrams and cases to impart the knowledge about the complexities faced during the management of projects
- Several review questions after each chapter to appraise the perfor mance of the students
- Index terms at the end of the book to provide the definitions of project -related terms

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

StarEdu Solutions India Pvt. Ltd.
450, North Brand Blvd,Suite 600, Glendale, CA 91203
support@staredusolutions.org
www.staredusolutions.org

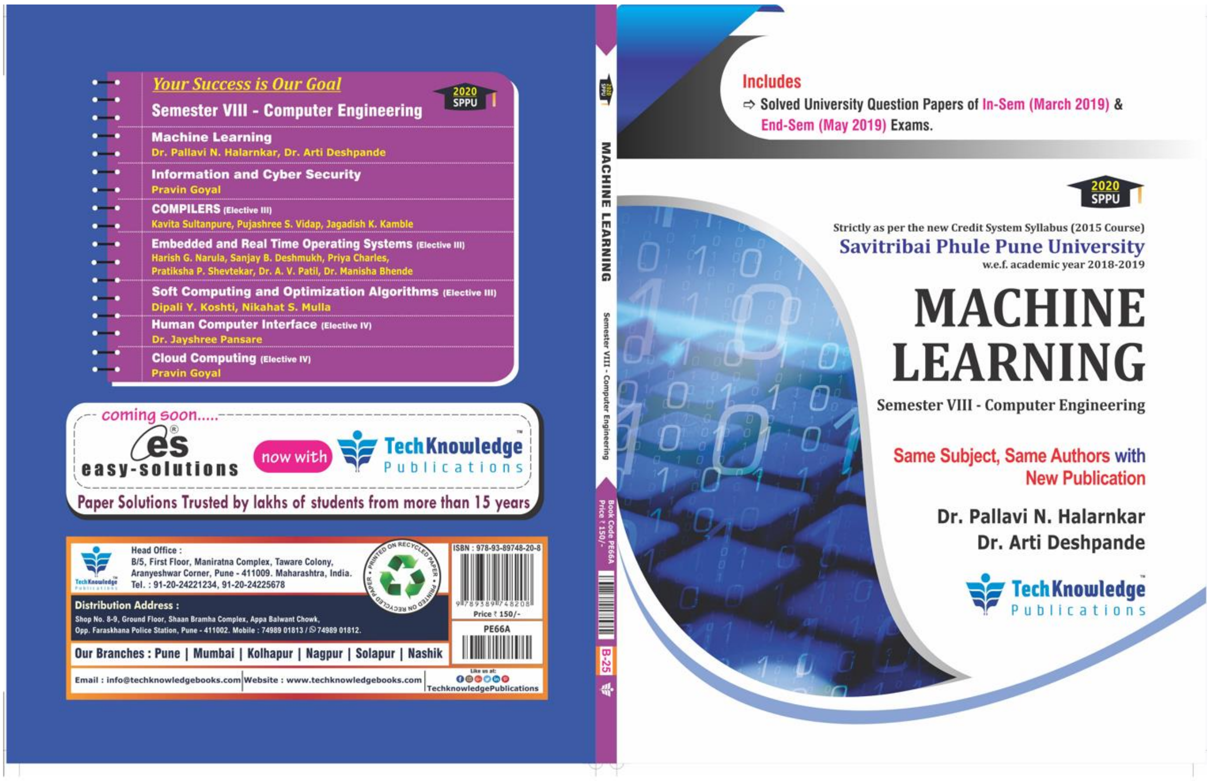
STAREDU SOLUTIONS

ISBN: 978-93-86765-71-0
9 789386 765710

Books by Author – Dr Arti Deshpande

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE ★
BANDRA
MUMBAI-50.



Your Success is Our Goal

2020 SPPU

Semester VIII - Computer Engineering

Machine Learning
Dr. Pallavi N. Halarnkar, Dr. Arti Deshpande

Information and Cyber Security
Pravin Goyal

COMPILERS (Elective III)
Kavita Sultanpure, Pujashree S. Vidap, Jagadish K. Kamble

Embedded and Real Time Operating Systems (Elective III)
Harish G. Narula, Sanjay B. Deshmukh, Priya Charles,
Pratiksha P. Shevtekar, Dr. A. V. Patil, Dr. Manisha Bhende

Soft Computing and Optimization Algorithms (Elective III)
Dipali Y. Koshti, Nikahat S. Mulla

Human Computer Interface (Elective IV)
Dr. Jayshree Pansare

Cloud Computing (Elective IV)
Pravin Goyal

TechKnowledge™ Publications
Head Office :
B/5, First Floor, Maniratna Complex, Taware Colony,
Aranyeshwar Corner, Pune - 411009. Maharashtra, India.
Tel. : 91-20-24221234, 91-20-24225678

PRINTED ON RECYCLED PAPER • PRINTED ON RECYCLED PAPER •

ISBN : 978-93-89748-20-8
9789389748208
Price ₹ 150/-
PE66A

Distribution Address :
Shop No. 8-9, Ground Floor, Shaan Bramha Complex, Appa Balwant Chowk,
Opp. Faraskhana Police Station, Pune - 411002. Mobile : 74989 01813 / ✆ 74989 01812.

Our Branches : Pune | Mumbai | Kolhapur | Nagpur | Solapur | Nashik

Email : info@techknowledgebooks.com | Website : www.techknowledgebooks.com | Like us at: TechknowledgePublications

2020 SPPU
MACHINE LEARNING
Semester VIII - Computer Engineering
Book Code PE66A
Price ₹ 150/-
B-25

Includes
⇨ Solved University Question Papers of In-Sem (March 2019) & End-Sem (May 2019) Exams.

2020 SPPU

Strictly as per the new Credit System Syllabus (2015 Course)
Savitribai Phule Pune University
w.e.f. academic year 2018-2019

MACHINE LEARNING

Semester VIII - Computer Engineering

Same Subject, Same Authors with New Publication

Dr. Pallavi N. Halarnkar
Dr. Arti Deshpande

TechKnowledge™ Publications

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE ★ BANDRA MUMBAI-50.

Machine Learning

(Code : 410250)

Semester VIII - Computer Engineering
(Savitribai Phule Pune University)

Strictly as per the New Credit System Syllabus (2015 Course)
Savitribai Phule Pune University w.e.f. Academic Year 2018-2019

Dr. Pallavi N. Halarnkar

Department of Computer Engineering
Thadomal Shahani Engineering College, Mumbai.
Maharashtra, India.

Dr. Arti Deshpande

Department of Computer Engineering
Thadomal Shahani Engineering College,
Mumbai.
Maharashtra, India.

TechKnowledge™
Publications

PE66A Price ₹ 150/-

(Book Code : PE66A)

Machine Learning

Dr. Pallavi N. Halarnkar, Dr. Arti Deshpande

(Semester VIII - Computer Engineering) (Savitribai Phule Pune University)

First Edition : February 2019

Second Revised Edition : January 2020 (TechKnowledge Publications)

This edition is for sale in India, Bangladesh, Bhutan, Maldives, Nepal, Pakistan, Sri Lanka and designated countries in South-East Asia. Sale and purchase of this book outside of these countries is unauthorized by the publisher.

Printed at : 37/2, Ashtavinayak Industrial Estate,
Near Pari Company,
Narhe, Pune, Maharashtra State, India.
Pune – 411041

ISBN : 978-93-89748-20-8

Published by

TechKnowledge Publications

Head Office : B/5, First floor, Maniratna Complex,
Taware Colony, Aranyeshwar Corner,
Pune - 411 009. Maharashtra State, India
Ph : 91-20-24221234, 91-20-24225678.
Email : info@techknowledgebooks.com,
Website : www.techknowledgebooks.com

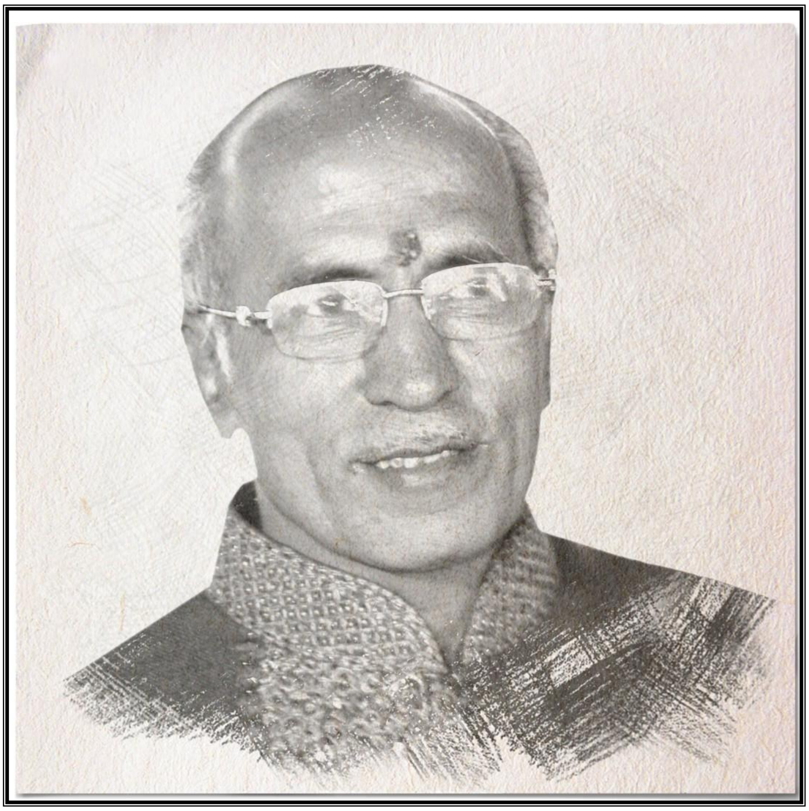
[410250] (FID : PE66) (Book Code : PE66A)

We dedicate this Publication soulfully and wholeheartedly,

in loving memory of our beloved founder director,

Late Shri. Pradeepji Lalchandji Lunawat,

who will always be an inspiration, a positive force and strong support behind us.



"My work is my prayer to God"

- Lt. Shri. Pradeepji L. Lunawat

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

Soulful Tribute and Gratitude for all Your
Sacrifices, Hardwork and 40 years of Strong Vision…

Preface

Dear students,

We are extremely happy to present the book of **"Machine Learning"** for you. We have divided the subject into small chapters so that the topics can be arranged and understood properly. The topics within the chapters have been arranged in a proper sequence to ensure smooth flow of the subject.

We present this book in the loving memory of **Late Shri. Pradeepji Lunawat,** our source of inspiration and a strong foundation of **"TechKnowledge Publications"**. He will always be remembered in our heart and motivate us to achieve our milestone.

We are thankful to Mr. Shital Bhandari, Shri. Arunoday Kumar and Shri. Chandroday Kumar for the encouragement and support that they have extended. We are also thankful to the staff members of TechKnowledge Publications and others for their efforts to make this book as good as it is. We have jointly made every possible efforts to eliminate all the errors in this book. However if you find any, please let us know, because that will help us to improve further.

We are also thankful to my family members and friends for patience and encouragement.

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA MUMBAI

Authors

❑❑❑

(Book Code : PE66A)

Syllabus

Savitribai Phule Pune University
Fourth Year of Computer Engineering (2015 Course)

410250 : Machine Learning

Teaching Scheme : **TH : 03 Hours/Week**	**Credit** **03**	**Examination Scheme :** **In-Sem (Paper) : 30 Marks** **End-Sem (Paper) : 70 Marks**

Prerequisite Courses : 207003 - Engineering Mathematics III

Companion Course:410254 - Laboratory Practice III

Course Objectives

- To understand human learning aspects and relate it with machine learning concepts.
- To understand nature of the problem and apply machine learning algorithm.
- To find optimized solution for given problem.

Course Outcomes

On completion of the course, student will be able to -

- Distinguish different learning based applications.
- Apply different preprocessing methods to prepare training data set for machine learning.
- Design and implement supervised and unsupervised machine learning algorithm.
- Implement different learning models.
- Learn Meta classifiers and deep learning concepts.

Course Contents

Unit I : Introduction to Machine Learning **(08 Hours)**

Classic and adaptive machines, Machine learning matters, Beyond machine learning-deep learning and bio inspired adaptive systems, Machine learning and Big data.

Important Elements of Machine Learning - Data formats, Learnability, Statistical learning approaches, Elements of information theory. **(Refer chapter 1)**

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

Unit II : Feature Selection **(08 Hours)**

Scikit - learn Dataset, Creating training and test sets, managing categorical data, Managing missing features, Data scaling and normalization, Feature selection and Filtering, Principle Component Analysis(PCA) - non negative matrix factorization, Sparse PCA, Kernel PCA. Atom Extraction and Dictionary Learning. **(Refer chapter 2)**

(Book Code : PE66A)

Unit III : Regression **(08 Hours)**

Linear Regression : Linear models, A bi-dimensional example, Linear Regression and higher dimensionality, Ridge, Lasso and ElasticNet, Robust regression with random sample consensus, Polynomial regression, Isotonic regression.

Logistic Regression : Linear classification, Logistic regression, Implementation and Optimizations, Stochastic gradient descendent algorithms, Finding the optimal hyper-parameters through grid search, Classification metric, ROC Curve.

(Refer chapter 3)

Unit IV : Naïve Bayes and Support Vector Machine **(08 Hours)**

Bayes´ Theorom, Naïve Bayes´ Classifiers, Naïve Bayes in Scikit - learn- Bernoulli Naïve Bayes, Multinomial Naïve Bayes, and Gaussian Naïve Bayes.

Support Vector Machine(SVM) : Linear Support Vector Machines, Scikit- learn implementation-Linear Classification, Kernel based classification, Non-linear Examples. Controlled Support Vector Machines, Support Vector Regression. **(Refer chapter 4)**

Unit V : Decision Trees and Ensemble Learning **(08 Hours)**

Decision Trees : Impurity measures, Feature Importance. Decision Tree Classification with Scikit-learn, Ensemble Learning-Random Forest, AdaBoost, Gradient Tree Boosting, Voting Classifier.

Clustering Fundamentals: Basics, K-means: Finding optimal number of clusters, DBSCAN, Spectral Clustering. Evaluation methods based on Ground Truth- Homogeneity, Completeness, Adjusted Rand Index.

Introduction to Meta Classifier : Concepts of Weak and eager learner, Ensemble methods, Bagging, Boosting, Random Forests. **(Refer chapter 5)**

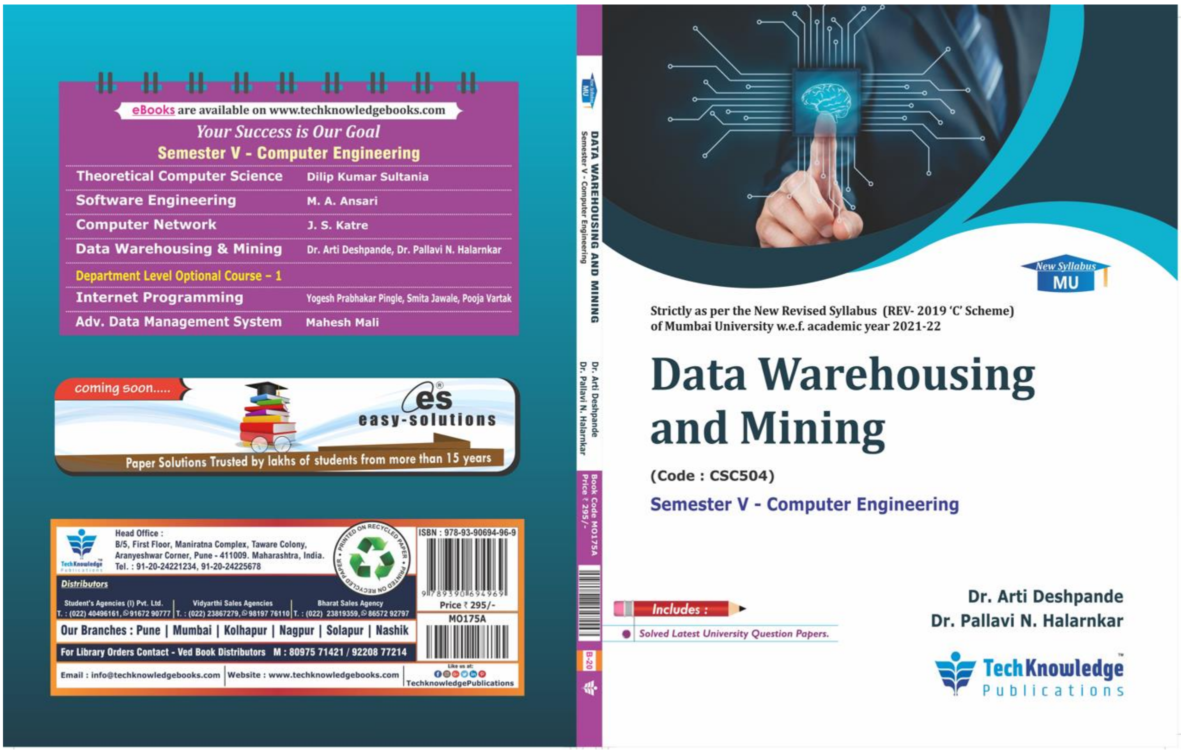
Unit VI : Clustering Techniques **(08 Hours)**

Hierarchical Clustering, Expectation maximization clustering, Agglomerative Clustering-Dendrograms, Agglomerative clustering in Scikit- learn, Connectivity Constraints.

Introduction to Recommendation Systems : Naïve User based systems, Content based Systems, Model free collaborative filtering-singular value decomposition, alternating least squares.

Fundamentals of Deep Networks: Defining Deep learning, common architectural principles of deep networks, building blocks of deep networks. **(Refer chapter 6)**

❑❑❑



TechKnowledge Publications

Head Office :
B/5, First Floor, Maniratna Complex, Taware Colony,
Aranyeshwar Corner, Pune - 411009. Maharashtra, India.
Tel. : 91-20-24221234, 91-20-24225678

PRINTED ON RECYCLED PAPER • PRINTED ON RECYCLED PAPER

Distributors

Student's Agencies (I) Pvt. Ltd. T. : (022) 40496161, 91672 90777
Vidyarthi Sales Agencies T. : (022) 23867279, 98197 76110
Bharat Sales Agency T. : (022) 23819359, 86572 92797

Our Branches : Pune | Mumbai | Kolhapur | Nagpur | Solapur | Nashik

For Library Orders Contact - Ved Book Distributors M : 80975 71421 / 92208 77214

Email : info@techknowledgebooks.com | Website : www.techknowledgebooks.com | Like us at: TechknowledgePublications

ISBN : 978-93-90694-96-9
9 789390 694969
Price ₹ 295/-
MO175A

New Syllabus MU

DATA WAREHOUSING AND MINING
Semester V - Computer Engineering

Dr. Arti Deshpande
Dr. Pallavi N. Halarnkar

Book Code MO175A
Price ₹ 295/-

B-20

New Syllabus MU

Strictly as per the New Revised Syllabus (REV- 2019 'C' Scheme)
of Mumbai University w.e.f. academic year 2021-22

Data Warehousing and Mining

(Code : CSC504)

Semester V - Computer Engineering

Includes :

- *Solved Latest University Question Papers.*

Dr. Arti Deshpande
Dr. Pallavi N. Halarnkar

TechKnowledge™ Publications

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE ★ BANDRA MUMBAI-50.

Data Warehousing and Mining

(Code : CSC504)

Semester V : Computer Engineering (Mumbai University)

Strictly as per the New Revised Syllabus (Rev-2019 'C' Scheme)

of Mumbai University w.e.f. academic year 2021-2022

(As per Choice Based Credit and Grading System)

Dr. Arti Deshpande

Department of Computer Engineering

Thadomal Shahani Engineering College, Mumbai.

Maharashtra, India

Dr. Pallavi Halarnkar

TechKnowledge™
Publications

MO175A Price ₹ 295/-

Data Warehousing and Mining (CSC504)

Dr. Arti Deshpande, Dr. Pallavi Halarnkar

Semester V : Computer Engineering (Mumbai University)

First Printed in India : February 2019

First Edition : July 2021

ISBN : *978-93-90694-96-9*

Published By

TECHKNOWLEDGE PUBLICATIONS

Printed @

37/2, Ashtavinayak Industrial Estate,
Near Pari Company,
Narhe, Pune, Maharashtra State, India.
Pune - 411041

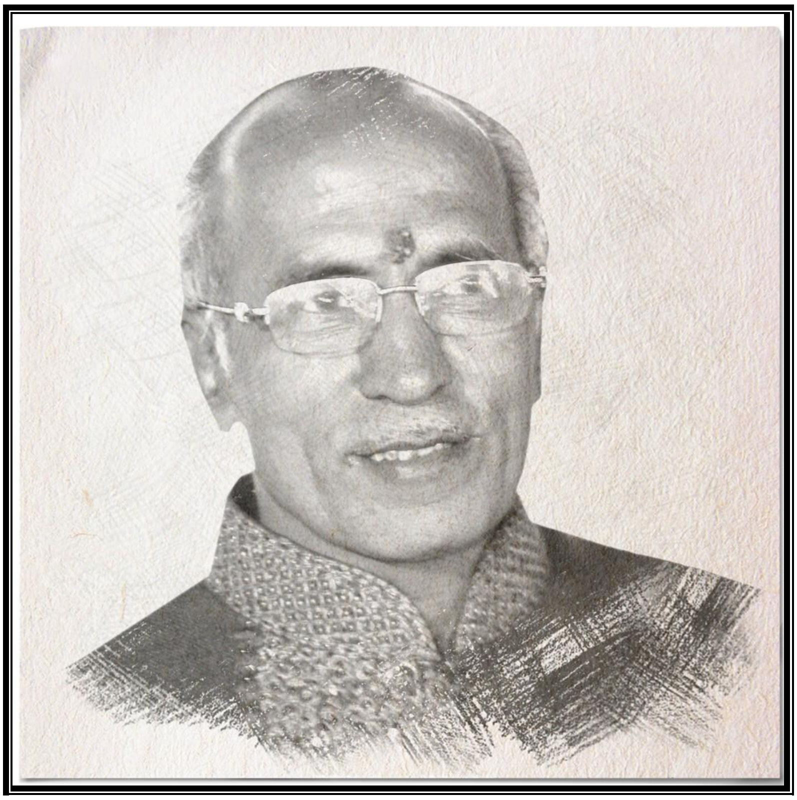
Head Office

B/5, First floor, Maniratna Complex, Taware Colony,
Aranyeshwar Corner, Pune - 411 009.
Maharashtra State, India
Ph : 91-20-24221234, 91-20-24225678.
Email : info@techknowledgebooks.com,
Website : www.techknowledgebooks.com

Subject Code : CSC504

Book Code : MO175A

We dedicate this Publication soulfully and wholeheartedly,
in loving memory of our beloved founder director,
Late Shri. Pradeepji Lalchandji Lunawat,
who will always be an inspiration, a positive force and strong support
behind us.



"My work is my prayer to God"

– Lt. Shri. Pradeepji L. Lunawat

Soulful Tribute and Gratitude for all Your
Sacrifices, Hardwork and 40 years of Strong Vision…

Preface

My Dear Students,

We are extremely happy to come out with this book on **" Data Warehousing and Mining"** for you. The topics within the chapters have been arranged in a proper sequence to ensure smooth flow of the subject.

We present this book in the loving memory of Late Shri. Pradeepji Lunawat, our source of inspiration and a strong foundation of "TechKnowledge Publications". He will always be remembered in our heart and motivate us to achieve our milestone.

We are thankful to Prof. J. S. Katre, Mr. Shital Bhandari, Prof. Arunoday Kumar and Shri. Chandroday Kumar for the encouragement and support that they have extended. We also thankful to Seema Lunawat for technology enhanced reading, E-books support and the staff members of TechKnowledge Publications for their efforts to make this book as good as it is. We have jointly made every possible efforts to eliminate all the errors in this book. However if you find any, please let me know, because that will help me to improve further.

We are thankful to my family members and friends for patience and encouragement.

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI

- Authors

❑❑❑

SYLLABUS

Mumbai University
Third Year of Computer Engineering (2019 Course)

Subject Code	Subject Name	Credits
CSC504	Data Warehousing and Mining	03

Prerequisite : Database Concepts

Course Objectives :

1	To identify the significance of Data Warehousing and Mining.
2	To analyze data, choose relevant models and algorithms for respective applications.
3	To study web data mining.
4	To develop research interest towards advances in data mining.

Course Outcomes : At the end of the course, the student will be able to

1	Understand data warehouse fundamentals and design data warehouse with dimensional modelling and apply OLAP operations.
2	Understand data mining principles and perform Data preprocessing and Visualization.
3	Identify appropriate data mining algorithms to solve real world problems.
4	Compare and evaluate different data mining techniques like classification, prediction, clustering and association rule mining.
5.	Describe complex information and social networks with respect to web mining

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA MUMBAI-50.

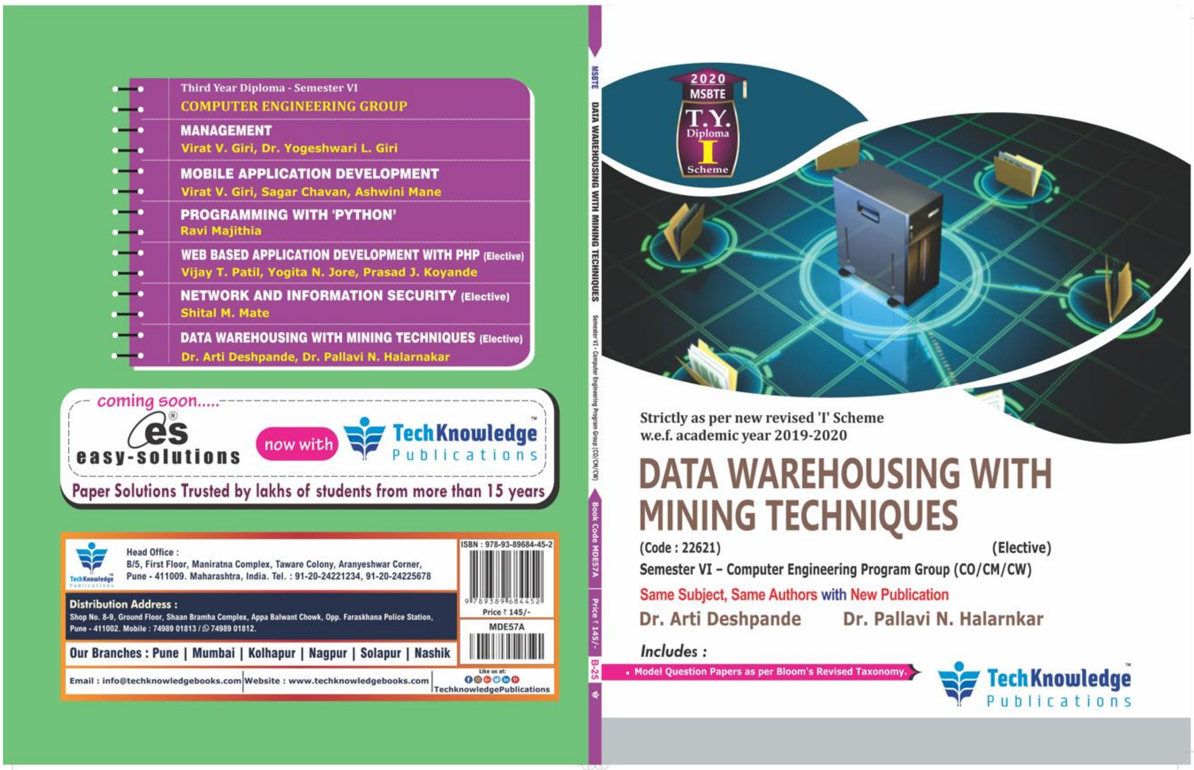
Detailed Syllabus :

Module	Course Module / Contents	Periods
1	**Data Warehousing Fundamentals**	08
	Introduction to Data Warehouse, Data warehouse architecture, Data warehouse versus Data Marts, E-R Modeling versus Dimensional Modeling, Information Package Diagram, Data Warehouse Schemas; Star Schema, Snowflake Schema, Factless Fact Table, Fact Constellation Schema. Update to the dimension tables. Major steps in ETL process, OLTP versus OLAP, OLAP operations: Slice, Dice, Rollup, Drilldown and Pivot. **(Refer Chapter 1)**	
2	**Introduction to Data Mining, Data Exploration and Data Pre-processing**	08
	Data Mining Task Primitives, Architecture, KDD process, Issues in Data Mining, Applications of Data Mining, Data Exploration: Types of Attributes, Statistical Description of Data, Data Visualization, Data Preprocessing: Descriptive data summarization, Cleaning, Integration & transformation, Data reduction, Data Discretization and Concept hierarchy generation. **(Refer Chapter 2)**	
3	**Classification**	06
	Basic Concepts, Decision Tree Induction, Naïve Bayesian Classification, Accuracy and Error measures, Evaluating the Accuracy of a Classifier: Holdout & Random Subsampling, Cross Validation, Bootstrap. **(Refer Chapter 3)**	
4	**Clustering**	06
	Types of data in Cluster analysis, Partitioning Methods (k-Means, k-Medoids), Hierarchical Methods (Agglomerative, Divisive). **(Refer Chapter 4)**	
5	**Mining Frequent Patterns and Associations**	06
	Market Basket Analysis, Frequent Item sets, Closed Item sets, and Association Rule, Frequent Pattern Mining, Apriori Algorithm, Association Rule Generation, Improving the Efficiency of Apriori, Mining Frequent Itemsets without candidate generation, Introduction to Mining Multilevel Association Rules and Mining Multidimensional Association Rules. **(Refer Chapter 5)**	
6	**Web Mining**	05
	Introduction, Web Content Mining: Crawlers, Harvest System, Virtual Web View, Personalization, Web Structure Mining: Page Rank, Clever, Web Usage Mining **(Refer Chapter 6)**	

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE ★ BANDRA MUMBAI-50.

❑❑❑



Third Year Diploma - Semester VI

COMPUTER ENGINEERING GROUP

MANAGEMENT
Virat V. Giri, Dr. Yogeshwari L. Giri

MOBILE APPLICATION DEVELOPMENT
Virat V. Giri, Sagar Chavan, Ashwini Mane

PROGRAMMING WITH 'PYTHON'
Ravi Majithia

WEB BASED APPLICATION DEVELOPMENT WITH PHP (Elective)
Vijay T. Patil, Yogita N. Jore, Prasad J. Koyande

NETWORK AND INFORMATION SECURITY (Elective)
Shital M. Mate

DATA WAREHOUSING WITH MINING TECHNIQUES (Elective)
Dr. Arti Deshpande, Dr. Pallavi N. Halarnakar

TechKnowledge™ Publications
Head Office :
B/5, First Floor, Maniratna Complex, Taware Colony, Aranyeshwar Corner,
Pune - 411009. Maharashtra, India. Tel. : 91-20-24221234, 91-20-24225678

Distribution Address :
Shop No. 8-9, Ground Floor, Shaan Bramha Complex, Appa Balwant Chowk, Opp. Faraskhana Police Station,
Pune - 411002. Mobile : 74989 01813 / ✆ 74989 01812.

Our Branches : Pune | Mumbai | Kolhapur | Nagpur | Solapur | Nashik

Email : info@techknowledgebooks.com | Website : www.techknowledgebooks.com | Like us at: TechknowledgePublications

ISBN : 978-93-89684-45-2
9 789389 684452
Price ₹ 145/-
MDE57A

MSBTE DATA WAREHOUSING WITH MINING TECHNIQUES Semester VI - Computer Engineering Program Group (CO/CM/CW) Book Code MDE57A | Price ₹ 145/- | B-25

2020 MSBTE T.Y. Diploma I Scheme

Strictly as per new revised 'I' Scheme
w.e.f. academic year 2019-2020

DATA WAREHOUSING WITH MINING TECHNIQUES

(Code : 22621) (Elective)

Semester VI – Computer Engineering Program Group (CO/CM/CW)

Same Subject, Same Authors with New Publication

Dr. Arti Deshpande Dr. Pallavi N. Halarnkar

Includes :

- Model Question Papers as per Bloom's Revised Taxonomy.

TechKnowledge™ Publications

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA MUMBAI-50.

Data Warehousing and Mining

(Code - CSC603)

Semester VI - Computer Engineering

(Mumbai University)

Strictly as per the Choice Based Credit and Grading System (Revise 2016) of Mumbai University w.e.f. academic year 2018-2019

Dr. Arti Deshpande

Department of Computer Engineering
Thadomal Shahani Engineering College,
Mumbai.
Maharashtra, India

Dr. Pallavi Halarnkar

Department of Computer Engineering
Thadomal Shahani Engineering College,
Mumbai.
Maharashtra, India.

TechKnowledge™ Publications

ME40A Price ₹ 325/-

(Book Code : ME40A)

Data Warehousing and Mining
Dr. Arti Deshpande, Dr. Pallavi Halarnkar
(Semester VI - Computer Engineering, Mumbai University)

First Edition : February 2019
Second Revised Edition : January 2020 **(TechKnowledge Publications)**

Printed at : 37/2, Ashtavinayak Industrial Estate,
Near Pari Company, Narhe, Pune, Maharashtra State, India.
Pune – 411041.

ISBN : 978-93-89748-90-1

Published by

TechKnowledge Publications

Head Office : B/5, First floor, Maniratna Complex, Taware Colony, Aranyeshwar Corner,
Pune - 411 009. Maharashtra State, India
Ph : 91-20-24221234, 91-20-24225678.
Email : info@techknowledgebooks.com,
Website : www.techknowledgebooks.com

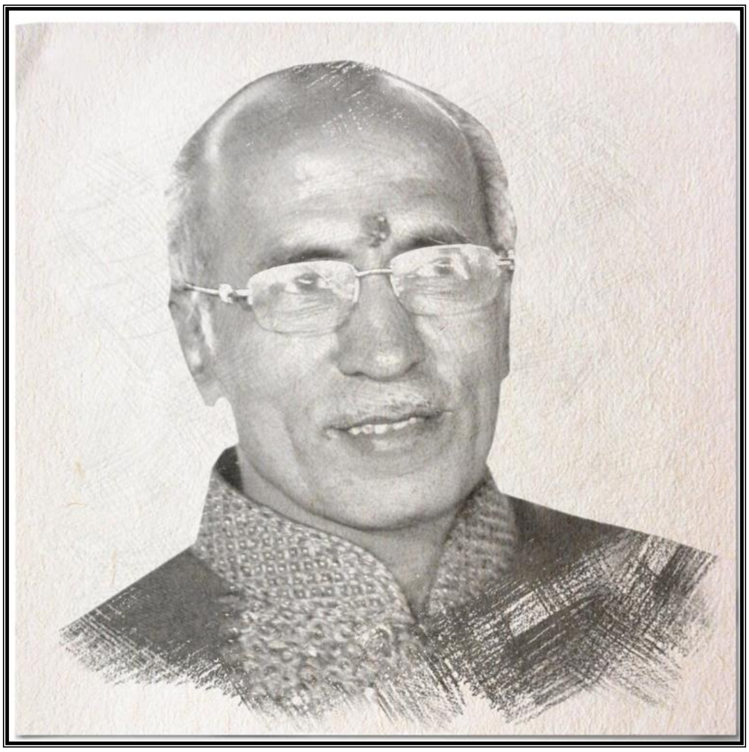
Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA MUMBAI-50.

[CSC603] (FID : ME40) (Book Code : ME40A)

(Book Code : ME40A)

We dedicate this Publication soulfully and wholeheartedly, in loving memory of our beloved founder director, ***Late Shri. Pradeepji Lalchandji Lunawat,*** *who will always be an inspiration, a positive force and strong support behind us.*



"My work is my prayer to God"

- Lt. Shri. Pradeepji L. Lunawat

Soulful Tribute and Gratitude for all Your Sacrifices, Hardwork and 40 years of Strong Vision…

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

(Book Code : ME40A)

Preface

Dear Students,

We are extremely happy to present the book of **“Data Warehousing and Mining”** for you. We have divided the subject into small chapters so that the topics can be arranged and understood properly. The topics within the chapters have been arranged in a proper sequence to ensure smooth flow of the subject.

We present this book in the loving memory of **Late. Shri. Pradeepji Lunawat**, our source of inspiration and a strong foundation of **“TechKnowledge Publications”.** He will always be remembered in our hearts and motivate us to achieve our new milestone.

We are thankful to Prof. Arunoday Kumar, Mr. Shital Bhandari and Shri. Chandroday Kumar for the encouragement and support that they have extended. We also thankful to the staff members of TechKnowledge Publications for their efforts to make this book as good as it is. We have made every possible efforts to eliminate all the errors in this book. However if you find any, please let us know, because that will help us to improve the book quality further.

We are thankful to my family members and friends for their patience and encouragement.

- Authors

❑❑❑

(Book Code : ME40A)

Syllabus

Course Code	Course Name	Credits
CSC603	**Data Warehousing and Mining**	**4**

Course Objectives :

1. To identify the scope and essentiality of Data Warehousing and Mining.
2. To analyze data, choose relevant models and algorithms for respective applications.
3. To study spatial and web data mining.
4. To develop research interest towards advances in data mining.

Course Outcomes : On successful completion of course learner will be able to :

1. Understand Data Warehouse fundamentals, Data Mining Principles.
2. Design data warehouse with dimensional modelling and apply OLAP operations.
3. Identify appropriate data mining algorithms to solve real world problems.
4. Compare and evaluate different data mining techniques like classification, prediction, clustering and association rule mining.
5. Describe complex data types with respect to spatial and web mining.
6. Benefit the user experiences towards research and innovation.

Prerequisite : Basic database concepts, Concepts of algorithm design and analysis.

Module No.	Topics	Hrs.
1.0	**Introduction to Data Warehouse and Dimensional modeling :** Introduction to Strategic Information, Need for Strategic Information, Features of Data Warehouse, Data warehouses versus Data Marts, Top-down versus Bottom-up approach. Data warehouse architecture, metadata, E-R modelling versus Dimensional Modelling, Information Package Diagram, STAR schema, STAR schema keys, Snowflake Schema, Fact Constellation Schema, Factless Fact tables, Update to the dimension tables, Aggregate fact tables.**(Refer chapter 1)**	**8**
2.0	**ETL Process and OLAP :** Major steps in ETL process, Data extraction : Techniques, Data transformation : Basic tasks, Major transformation types, Data Loading : Applying Data, OLTP Vs OLAP, OLAP definition, Dimensional Analysis, Hypercubes, OLAP operations : Drill down, Roll up, Slice, Dice and Rotation, OLAP models : MOLAP, ROLAP. **(Refer chapter 2)**	**8**

(Book Code : ME40A)

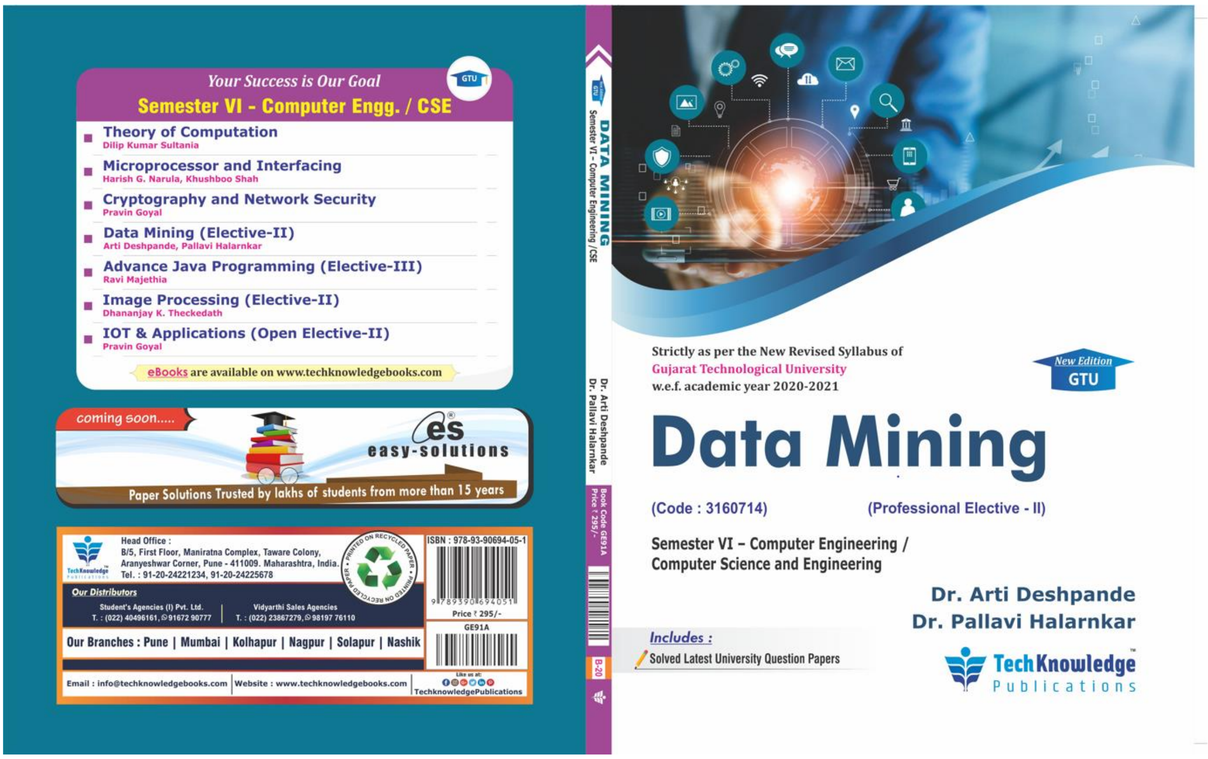
Module No.	Topics	Hrs.
3.0	**Introduction to Data Mining, Data Exploration and Preprocessing :** Data Mining Task Primitives, Architecture, Techniques, KDD process, Issues in Data Mining, Applications of Data Mining, Data Exploration : Types of Attributes, Statistical Description of Data, Data Visualization, Data Preprocessing : Cleaning, Integration, Reduction : Attribute subset selection, Histograms, Clustering and Sampling, Data Transformation and Data Discretization : Normalization, Binning, Concept hierarchy generation, Concept Description : Attribute oriented Induction for Data Characterization. **(Refer chapter 3)**	10
4.0	**Classification, Prediction and Clustering :** Basic Concepts, Decision Tree using Information Gain, Induction : Attribute Selection Measures, Tree pruning, Bayesian Classification : Naive Bayes, Classifier Rule - Based Classification : Using IF-THEN Rules for classification, Prediction : Simple linear regression, Multiple linear regression Model Evaluation and Selection : Accuracy and Error measures, Holdout, Random Sampling, Cross Validation, Bootstrap, Clustering : Distance Measures, Partitioning Methods (k-Means, k-Medoids), Hierarchical Methods (Agglomerative, Divisive). **(Refer chapter 4)**	12
5.0	**Mining Frequent Patterns and Association Rules :** Market Basket Analysis, Frequent Item sets, Closed Item sets and Association Rule, Frequent Pattern Mining, Efficient and Scalable Frequent Item set Mining Methods : Apriori Algorithm, Association Rule Generation, Improving the Efficiency of Apriori, FP growth, Mining frequent Itemsets using Vertical Data Format, Introduction to Mining Multilevel Association Rules and Multidimensional Association Rules. **(Refer chapter 5)**	8
6.0	**Spatial and Web Mining :** Spatial Data, Spatial Vs. Classical Data Mining, Spatial Data Structures, Mining Spatial Association and Co-location Patterns, Spatial Clustering Techniques : CLARANS Extension, Web Mining : Web Content Mining, Web Structure Mining, Web Usage mining, Applications of Web Mining. **(Refer chapter 6)**	6
	Total	**52**

❑❑❑

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

(Book Code : ME40A)



Your Success is Our Goal

Semester VI - Computer Engg. / CSE

TechKnowledge Publications

Head Office :
B/5, First Floor, Maniratna Complex, Taware Colony,
Aranyeshwar Corner, Pune - 411009. Maharashtra, India.
Tel. : 91-20-24221234, 91-20-24225678

PRINTED ON RECYCLED PAPER

Our Distributors

Student's Agencies (I) Pvt. Ltd.
T. : (022) 40496161, 91672 90777

Vidyarthi Sales Agencies
T. : (022) 23867279, 98197 76110

Our Branches : Pune | Mumbai | Kolhapur | Nagpur | Solapur | Nashik

ISBN : 978-93-90694-05-1

Price ₹ 295/-

GE91A

Email : info@techknowledgebooks.com | Website : www.techknowledgebooks.com | Like us at: TechknowledgePublications

DATA MINING
Semester VI - Computer Engineering / CSE
Dr. Arti Deshpande
Dr. Pallavi Halarnkar
Book Code GE91A
Price ₹ 295/-
B-20

Strictly as per the New Revised Syllabus of
Gujarat Technological University
w.e.f. academic year 2020-2021

New Edition GTU

Data Mining

(Code : 3160714) (Professional Elective - II)

Semester VI – Computer Engineering /
Computer Science and Engineering

Dr. Arti Deshpande
Dr. Pallavi Halarnkar

Includes :
Solved Latest University Question Papers

TechKnowledge Publications

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA MUMBAI-50.

Data Mining

(Code : 3160714) (Professional Elective-II)

Semester VI - Computer Engineering/Computer Science and Engineering,
(Gujarat Technological University)

Strictly as per the New Revised Syllabus of

Gujarat Technological University w.e.f. academic year 2020-2021

Dr. Arti Deshpande
Assistant Professor,
Department of Computer Engineering
Thadomal Shahani Engineering College,
Mumbai.
Maharashtra, India

Dr. Pallavi N. Halarnkar
Associate Professor,
Department of Computer Engineering
Thadomal Shahani Engineering College,
Mumbai.
Maharashtra, India.

TechKnowledge™
Publications

GE91A Price ₹ 295/-

Data Mining (3160714)

Dr. Arti Deshpande, Dr. Pallavi N. Halarnkar

Semester VI – Computer Engineering/Computer Science and Engineering, (Gujarat Technological University)

First Printed In India : April 2010 (Mumbai University)

First Edition : February 2021

ISBN : 978-93-90694-05-1

Published By

TechKnowledge Publications

Printed @

37/2, Ashtavinayak Industrial Estate,
Near Pari Company,
Narhe, Pune, Maharashtra State, India.
Pune - 411041

Head Office

B/5, First floor, Maniratna Complex,
Taware Colony, Aranyeshwar Corner,
Pune - 411 009. Maharashtra State, India
Ph : 91-20-24221234, 91-20-24225678.
Email : info@techknowledgebooks.com,
Website : www.techknowledgebooks.com

Subject Code : 3160714
Book Code : GE91A

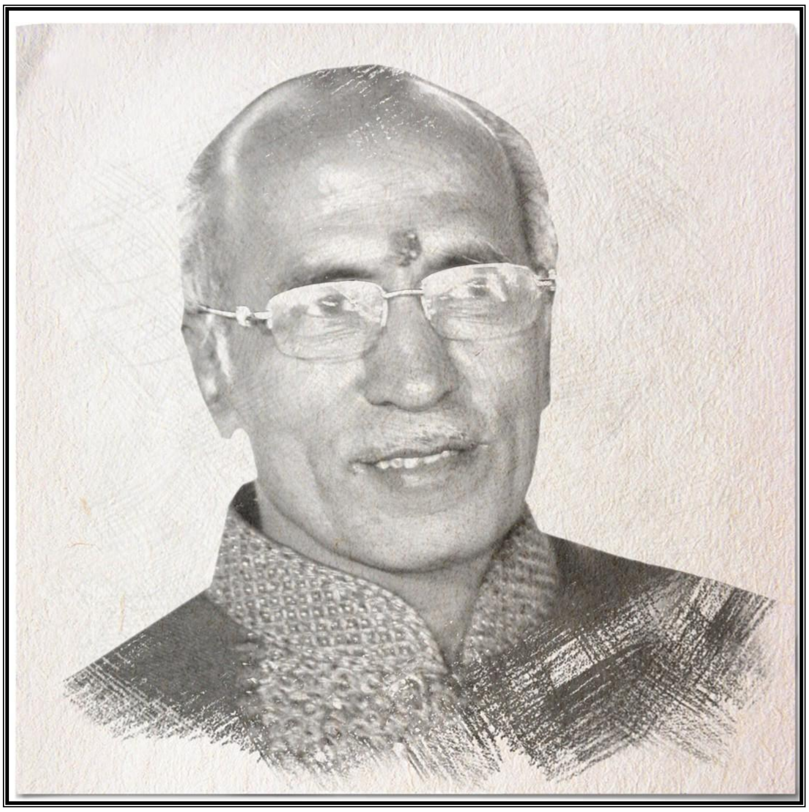
We dedicate this Publication soulfully and wholeheartedly,

in loving memory of our beloved founder director,

Late Shri. Pradeepji Lalchandji Lunawat,

who will always be an inspiration, a positive force and strong support

behind us.



"My work is my prayer to God"

– Lt. Shri. Pradeepji L. Lunawat

Soulful Tribute and Gratitude for all Your

Sacrifices, Hardwork and 40 years of Strong Vision…

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

Preface

My Dear Students,

We are extremely happy to come out with this book on **Data Mining** for you. The topics within the chapters have been arranged in a proper sequence to ensure smooth flow of the subject.

We present this book in the loving memory of **Late Shri. Pradeepji Lunawat,** our source of inspiration and a strong foundation of **"TechKnowledge Publications"**. He will always be remembered in our heart and motivate us to achieve our milestone.

We are thankful to Shri. J. S. Katre, Shri. Shital Bhandari, Shri. Arunoday Kumar and Shri. Chandroday Kumar for the encouragement and support that they have extended. We are also thankful to Seema Lunawat for technology enhanced reading, E-books support and the staff members of TechKnowledge Publications for their efforts to make this book as good as it is. We have jointly made every possible efforts to eliminate all the errors in this book. However if you find any, please let us know, because that will help us to improve further.

We are also thankful to our family members and friends for their patience and encouragement.

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

- Authors

SYLLABUS

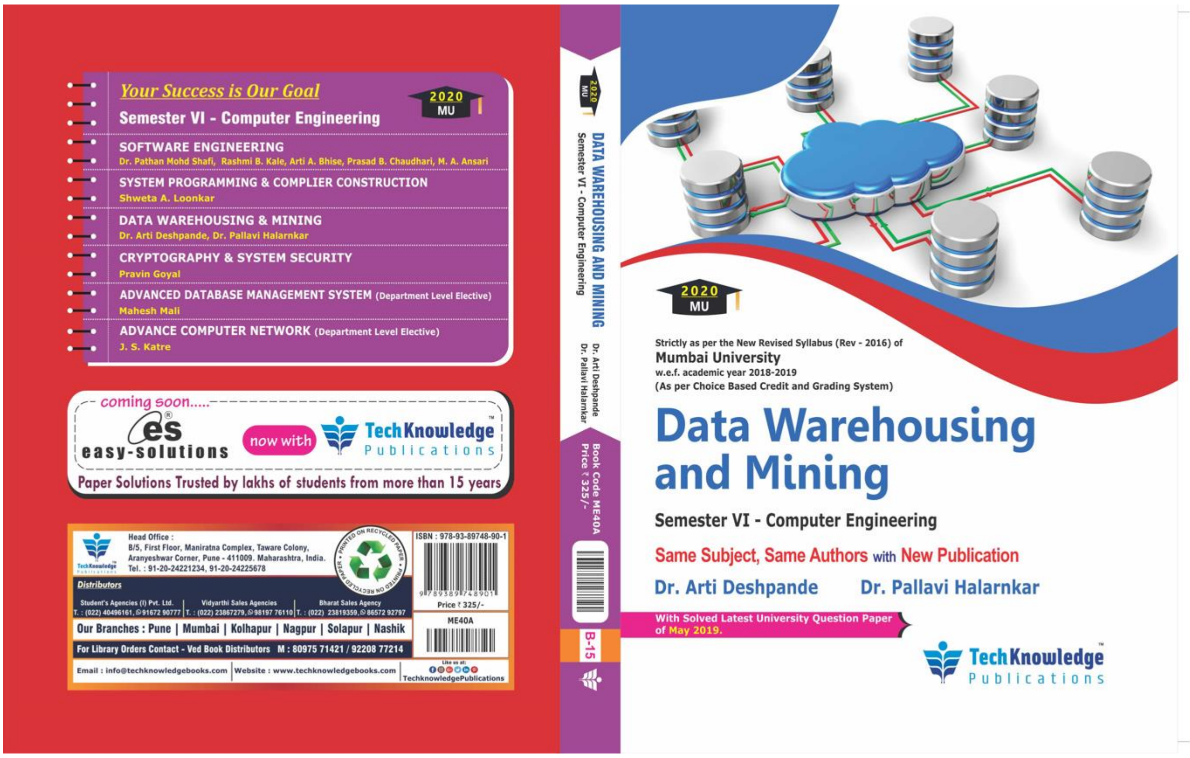
Gujarat Technological University
Sixth Semester of Computer Engineering / Computer Science and Engineering
Data Mining (Code : 3160714)

Teaching and Examination Scheme

Teaching scheme			**Credits**	**Examination Marks**				**Total Marks**
L	T	P	C	Theory Marks		Practical Marks		
				ESE (E)	PA (M)	ESE (V)	PA (I)	
3	0	2	4	70	30	30	20	150

Sr. No.	**Content**	**Total Hours**
1.	**Introduction to data mining (DM) :** Motivation for Data Mining - Data Mining-Definition and Functionalities – Classification of DM Systems - DM task primitives - Integration of a Data Mining system with a Database or a Data Warehouse - Issues in DM – KDD Process **(Refer Chapter 1)**	3
2.	**Data Pre-processing :** Data summarization, data cleaning, data integration and transformation, data reduction, data discretization and concept hierarchy generation, feature extraction, feature transformation, feature selection, introduction to Dimensionality Reduction, CUR decomposition **(Refer Chapter 2)**	4
3.	**Concept Description, Mining Frequent Patterns, Associations and Correlations :** What is concept description? - Data Generalization and summarization-based characterization - Attribute relevance - class comparisons, Basic concept, efficient and scalable frequent item-set mining methods, mining various kind of association rules, from association mining to correlation analysis, Advanced Association Rule Techniques, Measuring the Quality of Rules. **(Refer Chapter 3)**	10
4.	**Classification and Prediction :** Classification vs. prediction, Issues regarding classification and prediction, Statistical-Based Algorithms, Distance-Based Algorithms, Decision Tree-Based Algorithms, Neural Network-Based Algorithms, Rule-Based Algorithms, Combining Techniques, accuracy and error measures, evaluation of the accuracy of a classifier or predictor. Neural Network Prediction methods: Linear and nonlinear regression, Logistic Regression Introduction of tools such as DB Miner / WEKA / DTREG DM Tools **(Refer Chapter 4)**	10
5.	**Cluster Analysis :** Clustering: Problem Definition, Clustering Overview, Evaluation of Clustering Algorithms, Partitioning Clustering -K-Means Algorithm, K-Means Additional issues, PAM Algorithm; Hierarchical Clustering – Agglomerative Methods and divisive methods, Basic Agglomerative Hierarchical Clustering, Strengths and Weakness; Outlier Detection, Clustering high dimensional data, clustering Graph and Network data. **(Refer Chapter 5)**	10
8.	**Web mining and other Data Mining :** Web Mining: Introduction to Web Mining, Web content mining, Web usage mining, Web Structure mining, Web log structure and issues regarding web logs, Spatial Data Mining, Temporal Mining, And Multimedia Mining. Applications of Distributed and parallel Data Mining. **(Refer Chapter 6)**	5

❑❑❑



Your Success is Our Goal

2020 MU

Semester VI - Computer Engineering

SOFTWARE ENGINEERING
Dr. Pathan Mohd Shafi, Rashmi B. Kale, Arti A. Bhise, Prasad B. Chaudhari, M. A. Ansari

SYSTEM PROGRAMMING & COMPLIER CONSTRUCTION
Shweta A. Loonkar

DATA WAREHOUSING & MINING
Dr. Arti Deshpande, Dr. Pallavi Halarnkar

CRYPTOGRAPHY & SYSTEM SECURITY
Pravin Goyal

ADVANCED DATABASE MANAGEMENT SYSTEM (Department Level Elective)
Mahesh Mali

ADVANCE COMPUTER NETWORK (Department Level Elective)
J. S. Katre

Head Office :
B/5, First Floor, Maniratna Complex, Taware Colony,
Aranyeshwar Corner, Pune - 411009. Maharashtra, India.
Tel. : 91-20-24221234, 91-20-24225678

PRINTED ON RECYCLED PAPER

Distributors
Student's Agencies (I) Pvt. Ltd. T. : (022) 40496161, 91672 90777
Vidyarthi Sales Agencies T. : (022) 23867279, 98197 76110
Bharat Sales Agency T. : (022) 23819359, 86572 92797

Our Branches : Pune | Mumbai | Kolhapur | Nagpur | Solapur | Nashik

For Library Orders Contact - Ved Book Distributors M : 80975 71421 / 92208 77214

Email : info@techknowledgebooks.com | Website : www.techknowledgebooks.com | Like us at: TechknowledgePublications

ISBN : 978-93-89748-90-1
9789389748901
Price ₹ 325/-
ME40A

2020 MU
DATA WAREHOUSING AND MINING
Semester VI - Computer Engineering
Dr. Arti Deshpande
Dr. Pallavi Halarnkar
Book Code ME40A
Price ₹ 325/-
B-15

2020 MU

Strictly as per the New Revised Syllabus (Rev - 2016) of
Mumbai University
w.e.f. academic year 2018-2019
(As per Choice Based Credit and Grading System)

Data Warehousing and Mining

Semester VI - Computer Engineering

Same Subject, Same Authors with New Publication

Dr. Arti Deshpande   Dr. Pallavi Halarnkar

With Solved Latest University Question Paper of May 2019.

TechKnowledge Publications

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA MUMBAI-50.

Data Warehousing and Mining

(Code - CSC603)

Semester VI - Computer Engineering

(Mumbai University)

Strictly as per the Choice Based Credit and Grading System (Revise 2016) of Mumbai University w.e.f. academic year 2018-2019

Dr. Arti Deshpande

Department of Computer Engineering
Thadomal Shahani Engineering College,
Mumbai.
Maharashtra, India

Dr. Pallavi Halarnkar

Department of Computer Engineering
Thadomal Shahani Engineering College,
Mumbai.
Maharashtra, India.

TechKnowledge™
Publications

ME40A Price ₹ 325/-

(Book Code : ME40A)

Data Warehousing and Mining
Dr. Arti Deshpande, Dr. Pallavi Halarnkar
(Semester VI - Computer Engineering, Mumbai University)

First Edition : February 2019
Second Revised Edition : January 2020 **(TechKnowledge Publications)**

Printed at : 37/2, Ashtavinayak Industrial Estate,
Near Pari Company, Narhe, Pune, Maharashtra State, India.
Pune – 411041.

ISBN : 978-93-89748-90-1

Published by

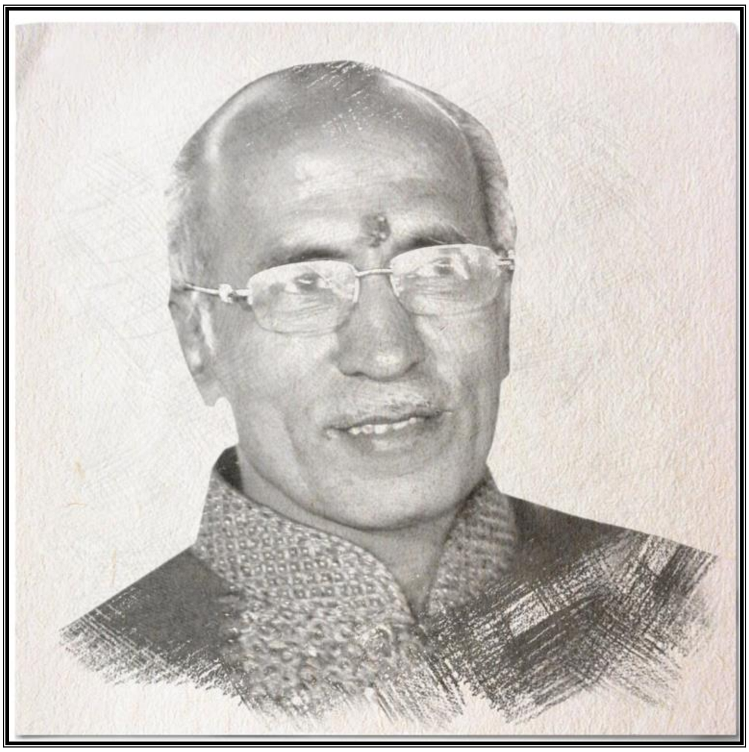
TechKnowledge Publications

Head Office : B/5, First floor, Maniratna Complex, Taware Colony, Aranyeshwar Corner,
Pune - 411 009. Maharashtra State, India
Ph : 91-20-24221234, 91-20-24225678.
Email : info@techknowledgebooks.com,
Website : www.techknowledgebooks.com

[CSC603] (FID : ME40) (Book Code : ME40A)

(Book Code : ME40A)

We dedicate this Publication soulfully and wholeheartedly, in loving memory of our beloved founder director, ***Late Shri. Pradeepji Lalchandji Lunawat,*** *who will always be an inspiration, a positive force and strong support behind us.*



"My work is my prayer to God"

- Lt. Shri. Pradeepji L. Lunawat

Soulful Tribute and Gratitude for all Your Sacrifices, Hardwork and 40 years of Strong Vision…

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

(Book Code : ME40A)

Preface

Dear Students,

We are extremely happy to present the book of **"Data Warehousing and Mining"** for you. We have divided the subject into small chapters so that the topics can be arranged and understood properly. The topics within the chapters have been arranged in a proper sequence to ensure smooth flow of the subject.

We present this book in the loving memory of **Late. Shri. Pradeepji Lunawat**, our source of inspiration and a strong foundation of **"TechKnowledge Publications".** He will always be remembered in our hearts and motivate us to achieve our new milestone.

We are thankful to Prof. Arunoday Kumar, Mr. Shital Bhandari and Shri. Chandroday Kumar for the encouragement and support that they have extended. We also thankful to the staff members of TechKnowledge Publications for their efforts to make this book as good as it is. We have made every possible efforts to eliminate all the errors in this book. However if you find any, please let us know, because that will help us to improve the book quality further.

We are thankful to my family members and friends for their patience and encouragement.

- Authors

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

❑❑❑

(Book Code : ME40A)

Syllabus

Course Code	Course Name	Credits
CSC603	**Data Warehousing and Mining**	**4**

Course Objectives :

1. To identify the scope and essentiality of Data Warehousing and Mining.
2. To analyze data, choose relevant models and algorithms for respective applications.
3. To study spatial and web data mining.
4. To develop research interest towards advances in data mining.

Course Outcomes : On successful completion of course learner will be able to :

1. Understand Data Warehouse fundamentals, Data Mining Principles.
2. Design data warehouse with dimensional modelling and apply OLAP operations.
3. Identify appropriate data mining algorithms to solve real world problems.
4. Compare and evaluate different data mining techniques like classification, prediction, clustering and association rule mining.
5. Describe complex data types with respect to spatial and web mining.
6. Benefit the user experiences towards research and innovation.

Prerequisite : Basic database concepts, Concepts of algorithm design and analysis.

Module No.	Topics	Hrs.
1.0	**Introduction to Data Warehouse and Dimensional modeling :** Introduction to Strategic Information, Need for Strategic Information, Features of Data Warehouse, Data warehouses versus Data Marts, Top-down versus Bottom-up approach. Data warehouse architecture, metadata, E-R modelling versus Dimensional Modelling, Information Package Diagram, STAR schema, STAR schema keys, Snowflake Schema, Fact Constellation Schema, Factless Fact tables, Update to the dimension tables, Aggregate fact tables.**(Refer chapter 1)**	**8**
2.0	**ETL Process and OLAP :** Major steps in ETL process, Data extraction : Techniques, Data transformation : Basic tasks, Major transformation types, Data Loading : Applying Data, OLTP Vs OLAP, OLAP definition, Dimensional Analysis, Hypercubes, OLAP operations : Drill down, Roll up, Slice, Dice and Rotation, OLAP models : MOLAP, ROLAP. **(Refer chapter 2)**	**8**

(Book Code : ME40A)

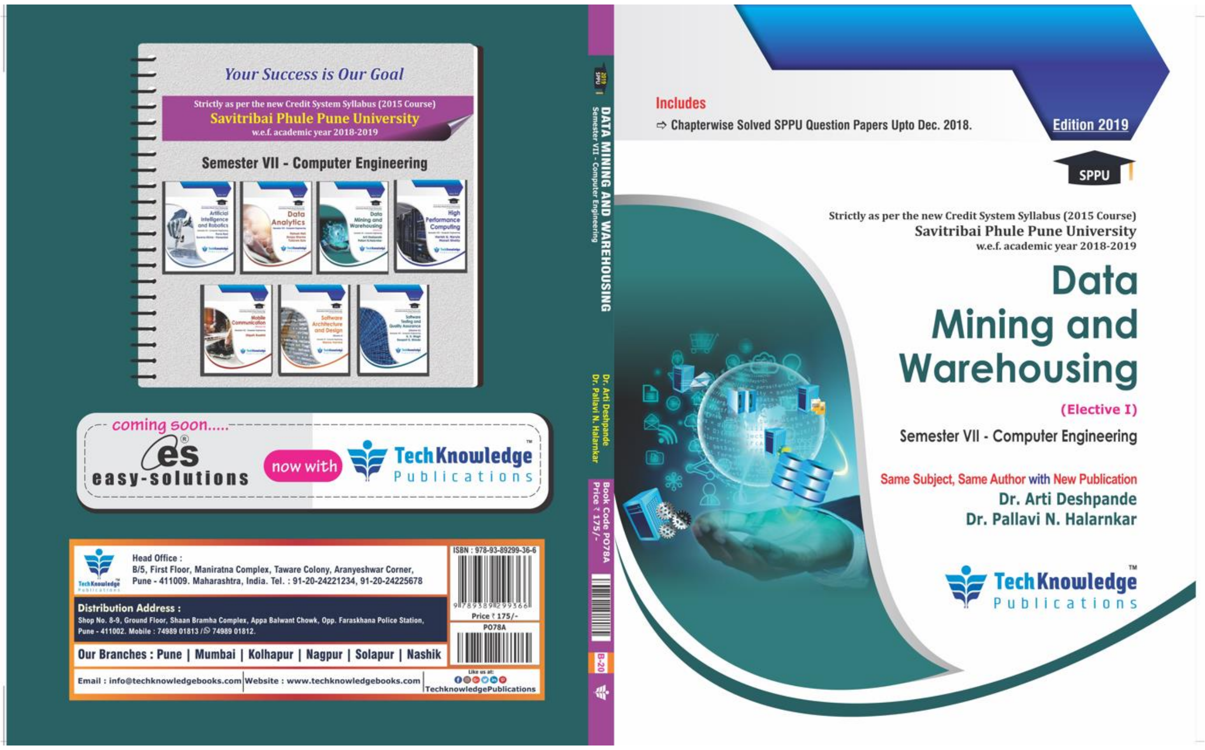
Module No.	Topics	Hrs.
3.0	**Introduction to Data Mining, Data Exploration and Preprocessing :** Data Mining Task Primitives, Architecture, Techniques, KDD process, Issues in Data Mining, Applications of Data Mining, Data Exploration : Types of Attributes, Statistical Description of Data, Data Visualization, Data Preprocessing : Cleaning, Integration, Reduction : Attribute subset selection, Histograms, Clustering and Sampling, Data Transformation and Data Discretization : Normalization, Binning, Concept hierarchy generation, Concept Description : Attribute oriented Induction for Data Characterization. **(Refer chapter 3)**	**10**
4.0	**Classification, Prediction and Clustering :** Basic Concepts, Decision Tree using Information Gain, Induction : Attribute Selection Measures, Tree pruning, Bayesian Classification : Naive Bayes, Classifier Rule - Based Classification : Using IF-THEN Rules for classification, Prediction : Simple linear regression, Multiple linear regression Model Evaluation and Selection : Accuracy and Error measures, Holdout, Random Sampling, Cross Validation, Bootstrap, Clustering : Distance Measures, Partitioning Methods (k-Means, k-Medoids), Hierarchical Methods (Agglomerative, Divisive). **(Refer chapter 4)**	**12**
5.0	**Mining Frequent Patterns and Association Rules :** Market Basket Analysis, Frequent Item sets, Closed Item sets and Association Rule, Frequent Pattern Mining, Efficient and Scalable Frequent Item set Mining Methods : Apriori Algorithm, Association Rule Generation, Improving the Efficiency of Apriori, FP growth, Mining frequent Itemsets using Vertical Data Format, Introduction to Mining Multilevel Association Rules and Multidimensional Association Rules. **(Refer chapter 5)**	**8**
6.0	**Spatial and Web Mining :** Spatial Data, Spatial Vs. Classical Data Mining, Spatial Data Structures, Mining Spatial Association and Co-location Patterns, Spatial Clustering Techniques : CLARANS Extension, Web Mining : Web Content Mining, Web Structure Mining, Web Usage mining, Applications of Web Mining. **(Refer chapter 6)**	**6**
	Total	**52**

❑❑❑

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA MUMBAI-50.

(Book Code : ME40A)



Includes

⇨ Chapterwise Solved SPPU Question Papers Upto Dec. 2018.

Edition 2019

SPPU

Strictly as per the new Credit System Syllabus (2015 Course)
Savitribai Phule Pune University
w.e.f. academic year 2018-2019

Data Mining and Warehousing

(Elective I)

Semester VII - Computer Engineering

Same Subject, Same Author with New Publication

Dr. Arti Deshpande
Dr. Pallavi N. Halarnkar

TechKnowledge Publications

2019 SPPU

DATA MINING AND WAREHOUSING
Semester VII - Computer Engineering

Dr. Arti Deshpande
Dr. Pallavi N. Halarnkar

Book Code PO78A
Price ₹ 175/-

B-20

Your Success is Our Goal

Strictly as per the new Credit System Syllabus (2015 Course)
Savitribai Phule Pune University
w.e.f. academic year 2018-2019

Semester VII - Computer Engineering

Artificial Intelligence and Robotics

Data Analytics

Data Mining and Warehousing

High Performance Computing

Mobile Communication

Software Architecture and Design

Software Testing and Quality Assurance

coming soon.....

es® easy-solutions now with TechKnowledge™ Publications

Head Office :
B/5, First Floor, Maniratna Complex, Taware Colony, Aranyeshwar Corner,
Pune - 411009. Maharashtra, India. Tel. : 91-20-24221234, 91-20-24225678

Distribution Address :
Shop No. 8-9, Ground Floor, Shaan Bramha Complex, Appa Balwant Chowk, Opp. Faraskhana Police Station,
Pune - 411002. Mobile : 74989 01813 / ✆ 74989 01812.

Our Branches : Pune | Mumbai | Kolhapur | Nagpur | Solapur | Nashik

Email : info@techknowledgebooks.com | Website : www.techknowledgebooks.com | Like us at: TechknowledgePublications

ISBN : 978-93-89299-36-6

9 789389 299366

Price ₹ 175/-

PO78A

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA MUMBAI-50.

Data Mining and Warehousing

Elective I

(Code : 410244(D))

Semester VII - Computer Engineering

(Savitribai Phule Pune University)

Strictly as per the New Credit System Syllabus (2015 Course)
Savitribai Phule Pune University w.e.f. academic year 2018-2019

Dr. Arti Deshpande

Assistant Professor,
Department of Computer Engineering
Thadomal Shahani Engineering College, Mumbai.
Maharashtra, India.

Dr. Pallavi N. Halarnkar

Associate Professor,
Department of Computer Engineering
Thadomal Shahani Engineering College, Mumbai.
Maharashtra, India.

TechKnowledge™
Publications

PO78A Price ₹ 175/-

Data Mining and Warehousing

Dr. Arti Deshpande, Dr. Pallavi N. Halarnkar

(Semester VII - Computer Engineering) (Savitribai Phule Pune University)

First Edition : July 2018
Second Revised Edition : July 2019 **(TechKnowledge Publications)**

This edition is for sale in India, Bangladesh, Bhutan, Maldives, Nepal, Pakistan, Sri Lanka and designated countries in South-East Asia. Sale and purchase of this book outside of these countries is unauthorized by the publisher.

Printed at :: 37/2, Ashtvinayak Industrial Estate, Near Pari Company,
Narhe, Pune, Maharashtra State India, Pune - 411041

ISBN 978-93-89299-36-6

Published by
TechKnowledge Publications

Head Office : B/5, First floor, Maniratna Complex,
Taware Colony, Aranyeshwar Corner,
Pune - 411 009. Maharashtra State,
India. Ph : 91-20-24221234, 91-20-24225678.

[410244(D)] (FID : PO78) (Book Code : PO78A)

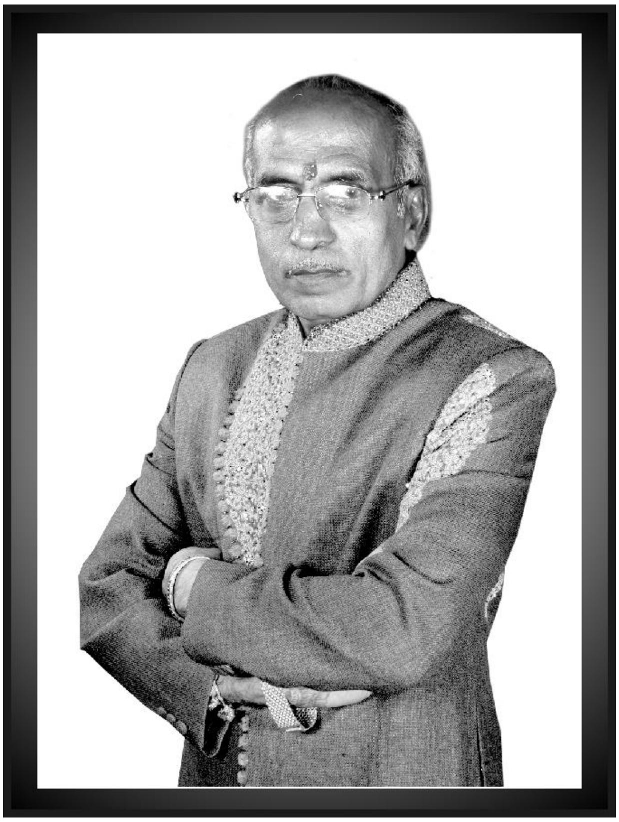
We dedicate this Publication soulfully and wholeheartedly,

in loving memory of our beloved founder director

Late. Shri. Pradeepsheth Lalchandji Lunawat,

who will always be an inspiration, a positive force and strong support

behind us.



Lt. Shri. Pradeepji L. Lunawat

Soulful Tribute and Gratitude for all Your

Sacrifices, Hardwork and 40 years of Strong Vision……

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai-400 050

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

Preface

Dear Students,

We are extremely happy to present the book of **"Data Mining and Warehousing"** for you. We have divided the subject into small chapters so that the topics can be arranged and understood properly. The topics within the chapters have been arranged in a proper sequence to ensure smooth flow of the subject.

We present this book in the loving memory **of Late. Shri. Pradeepji Lunawat**, our source of inspiration and a strong foundation of **"TechKnowledge Publications"**. He will always be remembered in our heart and motivate us to achieve our milestone.

We are thankful to Shri. J. S. Katre, Shri. Shital Bhandari, Shri. Arunoday Kumar and Shri. Chandroday Kumar for the encouragement and support that they have extended. We are also thankful to the staff members of TechKnowledge Publications and others for their efforts to make this book as good as it is.

We have jointly made every possible effort to eliminate all the errors in this book. However if you find any, please let us know, because that will help us to improve further.

- Dr. Arti Deshpande

- Dr. Pallavi Halarnkar

❏❏❏

Syllabus

Savitribai Phule Pune University
Fourth Year of Computer Engineering (2015 Course)
Elective I

410244(D) : Data Mining and Warehousing

Teaching Scheme : TH : 03 Hours/Week	**Credit** 03	**Examination Scheme :** In-Sem (Paper) : 30 Marks End-Sem (Paper) : 70 Marks

Pre-requisites Courses

310242-Database Management Systems, 310244 - Information Systems and Engineering Economics

Companion Course : 410247- Laboratory Practice II

Course Objectives

- To understand the fundamentals of Data Mining.
- To identify the appropriateness and need of mining the data.
- To learn the preprocessing, mining and post processing of the data.
- To understand various methods, techniques and algorithms in data mining.

Course Outcomes

On completion of the course the student should be able to :

- Apply basic, intermediate and advanced techniques to mine the data.
- Analyze the output generated by the process of data mining.
- Explore the hidden patterns in the data.
- Optimize the mining process by choosing best data mining technique.

Course Contents

Unit I : Introduction (08 Hours)

Data Mining, Data Mining Task Primitives, Data : Data, Information and Knowledge; Attribute Types : Nominal, Binary, Ordinal and Numeric attributes, Discrete versus Continuous Attributes; Introduction to Data Preprocessing, Data Cleaning : Missing values, Noisy data; Data integration : Correlation analysis; transformation : Min-max normalization, z-score normalization and decimal scaling; data reduction : Data Cube Aggregation, Attribute Subset Selection, sampling; and Data Discretization : Binning, Histogram Analysis **(Refer chapter 1)**

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE ★ BANDRA MUMBAI-50.

Unit II : Data Warehouse (08 Hours)

Data Warehouse, Operational Database Systems and Data Warehouses (OLTP Vs OLAP), A Multidimensional Data Model: Data Cubes, Stars, Snowflakes, and Fact Constellations Schemas; OLAP Operations in the Multidimensional Data Model, Concept Hierarchies, Data Warehouse Architecture, The Process of Data Warehouse Design, A three-tier data warehousing architecture, Types of OLAP Servers : ROLAP versus MOLAP versus HOLAP. **(Refer chapter 2)**

Unit III : Measuring Data Similarity and Dissimilarity (08 Hours)

Measuring Data Similarity and Dissimilarity, Proximity Measures for Nominal Attributes and Binary Attributes, interval scaled; Dissimilarity of Numeric Data : Minskowski Distance, Euclidean distance and Manhattan distance; Proximity Measures for Categorical, Ordinal Attributes, Ratio scaled variables; Dissimilarity for Attributes of Mixed Types, Cosine Similarity. **(Refer chapter 3)**

Unit IV : Association Rules Mining (08 Hours)

Market basket Analysis, Frequent item set, Closed item set, Association Rules, a-priori Algorithm, Generating Association Rules from Frequent Item sets, Improving the Efficiency of a-priori, Mining Frequent Item sets without Candidate Generation : FP Growth Algorithm; Mining Various Kinds of Association Rules : Mining multilevel association rules, constraint based association rule mining, Meta rule-Guided Mining of Association Rules.

(Refer chapter 4)

Unit V : Classification (08 Hours)

Introduction to : Classification and Regression for Predictive Analysis, Decision Tree Induction, Rule-Based Classification : using IF-THEN Rules for Classification, Rule Induction Using a Sequential Covering Algorithm. Bayesian Belief Networks, Training Bayesian Belief Networks, Classification Using Frequent Patterns, Associative Classification, Lazy Learners-k-Nearest-Neighbor Classifiers, Case-Based Reasoning. **(Refer chapter 5)**

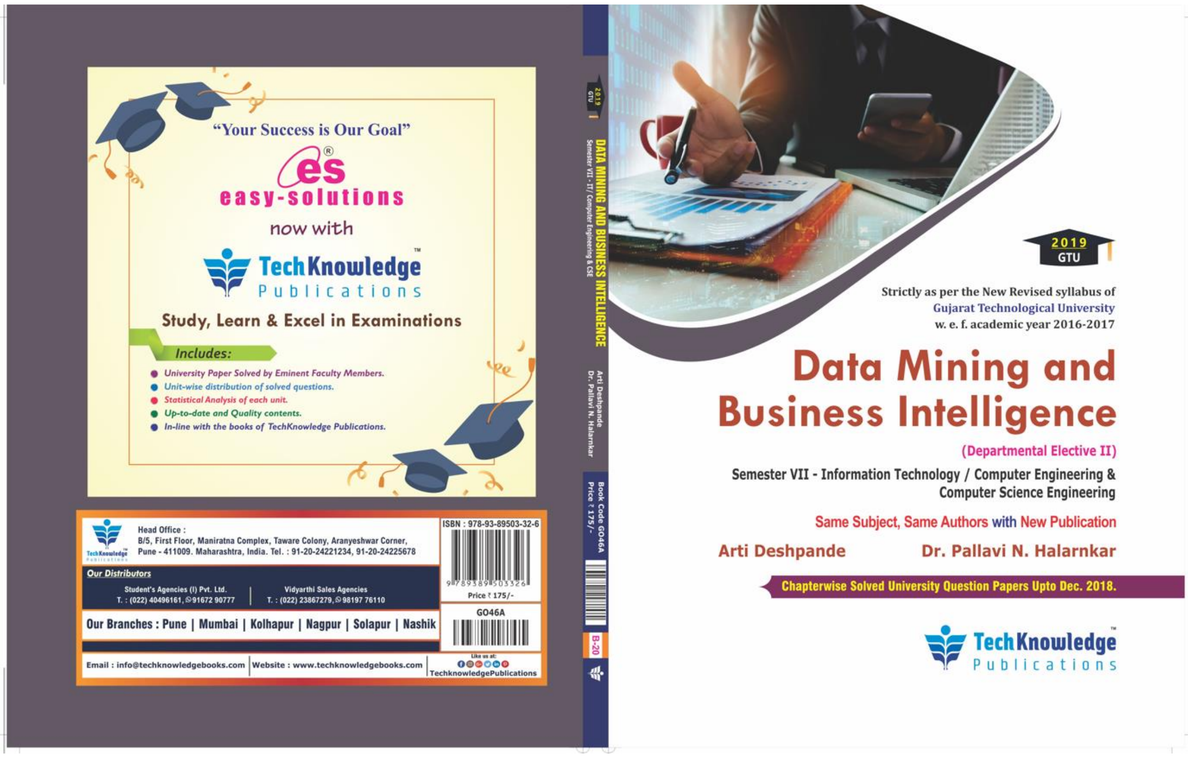
Unit VI : Multiclass Classification (08 Hours)

Multiclass Classification, Semi-Supervised Classification, Reinforcement learning, Systematic Learning, Wholistic learning and multi-perspective learning. Metrics for Evaluating Classifier Performance : Accuracy, Error Rate, precision, Recall, Sensitivity, Specificity; Evaluating the Accuracy of a Classifier : Holdout Method, Random Sub sampling and Cross-Validation. **(Refer chapter 6)**

❑❑❑

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.



TechKnowledge Publications™
Head Office :
B/5, First Floor, Maniratna Complex, Taware Colony, Aranyeshwar Corner,
Pune - 411009. Maharashtra, India. Tel. : 91-20-24221234, 91-20-24225678

Our Distributors
Student's Agencies (I) Pvt. Ltd.
T. : (022) 40496161, ✆ 91672 90777
Vidyarthi Sales Agencies
T. : (022) 23867279, ✆ 98197 76110

Our Branches : Pune | Mumbai | Kolhapur | Nagpur | Solapur | Nashik

ISBN : 978-93-89503-32-6
9 789389 503326
Price ₹ 175/-
GO46A

Email : info@techknowledgebooks.com | Website : www.techknowledgebooks.com | Like us at: TechknowledgePublications

2019 GTU
DATA MINING AND BUSINESS INTELLIGENCE
Semester VII - IT/ Computer Engineering & CSE
Arti Deshpande
Dr. Pallavi N. Halarnkar
Book Code GO46A
Price ₹ 175/-
B-20

2019 GTU

Strictly as per the New Revised syllabus of
Gujarat Technological University
w. e. f. academic year 2016-2017

Data Mining and Business Intelligence

(Departmental Elective II)

Semester VII - Information Technology / Computer Engineering &
Computer Science Engineering

Same Subject, Same Authors with New Publication

Arti Deshpande   Dr. Pallavi N. Halarnkar

Chapterwise Solved University Question Papers Upto Dec. 2018.

TechKnowledge Publications™

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

Data Mining and Business Intelligence

(Code - 2170715)

Semester VII - Information Technology / Computer Engineering & Computer Science Engineering (Departmental Elective-II)
(Gujarat Technological University)

Strictly as per the New Revised Syllabus of Gujarat Technological University w.e.f. academic year 2016-2017

Mrs. Arti Deshpande
ME (Comp. Engg.)
Thadomal Shahani Engineering College , Mumbai.

Dr. Pallavi N. Halarnkar
ME (Comp. Engg.)
Mukesh Patel School of Technology, Management and Engineering, Mumbai.

TechKnowledge™
Publications

GO46A Price ₹ 175/-

(Book Code : GO46A)

Data Mining and Business Intelligence

Mrs. Arti Deshpande, Dr. Pallavi N. Halarnkar

(Semester VII - Information Technology/ Computer Engineering &
Computer Science Engineering : Departmental Elective-II, GTU)

First Printed In India : April 2010 (Mumbai University)
First Edition : July 2016
Second Revised Edition : June 2017
Third Revised Edition : June 2018
Fourth Revised Edition : October 2019 **(TechKnowledge Publications)**

This edition is for sale in India, Bangladesh, Bhutan, Maldives, Nepal, Pakistan, Sri Lanka and designated countries in South-East Asia. Sale and purchase of this book outside of these countries is unauthorized by the publisher.

Printed at : 37/2, Ashtvinayak Industrial Estate, Near Pari Company,
Narhe, Pune, Maharashtra State India,
Pune – 411041

ISBN 978-93-89503-32-6

Published by
TechKnowledge Publications

Head Office : B/5, First floor, Maniratna Complex, Taware Colony, Aranyeshwar Corner,
Pune - 411 009. Maharashtra State, India
Ph : 91-20-24221234, 91-20-24225678.

[2K16GTUCOMP29] (FID : GO46) (Book Code : GO46A)

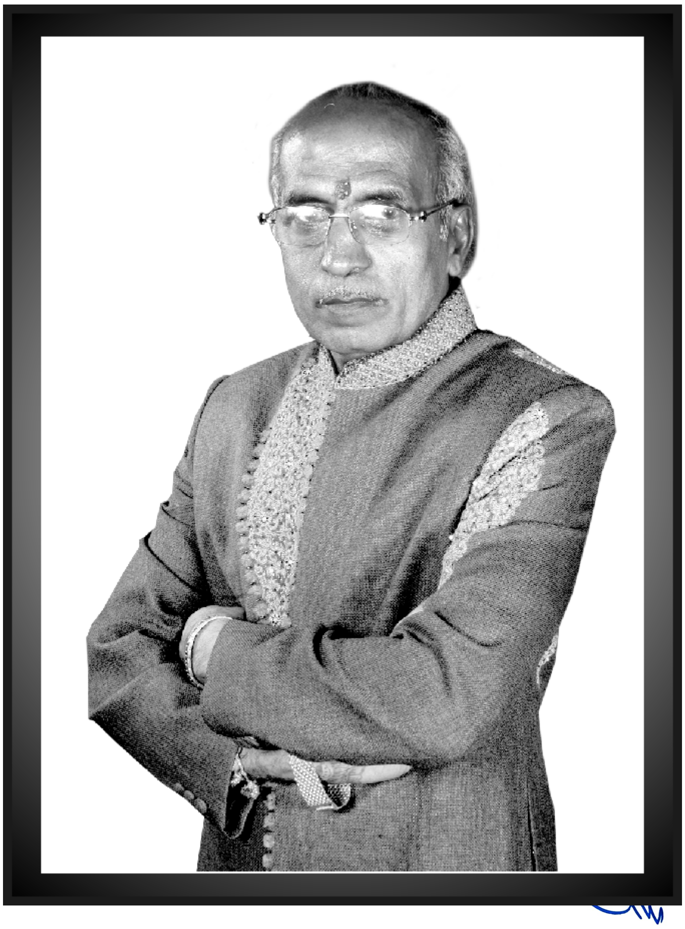
(Book Code : GO46A)

We dedicate this Publication soulfully and wholeheartedly,

in loving memory of our beloved founder director,

Late Shri. Pradeepsheth Lalchandji Lunawat,

who will always be an inspiration, a positive force and strong support behind us.



Lt. Shri. Pradeepji L. Lunawat

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

Soulful Tribute and Gratitude for all Your Sacrifices, Hardwork and 40 years of Strong Vision…

(Book Code : GO46A)

Preface

My Dear Students,

We are extremely happy to come out with this book on **"Data Mining & Business Intelligence"** for you. The topics within the chapters have been arranged in a proper sequence to ensure smooth flow of the subject.

We present this book in the loving memory of **Late Shri. Pradeepji Lunawat,** our source of inspiration and a strong foundation of **"TechKnowledge Publications"**. He will always be remembered in our heart and motivate us to achieve our milestone.

We are thankful to Shri. J. S. Katre, Mr. Shital Bhandari, Shri. Arunoday Kumar and Shri. Chandroday Kumar for the encouragement and support that they have extended. We are also thankful to the staff members of TechKnowledge Publications and others for their efforts to make this book as good as it is. We have jointly made every possible efforts to eliminate all the errors in this book. However if you find any, please let us know, because that will help us to improve further.

We are also thankful to my family members and friends for patience and encouragement.

– Arti Deshpande

– Pallavi N. Halankar

❑❑❑

(Book Code : GO46A)

Syllabus

Unit 1 : Overview and concepts Data Warehousing and Business Intelligence :

Why reporting and Analysing data, Raw data to valuable information-Lifecycle of Data - What is Business Intelligence - BI and DW in today's perspective - What is data warehousing - The building Blocks : Defining Features - Data warehouses and data 1marts - Overview of the components - Metadata in the data warehouse - Need for data warehousing - Basic elements of data warehousing - trends in data warehousing. **(Refer Chapter 1)**

Unit 2 : The Architecture of BI and DW :

BI and DW architectures and its types - Relation between BI and DW - OLAP (Online analytical processing) definitions - Difference between OLAP and OLTP - Dimensional analysis - What are cubes? Drill-down and roll-up - slice and dice or rotation - OLAP models - ROLAP versus MOLAP - defining schemas : Stars, snowflakes and fact constellations. **(Refer Chapter 2)**

Unit 3 : Introduction to data mining (DM) :

Motivation for Data Mining - Data Mining-Definition and Functionalities – Classification of DM Systems - DM task primitives - Integration of a Data Mining system with a Database or a Data Warehouse - Issues in DM – KDD Process. **(Refer Chapter 3)**

Unit 4 : Data Pre-processing :

Why to pre-process data? - Data cleaning: Missing Values, Noisy Data - Data Integration and transformation - Data Reduction : Data cube aggregation, Dimensionality reduction - Data Compression - Numerosity Reduction - Data Mining Primitives - Languages and System Architectures : Task relevant data - Kind of Knowledge to be mined - Discretization and Concept Hierarchy. **(Refer Chapter 4)**

Unit 5 : Concept Description and Association Rule Mining :

What is concept description? - Data Generalization and summarization-based characterization - Attribute relevance - class comparisons Association Rule Mining: Market basket analysis – basic concepts - Finding frequent item sets: Apriori algorithm - generating rules – Improved Apriori algorithm – Incremental ARM – Associative Classification – Rule Mining. **(Refer Chapter 5)**

Unit 6 : Classification and Prediction :

What is classification and prediction? – Issues regarding Classification and prediction :

Classification methods : Decision tree, Bayesian Classification, Rule based, CART, Neural Network

Prediction methods : Linear and nonlinear regression, Logistic Regression

Introduction of tools such as DB Miner /WEKA/DTREG DM Tools. **(Refer Chapter 6)**

(Book Code : GO46A)

Unit 7 : Data Mining for Business Intelligence Applications :

Data mining for business Applications like Balanced Scorecard, Fraud Detection, Click stream Mining, Market Segmentation, retail industry, telecommunications industry, banking & finance and CRM etc.

Data Analytics Life Cycle : Introduction to Big data Business Analytics - State of the practice in analytics role of data scientists

Key roles for successful analytic project : Main phases of life cycle - Developing core deliverables for stakeholders. **(Refer Chapter 7)**

Unit 8 : Advance topics :

Introduction and basic concepts of following topics.

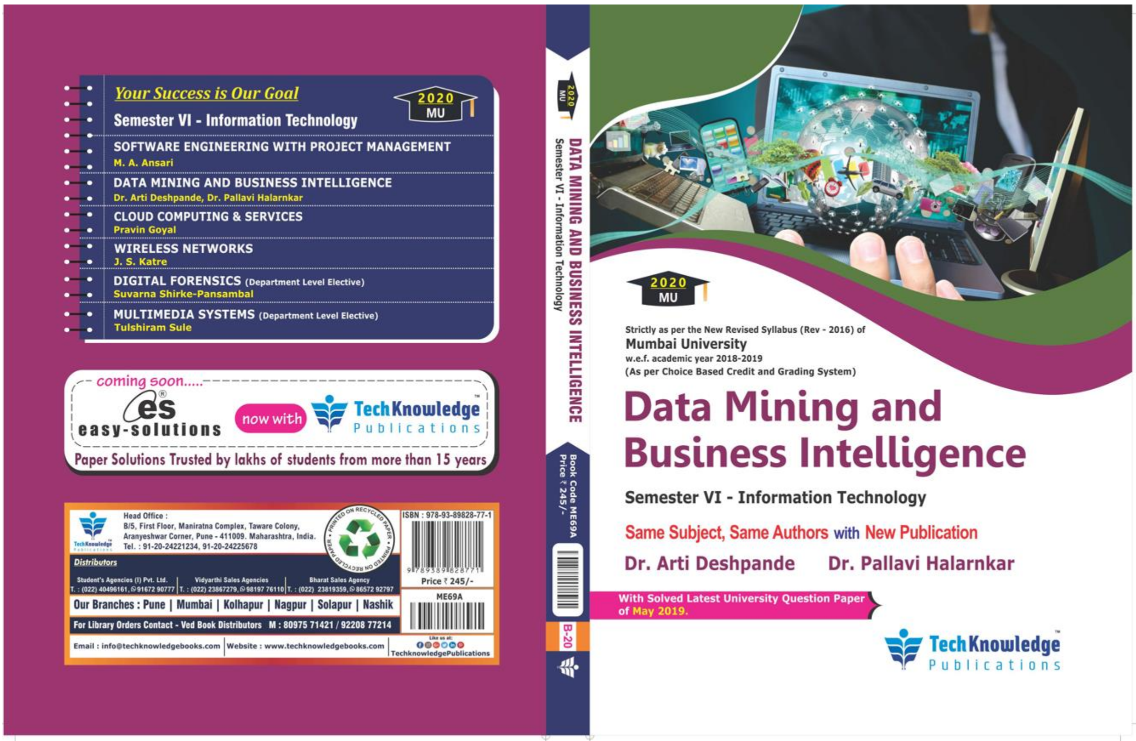
Clustering, Spatial mining, web mining, text mining,

Big Data : Introduction to big data : distributed file system – Big Data and its importance, Four Vs, Drivers for Big data, Big data analytics, Big data applications. Algorithms using map reduce, Matrix-Vector Multiplication by Map Reduce. Introduction to Hadoop architecture: Hadoop Architecture, Hadoop Storage: HDFS, Common Hadoop Shell commands , Anatomy of File Write and Read., NameNode, Secondary NameNode, and DataNode, Hadoop MapReduce paradigm, Map and Reduce tasks, Job, Task trackers – Cluster Setup – SSH & Hadoop Configuration – HDFS Administering – Monitoring & Maintenance. **(Refer Chapter 8)**

□□□

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.


Your Success is Our Goal

2020 MU

Semester VI - Information Technology

SOFTWARE ENGINEERING WITH PROJECT MANAGEMENT
M. A. Ansari

DATA MINING AND BUSINESS INTELLIGENCE
Dr. Arti Deshpande, Dr. Pallavi Halarnkar

CLOUD COMPUTING & SERVICES
Pravin Goyal

WIRELESS NETWORKS
J. S. Katre

DIGITAL FORENSICS (Department Level Elective)
Suvarna Shirke-Pansambal

MULTIMEDIA SYSTEMS (Department Level Elective)
Tulshiram Sule

coming soon.....
es easy-solutions now with TechKnowledge Publications
Paper Solutions Trusted by lakhs of students from more than 15 years

Head Office :
B/5, First Floor, Maniratna Complex, Taware Colony,
Aranyeshwar Corner, Pune - 411009. Maharashtra, India.
Tel. : 91-20-24221234, 91-20-24225678

PRINTED ON RECYCLED PAPER

Distributors

Student's Agencies (I) Pvt. Ltd. T. : (022) 40496161, 91672 90777 | Vidyarthi Sales Agencies T. : (022) 23867279, 98197 76110 | Bharat Sales Agency T. : (022) 23819359, 86572 92797

Our Branches : Pune | Mumbai | Kolhapur | Nagpur | Solapur | Nashik

For Library Orders Contact - Ved Book Distributors M : 80975 71421 / 92208 77214

Email : info@techknowledgebooks.com | Website : www.techknowledgebooks.com | Like us at: TechknowledgePublications

ISBN : 978-93-89828-77-1
9789389828771
Price ₹ 245/-
ME69A

2020 MU
DATA MINING AND BUSINESS INTELLIGENCE
Semester VI - Information Technology
Book Code ME69A
Price ₹ 245/-
B-20

2020 MU

Strictly as per the New Revised Syllabus (Rev - 2016) of
Mumbai University
w.e.f. academic year 2018-2019
(As per Choice Based Credit and Grading System)

Data Mining and Business Intelligence

Semester VI - Information Technology

Same Subject, Same Authors with New Publication

Dr. Arti Deshpande Dr. Pallavi Halarnkar

With Solved Latest University Question Paper of May 2019.

TechKnowledge Publications

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA MUMBAI-50.

Data Mining and Business Intelligence

(Code - ITC602)

Semester VI - Information Technology

(Mumbai University)

Strictly as per the Choice Based Credit and Grading System (Revise 2016) of Mumbai University w.e.f. academic year 2018-2019

Dr. Arti Deshpande

Department of Computer Engineering

Thadomal Shahani Engineering College, Mumbai.

Maharashtra, India.

Dr. Pallavi Halarnkar

Ph.D (Computer Engineering)

TechKnowledge™ Publications

ME69A Price ₹ 245/-

(Book Code : ME69A)

Data Mining and Business Intelligence

Dr. Arti Deshpande, Dr. Pallavi Halarnkar

(Semester VI - Information Technology, Mumbai University)

First Edition : February 2019
Second Revised Edition : February 2020 **(TechKnowledge Publications)**

Printed at : 37/2, Ashtavinayak Industrial Estate,
Near Pari Company, Narhe, Pune, Maharashtra State, India.
Pune – 411041.

ISBN : 978-93-89828-77-1

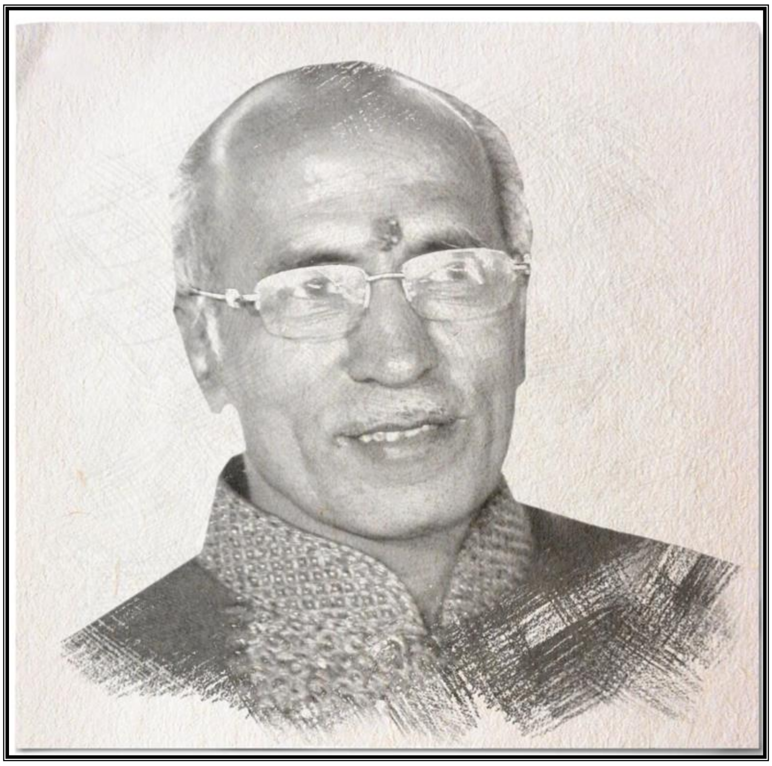
Published by : TechKnowledge Publications

Head Office : B/5, First floor, Maniratna Complex, Taware Colony, Aranyeshwar Corner,
Pune - 411 009. Maharashtra State, India
Ph : 91-20-24221234, 91-20-24225678.
Email : info@techknowledgebooks.com,
Website : www.techknowledgebooks.com

[ITC602] (FID : ME69) (Book Code : ME69A)

(Book Code : ME69A)

We dedicate this Publication soulfully and
wholeheartedly,
in loving memory of our beloved founder director,
Late Shri. Pradeepji Lalchandji Lunawat,
who will always be an inspiration, a positive force and
strong support behind us.



"My work is my prayer to God"

- Lt. Shri. Pradeepji L. Lunawat

Soulful Tribute and Gratitude for all Your
Sacrifices, Hardwork and 40 years of
Strong Vision…

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.
★

Preface

Dear students,

We are extremely happy to present the book on **“Data Mining and Business Intelligence”** for you. We have divided the subject into small chapters so that the topics can be arranged and understood properly. The topics within the chapters have been arranged in a proper sequence to ensure smooth flow of the subject.

We present this book in the loving memory of **Late. Shri. Pradeepji Lunawat**, our source of inspiration and a strong foundation of **“TechKnowledge Publications”.** He will always be remembered in our hearts and motivate us to achieve our new milestone.

We are thankful to Mr. Arunoday Kumar, Mr. Shital Bhandari and Mr. Chandroday Kumar for the encouragement and support that they have extended. We also thankful to the staff members of TechKnowledge Publications for their efforts to make this book as good as it is. We have made every possible efforts to eliminate all the errors in this book. However if you find any, please let us know, because that will help us to improve the book quality further.

We are also thankful to our family members and friends for their patience and encouragement.

- Authors

❑❑❑

(Book Code : ME69A)

Books by Author – Dr Arun Kulkarni

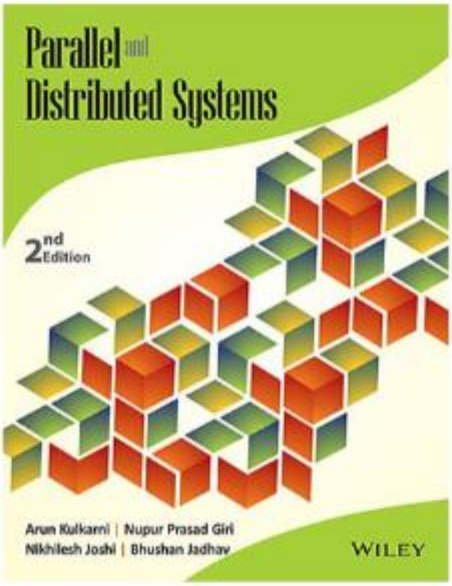
Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE ★
BANDRA
MUMBAI-50.

Parallel and Distributed Systems

by Bhushan Jadhav Arun Kulkarni, Nupur Prasad Giri, Nikhilesh Joshi (Author)

ISBN: **978-8126558674**


Parallel and Distributed Systems
2nd Edition
Arun Kulkarni | Nupur Prasad Giri
Nikhilesh Joshi | Bhushan Jadhav
WILEY

1.7 Classification Based on Characteristic Nature of Processing Elements

1.8 Performance Metrics

1.9 Parallel Programming Models

1.10 Serial and Parallel Algorithms

1.11 Parallelism

Chapter 2 Pipelining

2.1 Introduction

2.2 Pipeline Performance

2.3 Types of Pipeline

2.4 Pipeline Stage Design

2.5 Pipeline Hazards

2.6 Instruction Scheduling

Chapter 3 Synchronous Parallel Processing

3.1 Introduction

3.2 SIMD Architecture and Its Programming Principle

3.3 Single Instruction Multiple Data (SIMD) Parallel Algorithms

3.4 Data Mapping and Memory in Array Processor

3.5 Case Studies of SIMD Parallel Processors

Chapter 4 Introduction to Distributed Systems

4.1 Introduction

4.2 Definition

4.3 Goals of the Distributed System

4.4 Issues Related to the Distributed System

4.5 Types of Distributed System

4.6 Distributed System Models

4.7 Hardware Concept

4.8 Software Concept

4.9 Models of Middleware

4.10 Services Offered by Middleware System

4.11 Client–Server Model

Chapter 5 Communication

5.1 Introduction

5.2 Layered Protocols

5.3 Remote Procedure Call

5.4 Remote Object Invocation

5.5 Remote Method Invocation

5.6 Message-Oriented Communication

5.7 Stream-Oriented Communication

Chapter 6 Resource and Process Management

6.1 Resource Management in Distributed System

6.2 Desirable Features of Global Scheduling Algorithm

6.3 Scheduling in the Distributed System

6.4 Taxonomy of the Distributed Scheduling

6.5 Task Assignment Approach

6.6 Load Balancing Approach

6.7 Issues in Designing Load Balancing Algorithm

6.8 Load Sharing Approach

6.9 Introduction to Process Management

Chapter 7 Synchronization

7.1 Introduction: Clock Synchronization

7.2 Physical Clock

7.3 Logical Clock

7.4 Election Algorithms

7.5 Mutual Exclusion

7.6 Centralized Algorithm

7.7 Distributed Mutual Exclusion

Chapter 8 Replication, Consistency and Distributed File System

8.1 Introduction

8.2 Replication and Consistency

8.3 Replication Management

8.4 Distributed File Systems

8.5 Case Studies

Summary

Multiple Choice Questions

Short Answer Questions

Long Answer Questions

Answers

Practical No. 1

Practical No. 2

Practical No. 3

Practical No. 4

Practical No. 5

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE ★
BANDRA
MUMBAI-50.

Books by Author – Dr Bhushan Jadhav

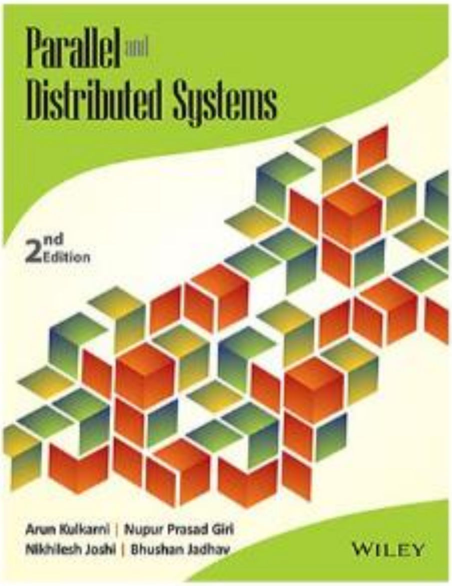
Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE ★
BANDRA
MUMBAI-50.

Parallel and Distributed Systems

by Bhushan Jadhav Arun Kulkarni, Nupur Prasad Giri, Nikhilesh Joshi (Author)

ISBN: **978-8126558674**


Parallel and
Distributed Systems
2nd Edition
Arun Kulkarni | Nupur Prasad Giri
Nikhilesh Joshi | Bhushan Jadhav
WILEY

1.7 Classification Based on Characteristic Nature of Processing Elements

1.8 Performance Metrics

1.9 Parallel Programming Models

1.10 Serial and Parallel Algorithms

1.11 Parallelism

Chapter 2 Pipelining

2.1 Introduction

2.2 Pipeline Performance

2.3 Types of Pipeline

2.4 Pipeline Stage Design

2.5 Pipeline Hazards

2.6 Instruction Scheduling

Chapter 3 Synchronous Parallel Processing

3.1 Introduction

3.2 SIMD Architecture and Its Programming Principle

3.3 Single Instruction Multiple Data (SIMD) Parallel Algorithms

3.4 Data Mapping and Memory in Array Processor

3.5 Case Studies of SIMD Parallel Processors

Chapter 4 Introduction to Distributed Systems

4.1 Introduction

4.2 Definition

4.3 Goals of the Distributed System

4.4 Issues Related to the Distributed System

4.5 Types of Distributed System

4.6 Distributed System Models

4.7 Hardware Concept

4.8 Software Concept

4.9 Models of Middleware

4.10 Services Offered by Middleware System

4.11 Client–Server Model

Chapter 5 Communication

5.1 Introduction

5.2 Layered Protocols

5.3 Remote Procedure Call

5.4 Remote Object Invocation

5.5 Remote Method Invocation

5.6 Message-Oriented Communication

5.7 Stream-Oriented Communication

Chapter 6 Resource and Process Management

6.1 Resource Management in Distributed System

6.2 Desirable Features of Global Scheduling Algorithm

6.3 Scheduling in the Distributed System

6.4 Taxonomy of the Distributed Scheduling

6.5 Task Assignment Approach

6.6 Load Balancing Approach

6.7 Issues in Designing Load Balancing Algorithm

6.8 Load Sharing Approach

6.9 Introduction to Process Management

Chapter 7 Synchronization

7.1 Introduction: Clock Synchronization

7.2 Physical Clock

7.3 Logical Clock

7.4 Election Algorithms

7.5 Mutual Exclusion

7.6 Centralized Algorithm

7.7 Distributed Mutual Exclusion

Chapter 8 Replication, Consistency and Distributed File System

8.1 Introduction

8.2 Replication and Consistency

8.3 Replication Management

8.4 Distributed File Systems

8.5 Case Studies

Summary

Multiple Choice Questions

Short Answer Questions

Long Answer Questions

Answers

Practical No. 1

Practical No. 2

Practical No. 3

Practical No. 4

Practical No. 5

Practical No. 6

Practical No. 7

Practical No. 8

Practical No. 9

Practical No. 10

Practical No. 11

Practical No. 12

Index

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

DISTRIBUTED COMPUTING

Includes Labs

Nupur Giri

Lata Ragha

Bhushan Jadhav

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA MUMBAI-50.

STAREDU SOLUTIONS

www.staredusolutions.org

DISTRIBUTED COMPUTING

Nupur Prasad Giri

Lata Ragha

Bhushan Jadhav

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

STAREDU SOLUTIONS

Author Profile

Dr. Nupur Prasad Giri is Professor and Head of Department, Computer Engineering at Vivekanand Education Society Institute of Technology, Mumbai. She has more than 25 years of industry and teaching experiences. Her doctoral research was in the multidisciplinary field of Multi Agent System and Mobile Computing, typically Cellular Networks. Her contributions in the field of Artificial Intelligence, Mobile and Distributed Computing have been published in many journals and International conferences. She has filed many patents.

She is also a committee member of many international conferences. She is recognized PhD guide in University of Mumbai. She has represented India as an Expert in Worldskills-2011, London as well as in Worldskills-2013, Germany in the category of Web Design. She has also been awarded Microsoft's "AI for Earth" Grant in the year 2018.

Dr. Lata L. Ragha is Professor and Head, Department of Computer Engineering at Fr. C. Rodrigues Institute of Technology, Vashi, Navi-Mumbai. She has received her Ph. D. degree from Jadavpur University, Kolkata in 2011. Her research interests include in the areas of Networking, Security, Internet Routing, and Data Mining. She has more than 100 research publications in International Journals and conferences. She is having 32 years of teaching experience. She is a member of Board of Studies, Mumbai University. She is associated with two colleges as member of Department Advisory Board. Two members have received their PhD degree under her guidance and two more have submitted their thesis under her guidance.

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

Dr. Bhushan A Jadhav is an Assistant Professor at Thadomal Sahani Engineering College, Bandra, Mumbai. He has more than 11 years of teaching experience and has completed his Ph.D in area of Cloud Computing &Big Data Analytics. His area of interest and research includes cutting edge technologies such as DevOps, Internet of Things (IOT), Cloud Computing, Big Data Analytics, Python Programming, R Programming and Advance Information Security. He has published more than 8 research papers in National and International journals and conferences He has published six books with various publishers. He is certified trainer for Star Certification's Python Programming, Cloud Computing and DevOps Courses.

Contents

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

Information Management

Asha Bharambe
Bhushan Jadhav
Anjali Yeole

TECHNICAL PUBLICATIONS
An Up-Thrust for Knowledge

WILEY

Information Management

Asha Bharambe
Assistant Professor
Vivekanand Education Society's Institute of Technology
Chembur (E), Mumbai

Bhushan Jadhav
Assistant Professor
Thadomal Shahani Engineering College
Bandra (W), Mumbai

Anjali Yeole
Assistant Professor
Vivekanand Education Society's Institute of Technology
Chembur (E), Mumbai

WILEY

*This book is dedicated to my parents Mr. Yadeo Barhate and Mrs. Usha Barhate,
my in-laws Mr. Laxman Bharambe and Mrs. Mangala Bharambe,
my husband Mr. Aniket Bharambe and my son Atharva.*

—Asha Bharambe

*This book is dedicated to my parents Mr. Ashok Jadhav and Mrs. Shanta Jadhav,
my wife Sonali Jadhav and my son Atharva.*

—Bhushan Jadhav

*This book is dedicated to my parents Mr. Devidas Pawar and Mrs. Alka Pawar,
my in-laws Mr. Madhukar Yeole and Late Kusum Yeole,
my husband Mr. Shrikant Yeole and my daughters Shrutika and Sanvee.*

—Anjali Yeole

About the Authors

Asha Bharambe is currently working as Assistant Professor at Vivekanand Education Society's Institute of Technology, Chembur (E), Mumbai. She is also pursuing PhD in Information Technology from Mumbai University. She has received her undergraduate degree in Computer Science from SNDT Women's University and postgraduate in Computer Engineering from Mumbai University. Prof. Bharambe has taught both graduate and undergraduate students and has a teaching experience of more than 15 years. She has published many research papers in national and international conferences on Data Mining, Big Data and Artificial Intelligence. Her areas of interest are Data Mining, Big Data and Natural Language Processing.

Bhushan Jadhav is currently working as Assistant Professor at Thadomal Shahani Engineering College, Bandra (W), Mumbai. He is also pursuing PhD in Computer Science Engineering from Thadomal Shahani Engineering College, Mumbai University. He received his undergraduate and postgraduate education from Mumbai University. Prof. Jadhav has taught both postgraduate and undergraduate students and has teaching experience of more than 8 years. He has published many research papers in national and international conferences on Virtualization and Cloud Computing. His book *Parallel and Distributed Systems* has received very good response at Mumbai University. His areas of interest are Database Management, Parallel Computing, Distributed Computing, Virtualization and Cloud Computing.

Anjali Yeole is currently working as Assistant Professor at Vivekanand Education Society's Institute of Technology, Chembur (E), Mumbai. She is also pursuing PhD in Computer Science Engineering from Thadomal Shahani Engineering College, Mumbai University. She received her undergraduate in Computer Science from SNDT University and postgraduate education in Computer Engineering from Veermata Jijabai Technological Institute, Mumbai University. Prof. Yeole has taught both postgraduate and undergraduate students and has teaching experience of more than 12 years. She has published many research papers in national and international conferences on Security, Cloud Computing and Web Applications. Her areas of interest are Parallel Computing, Distributed Computing, Virtualization and Cloud Computing, System Security, Computer Network, and Web Technology.

Contents

Chapter 2 Introduction to Big Data

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50.

Chapter 4 Information Governance 131

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai-400 050.

INTERNET OF EVERYTHING

Includes Labs and Cases

INTERNET OF EVERYTHING

Bhushan Jadhav

Anjali Yeole

Gopal Pardesi

Vaishali Khairnar

Dhananjay Kalbande

STAREDU SOLUTIONS

Dr. Bhushan Jadhav is an Assistant Professor at Thadomal Shahani Engineering College, Bandra, Mumbai. He has more than 11 years of teaching experience and has completed his Ph.D in area of Cloud Computing &Big Data Analytics. His area of interest and research includes cutting edge technologies such as DevOps, Internet of Things (IOT), Cloud Computing, Big Data Analytics, Python Programming, R Programming and Advance Information Security. He has published more than 8 research papers in National and International journals and conferences He has published six books with various publishers. He is certified trainer for Star Certification's Python Programming, Cloud Computing and DevOps Courses.

Dr. Anjali Yeole is currently working as Assistant Professor at Vivekanand Education Society's Institute of Technology, Chembur (E), Mumbai. She has received her PhD in Computer Science Engineering from Thadomal Shahani Engineering College under Mumbai University, India. Her research area is "Internet of Things". She received her undergraduate in Computer science from SNDT university and postgraduate education in Computer engineering from VJTI, Mumbai University. She has taught both Masters and undergraduate students and has teaching experience of more than 17 years. She has published many research papers in national and international conferences and journals on IoT, security, cloud computing and web applications. She has authored a book "Information management" Her areas of interest are IoT, Parallel Computing, Distributed Computing, Virtualization and Cloud Computing, System Security, Computer Network and Web Technology.

Dr. Gopal Pardesi is currently working as Associate Professor at Thadomal Shahani Engineering College, Bandra, Mumbai. He has more than 28 years of teaching experience. He has completed his Ph.D in the field of Information and Communication Technologies. His areas of interest are Wireless Networks, Project Management, Microprocessors and Microcontrollers, Embedded Systems and Machine learning.

Dr. Vaishali Khairnar is currently associated as Head of Department Information Technology Department at Terna College of Engineering, Nerul, Navi Mumbai . She has done her Doctoral work from Nirma University, Ahmedabad, Gujrat on Vehicular Ad-Hoc networks. She has more than 19 years of collective experience in industry, teaching as well as research. She has more than 50 papers published in national and international journals and conferences to her credit. Her area of interest includes wireless networks, IOT, Open source tools, Ad-Hoc Sensor networks, Internet Programming, Mobile Application and Development & Storage Network Management and Retrieval. She has authored four books published in India, UK and Singapore. She has been a resource person for several workshops and short term training programmes. She is paper reviewer on different editor bodies at national/ international journals and conferences. She is active member of CSI- Mumbai Chapter and ACM.

Dr. Dhananjay Kalbande is currently a Professor in Computer Engineering and Dean(Industry Relations), Sardar Patel Institute of Technology, Andheri (West),Mumbai, India. He was Head of the Department from April 2012 to Oct 2019. He completed his B.E. in Computer Technology from Nagpur University in 1997 and Master of Engineering in Information Technology in May 2005, from Vivekanand Education Society's Institute of Technology (VESIT), Mumbai University, Mumbai, India. He has obtained a Ph.D in Technology from University of Mumbai, Mumbai in 2011. He has been awarded a Post-Doctorate (PDF)from Tata Institute of Social Sciences (T.I.S.S.) in 2016. He has also been honoured as a Senior Research Fellow (SRF) on the NCW-TISS Project, funded by National Commission for Women, Govt. of India at T.I.S.S, from July 23,2016 to Oct 10, 2016. He was a Research Fellow on the CliX Project at T.I.S.S., funded by Tata Trust and M.I.T.(U.S.A.) from Feb 20,2017 to May 19,2017. He has over 19+ Years of experience in teaching & research. His research interests include Soft computing (Neural Networks, Fuzzy Logic) Computer Network, Human Machine Interaction Decision making and business Intelligence. Mobile application development for social cause. ICT for semi-rural development for social cause. He has authored four books namely Graphical User Interface (Pareen Publications), MIS (Pareen Publications), Human Machine Interaction and Digital Forensic with Wiley India Pvt. Ltd. He has delivered and conducted workshops, seminars, Tutorials and Expert Talks on NS2 Neural Network, VB.Net and ADO.Net, Transfer Learning. Skinzy is his brainchil which has been turned into reality into Healthcare Start-up operated from Mumba Skinzy's flagship product "DermaPhoto" is an AI based mobile application that ca detect skin diseases. He has patented 6 innovative ideas of research work.

Contents

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

CLOUD COMPUTING AND SERVICES

Bhushan A. Jadhav

Dr. Deven Shah

Arup Vithal

STAREDU SOLUTIONS

www.staredusolutions.org

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

CLOUD COMPUTING AND SERVICES

STAREDU
SOLUTIONS

Bhushan A. Jadhav | Dr. Deven Shah | Arup Vithal

Contents

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE
BANDRA
MUMBAI-50

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai-400 050.

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

Books by Author – Shilpa Ingoley

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

MU

Sem 4

Information Technology

NEW SYLLABUS 2020-21

Computer Network & Network Design

- Manoj S. Kavedia

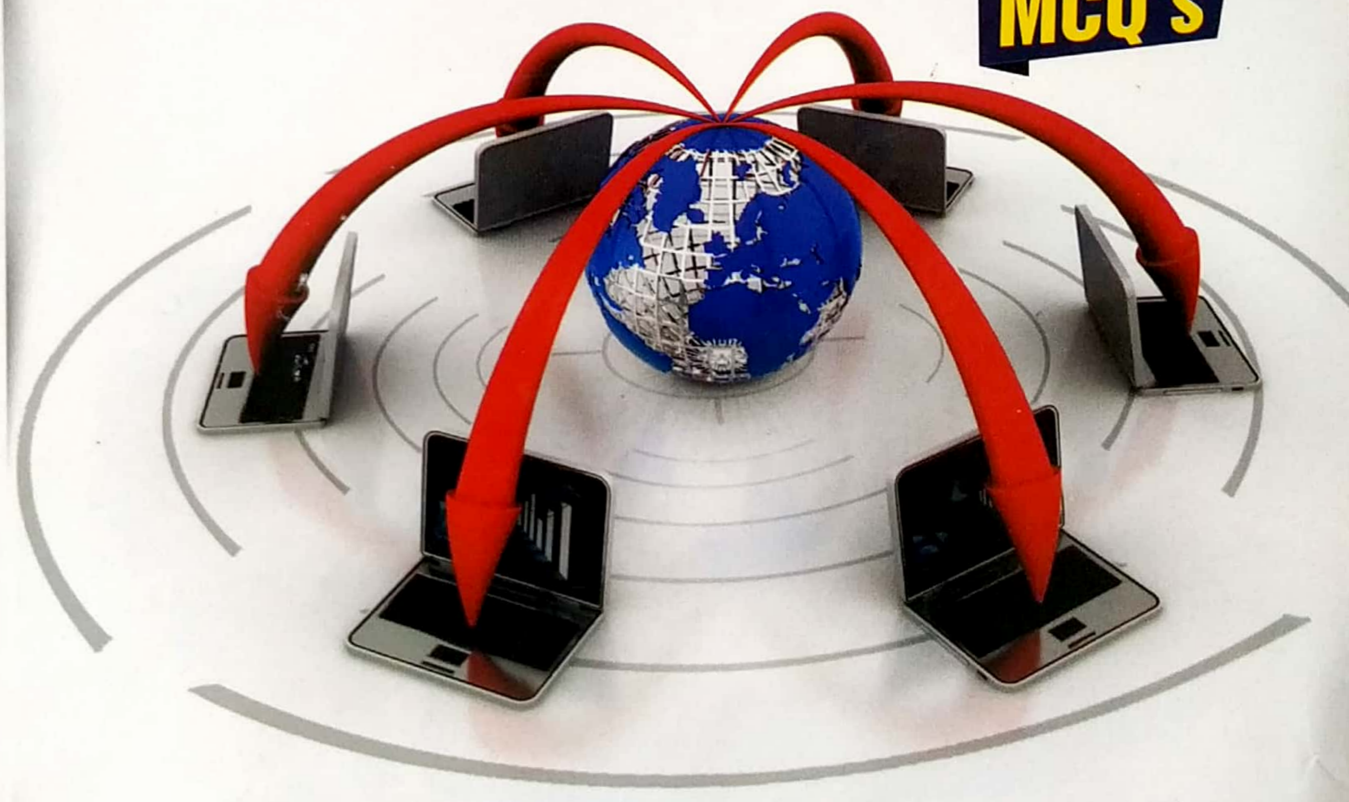
- Shilpa Ingoley

With All Latest Solved University Q. Papers

with MCQ's

2021 EDITION

Course Code (ITC402) (Compulsory Subject)



ISBN 978-93-90376-29-2

M4-27A

Price ₹ 365/-

TECH-NEO PUBLICATIONS

Where Authors Inspire Innovation

A Sachin Shah Venture

About the Authors

Er. Manoj S. Kavedia
Assistant Professor,
Electronics and Telecommunication Department.
Thadomal Shahani Engineering College, Bandra, Mumbai
(Total 24 years of Teaching Experience)

Shilpa Ingoley
Assistant Professor,
M.E. Computer Engineering (Mumbai University) Computer Department
Thadomal Shahani Engineering College, Bandra, Mumbai
(Total 21 years of Teaching Experience)

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE BANDRA MUMBAI-50.

MU Sem 4 Information Technology

Name of Subject	Author
Applied Mathematics-IV	Baphana R M (Adjunct Faculty, COEP, Pune)
Computer Network and Network Design	Manoj S. Kavedia, Shilpa Ingoley
Operating System	Manoj S. Kavedia
Automata Theory	Ashish Budhrani
Computer Organization and Architecture	Velankar Shrikant, Shah Urvashi

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

THADOMAL SHAHANI ENGINEERING COLLEGE BANDRA MUMBAI-50. ★

University of Mumbai

Computer Network and Network Design

(Course Code : ITC402)

Semester IV - Information Technology

Strictly as per the New Syllabus (REV-2019 'C' Scheme) of Mumbai University w.e.f. academic year 2020-2021

Er. Manoj S. Kavedia
Assistant Professor
Department of Electronics and Telecommunication
Thadomal Shahani Engineering College (TSEC), Bandra, Mumbai

Shilpa Ingoley
M.E. Computer Engineering (Mumbai University)
Assistant professor,
Computer Department
Thadomal Shahani Engineering College-Bandra

TECH-NEO PUBLICATIONS
Where Authors Inspire Innovation
A Sachin Shah Venture

M4-27

Index

MU Sem 5 Computer Engineering

New Syllabus 2021-22

Compulsory Subject (CSC503)

Computer Network

Manoj S. Kavedia Shilpa Ingoley

(Thadomal Shahani Engineering College, Bandra, Mumbai)

With Typical MCQ's

TECH-NEO PUBLICATIONS

Where Authors Inspire Innovation

A Sachin Shah Venture

www.techneobooks.in

info@techneobooks.in

- With Solved Latest UNIVERSITY QUESTION PAPERS.
- Self Explanatory Diagrams.
- Comparisons in Tabular Form.
- Multiple Choice Questions.

M5-55A

Price ₹ 325/-

About the Authors

Er. Manoj S. Kavedia

Assistant Professor,

Electronics and Telecommunication Department.

Thadomal Shahani Engineering College, Bandra, Mumbai

(Total 24 years of Teaching Experience)

Shilpa Ingoley

Assistant Professor,

M.E. Computer Engineering (Mumbai University) Computer Department

Thadomal Shahani Engineering College, Bandra, Mumbai

(Total 21 years of Teaching Experience)

Sem 5 Computer Engineering

Course Code	Compulsory Subjects
CSC501	Theoretical Computer Science
CSC502	Software Engineering
CSC503	Computer Network
CSC504	Data warehousing and Mining

Course Code	Departmental Elective
CSDO5012	Internet Programming
CSDO5013	Advance Database Management System

ISBN
978-93-90904-11-2
9 789390 904112

For Orders Contact :

Krishna Book Collections

Ground Floor, Krishna Niwas Building, Behind BEST Niwas Building, Near Napoo Hall,
Chandavarkar Road, Matunga East, Mumbai 400019. **E-mail :** dharmeshsota05@gmail.com
Mobile No.: Dharmesh Sota - 98207 41455, Tulsidas Sota - 98331 33921 / 98330 82745 /
98330 82761

Books are available on Flipkart amazon bookwalas.com the online book shop

University of Mumbai

Computer Network

(Code : CSC503)

Semester V - Computer Engineering

Strictly as per the Choice Based Credit and Grading System (Revise 2019) of Mumbai University w.e.f. academic year 2021-2022

Er. Manoj S. Kavedia
Assistant Professor,
Department of Electronics and Telecommunication
Thadomal Shahani Engineering College (TSEC), Bandra, Mumbai
ME Electronics
Pursing PHD in Internet of Things

Shilpa Ingoley
M.E. Computer Engineering (Mumbai University)
Assistant professor,
Computer Department
Thadomal Shahani Engineering College-Bandra

Dr. G. T. Thampi
PRINCIPAL
Thadomal Shahani Engineering College
Bandra (W), Mumbai - 400 050.

TECH-NEO™
PUBLICATIONS
Where Authors Inspire Innovation
A Sachin Shah Venture

M5-55

Index

□□□