

Data Clustering Charu Aggarwal

THANK YOU FOR READING **DATA CLUSTERING CHARU AGGARWAL** . AS YOU MAY KNOW, PEOPLE HAVE SEARCH HUNDREDS TIMES FOR THEIR FAVORITE NOVELS LIKE THIS DATA CLUSTERING CHARU AGGARWAL , BUT END UP IN MALICIOUS DOWNLOADS.

RATHER THAN READING A GOOD BOOK WITH A CUP OF TEA IN THE AFTERNOON, INSTEAD THEY ARE FACING WITH SOME HARMFUL VIRUS INSIDE THEIR LAPTOP.

DATA CLUSTERING CHARU AGGARWAL IS AVAILABLE IN OUR BOOK COLLECTION AN ONLINE ACCESS TO IT IS SET AS PUBLIC SO YOU CAN GET IT INSTANTLY. OUR BOOK SERVERS SPANS IN MULTIPLE LOCATIONS, ALLOWING YOU TO GET THE MOST LESS LATENCY TIME TO DOWNLOAD ANY OF OUR BOOKS LIKE THIS ONE. MERELY SAID, THE DATA CLUSTERING CHARU AGGARWAL IS UNIVERSALLY COMPATIBLE WITH ANY DEVICES TO READ

MINING HETEROGENEOUS INFORMATION NETWORKS - YIZHOU SUN 2012-08-15

REAL WORLD PHYSICAL AND ABSTRACT DATA OBJECTS ARE INTERCONNECTED, FORMING GIGANTIC, INTERCONNECTED NETWORKS. BY STRUCTURING THESE DATA OBJECTS AND INTERACTIONS BETWEEN THESE OBJECTS INTO MULTIPLE TYPES, SUCH NETWORKS BECOME SEMI-STRUCTURED HETEROGENEOUS INFORMATION NETWORKS. MOST REAL WORLD APPLICATIONS THAT HANDLE BIG DATA, INCLUDING INTERCONNECTED SOCIAL MEDIA AND SOCIAL NETWORKS, SCIENTIFIC, ENGINEERING, OR MEDICAL INFORMATION SYSTEMS, ONLINE E-COMMERCE SYSTEMS, AND MOST DATABASE SYSTEMS, CAN BE STRUCTURED INTO HETEROGENEOUS INFORMATION NETWORKS. THEREFORE, EFFECTIVE ANALYSIS OF LARGE-SCALE HETEROGENEOUS INFORMATION NETWORKS POSES AN INTERESTING BUT CRITICAL CHALLENGE. IN THIS MONOGRAPH, WE INVESTIGATE THE PRINCIPLES AND METHODOLOGIES OF MINING HETEROGENEOUS INFORMATION NETWORKS. DEPARTING FROM MANY EXISTING NETWORK MODELS THAT VIEW DATA AS HOMOGENEOUS GRAPHS OR NETWORKS, OUR SEMI-STRUCTURED HETEROGENEOUS INFORMATION NETWORK MODEL LEVERAGES THE RICH SEMANTICS OF TYPED NODES AND LINKS IN A NETWORK AND UNCOVERS SURPRISINGLY RICH KNOWLEDGE FROM INTERCONNECTED DATA. THIS SEMI-STRUCTURED HETEROGENEOUS NETWORK MODELING LEADS TO A SERIES OF NEW PRINCIPLES AND POWERFUL METHODOLOGIES FOR MINING INTERCONNECTED DATA, INCLUDING (1) RANK-BASED CLUSTERING AND CLASSIFICATION, (2) META-PATH-BASED SIMILARITY SEARCH AND MINING, (3) RELATION STRENGTH-AWARE MINING, AND MANY OTHER POTENTIAL DEVELOPMENTS. THIS MONOGRAPH INTRODUCES THIS NEW RESEARCH FRONTIER AND POINTS OUT SOME PROMISING RESEARCH DIRECTIONS.

MANAGING AND MINING GRAPH DATA - CHARU C. AGGARWAL 2010-02-02

MANAGING AND MINING GRAPH DATA IS A COMPREHENSIVE SURVEY BOOK IN GRAPH MANAGEMENT AND MINING. IT CONTAINS EXTENSIVE SURVEYS ON A VARIETY OF IMPORTANT GRAPH TOPICS SUCH AS GRAPH LANGUAGES, INDEXING, CLUSTERING, DATA GENERATION, PATTERN MINING, CLASSIFICATION, KEYWORD SEARCH, PATTERN MATCHING, AND PRIVACY. IT

ALSO STUDIES A NUMBER OF DOMAIN-SPECIFIC SCENARIOS SUCH AS STREAM MINING, WEB GRAPHS, SOCIAL NETWORKS, CHEMICAL AND BIOLOGICAL DATA. THE CHAPTERS ARE WRITTEN BY WELL KNOWN RESEARCHERS IN THE FIELD, AND PROVIDE A BROAD PERSPECTIVE OF THE AREA. THIS IS THE FIRST COMPREHENSIVE SURVEY BOOK IN THE EMERGING TOPIC OF GRAPH DATA PROCESSING. MANAGING AND MINING GRAPH DATA IS DESIGNED FOR A VARIED AUDIENCE COMPOSED OF PROFESSORS, RESEARCHERS AND PRACTITIONERS IN INDUSTRY. THIS VOLUME IS ALSO SUITABLE AS A REFERENCE BOOK FOR ADVANCED-LEVEL DATABASE STUDENTS IN COMPUTER SCIENCE AND ENGINEERING.

ON UNCERTAIN GRAPHS - ARIJIT KHAN 2022-05-31

LARGE-SCALE, HIGHLY INTERCONNECTED NETWORKS, WHICH ARE OFTEN MODELED AS GRAPHS, PERVADE BOTH OUR SOCIETY AND THE NATURAL WORLD AROUND US. UNCERTAINTY, ON THE OTHER HAND, IS INHERENT IN THE UNDERLYING DATA DUE TO A VARIETY OF REASONS, SUCH AS NOISY MEASUREMENTS, LACK OF PRECISE INFORMATION NEEDS, INFERENCE AND PREDICTION MODELS, OR EXPLICIT MANIPULATION, E.G., FOR PRIVACY PURPOSES. THEREFORE, UNCERTAIN, OR PROBABILISTIC, GRAPHS ARE INCREASINGLY USED TO REPRESENT NOISY LINKED DATA IN MANY EMERGING APPLICATION SCENARIOS, AND THEY HAVE RECENTLY BECOME A HOT TOPIC IN THE DATABASE AND DATA MINING COMMUNITIES. MANY CLASSICAL ALGORITHMS SUCH AS REACHABILITY AND SHORTEST PATH QUERIES BECOME #P-COMPLETE AND, THUS, MORE EXPENSIVE OVER UNCERTAIN GRAPHS. MOREOVER, VARIOUS COMPLEX QUERIES AND ANALYTICS ARE ALSO EMERGING OVER UNCERTAIN NETWORKS, SUCH AS PATTERN MATCHING, INFORMATION DIFFUSION, AND INFLUENCE MAXIMIZATION QUERIES. IN THIS BOOK, WE DISCUSS THE SOURCES OF UNCERTAIN GRAPHS AND THEIR APPLICATIONS, UNCERTAINTY MODELING, AS WELL AS THE COMPLEXITIES AND ALGORITHMIC ADVANCES ON UNCERTAIN GRAPHS PROCESSING IN THE CONTEXT OF BOTH CLASSICAL AND EMERGING GRAPH QUERIES AND ANALYTICS. WE EMPHASIZE THE CURRENT CHALLENGES AND HIGHLIGHT SOME FUTURE RESEARCH DIRECTIONS.

CONSTRAINED CLUSTERING - SUGATO BASU 2008-08-18

SINCE THE INITIAL WORK ON CONSTRAINED CLUSTERING, THERE HAVE BEEN NUMEROUS ADVANCES IN METHODS, APPLICATIONS, AND OUR UNDERSTANDING OF THE THEORETICAL PROPERTIES OF CONSTRAINTS AND CONSTRAINED CLUSTERING ALGORITHMS. BRINGING THESE DEVELOPMENTS TOGETHER, **CONSTRAINED CLUSTERING: ADVANCES IN ALGORITHMS, THEORY, AND APPLICATIONS** PRESENTS AN EXTENSIVE COLLECTION OF THE LATEST INNOVATIONS IN CLUSTERING DATA ANALYSIS METHODS THAT USE BACKGROUND KNOWLEDGE ENCODED AS CONSTRAINTS. **ALGORITHMS** THE FIRST FIVE CHAPTERS OF THIS VOLUME INVESTIGATE ADVANCES IN THE USE OF INSTANCE-LEVEL, PAIRWISE CONSTRAINTS FOR PARTITIONAL AND HIERARCHICAL CLUSTERING. **THE BOOK THEN EXPLORES OTHER TYPES OF CONSTRAINTS FOR CLUSTERING, INCLUDING CLUSTER SIZE BALANCING, MINIMUM CLUSTER SIZE, AND CLUSTER-LEVEL RELATIONAL CONSTRAINTS. THEORY** IT ALSO DESCRIBES VARIATIONS OF THE TRADITIONAL CLUSTERING UNDER CONSTRAINTS PROBLEM AS WELL AS APPROXIMATION ALGORITHMS WITH HELPFUL PERFORMANCE GUARANTEES. **APPLICATIONS** THE BOOK ENDS BY APPLYING CLUSTERING WITH CONSTRAINTS TO RELATIONAL DATA, PRIVACY-PRESERVING DATA PUBLISHING, AND VIDEO SURVEILLANCE DATA. IT DISCUSSES AN INTERACTIVE VISUAL CLUSTERING APPROACH, A DISTANCE METRIC LEARNING APPROACH, EXISTENTIAL CONSTRAINTS, AND AUTOMATICALLY GENERATED CONSTRAINTS. WITH CONTRIBUTIONS FROM INDUSTRIAL RESEARCHERS AND LEADING ACADEMIC EXPERTS WHO PIONEERED THE FIELD, THIS VOLUME DELIVERS THOROUGH COVERAGE OF THE CAPABILITIES AND LIMITATIONS OF CONSTRAINED CLUSTERING METHODS AS WELL AS INTRODUCES NEW TYPES OF CONSTRAINTS AND CLUSTERING ALGORITHMS.

DATA MINING - CHARU C. AGGARWAL 2016-10-09

THIS TEXTBOOK EXPLORES THE DIFFERENT ASPECTS OF DATA MINING FROM THE FUNDAMENTALS TO THE COMPLEX DATA TYPES AND THEIR APPLICATIONS, CAPTURING THE WIDE DIVERSITY OF PROBLEM DOMAINS FOR DATA MINING ISSUES. IT GOES BEYOND THE TRADITIONAL FOCUS ON DATA MINING PROBLEMS TO INTRODUCE ADVANCED DATA TYPES SUCH AS TEXT, TIME SERIES, DISCRETE SEQUENCES, SPATIAL DATA, GRAPH DATA, AND SOCIAL NETWORKS. UNTIL NOW, NO SINGLE BOOK HAS ADDRESSED ALL THESE TOPICS IN A COMPREHENSIVE AND INTEGRATED WAY. THE CHAPTERS OF THIS BOOK FALL INTO ONE OF THREE CATEGORIES: **FUNDAMENTAL CHAPTERS:** DATA MINING HAS FOUR MAIN PROBLEMS, WHICH CORRESPOND TO CLUSTERING, CLASSIFICATION, ASSOCIATION PATTERN MINING, AND OUTLIER ANALYSIS. THESE CHAPTERS COMPREHENSIVELY DISCUSS A WIDE VARIETY OF METHODS FOR THESE PROBLEMS. **DOMAIN CHAPTERS:** THESE CHAPTERS DISCUSS THE SPECIFIC METHODS USED FOR DIFFERENT DOMAINS OF DATA SUCH AS TEXT DATA, TIME-SERIES DATA, SEQUENCE DATA, GRAPH DATA, AND SPATIAL DATA. **APPLICATION CHAPTERS:** THESE CHAPTERS STUDY IMPORTANT APPLICATIONS SUCH AS STREAM MINING, WEB MINING, RANKING, RECOMMENDATIONS, SOCIAL NETWORKS, AND PRIVACY PRESERVATION. THE DOMAIN CHAPTERS ALSO HAVE AN APPLIED FLAVOR. APPROPRIATE FOR BOTH INTRODUCTORY AND ADVANCED DATA MINING COURSES, **DATA MINING: THE TEXTBOOK** BALANCES MATHEMATICAL

DETAILS AND INTUITION. IT CONTAINS THE NECESSARY MATHEMATICAL DETAILS FOR PROFESSORS AND RESEARCHERS, BUT IT IS PRESENTED IN A SIMPLE AND INTUITIVE STYLE TO IMPROVE ACCESSIBILITY FOR STUDENTS AND INDUSTRIAL PRACTITIONERS (INCLUDING THOSE WITH A LIMITED MATHEMATICAL BACKGROUND). NUMEROUS ILLUSTRATIONS, EXAMPLES, AND EXERCISES ARE INCLUDED, WITH AN EMPHASIS ON SEMANTICALLY INTERPRETABLE EXAMPLES. **PRAISE FOR DATA MINING: THE TEXTBOOK** - "AS I READ THROUGH THIS BOOK, I HAVE ALREADY DECIDED TO USE IT IN MY CLASSES. THIS IS A BOOK WRITTEN BY AN OUTSTANDING RESEARCHER WHO HAS MADE FUNDAMENTAL CONTRIBUTIONS TO DATA MINING, IN A WAY THAT IS BOTH ACCESSIBLE AND UP TO DATE. THE BOOK IS COMPLETE WITH THEORY AND PRACTICAL USE CASES. IT'S A MUST-HAVE FOR STUDENTS AND PROFESSORS ALIKE!" -- QIANG YANG, CHAIR OF COMPUTER SCIENCE AND ENGINEERING AT HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY "THIS IS THE MOST AMAZING AND COMPREHENSIVE TEXT BOOK ON DATA MINING. IT COVERS NOT ONLY THE FUNDAMENTAL PROBLEMS, SUCH AS CLUSTERING, CLASSIFICATION, OUTLIERS AND FREQUENT PATTERNS, AND DIFFERENT DATA TYPES, INCLUDING TEXT, TIME SERIES, SEQUENCES, SPATIAL DATA AND GRAPHS, BUT ALSO VARIOUS APPLICATIONS, SUCH AS RECOMMENDERS, WEB, SOCIAL NETWORK AND PRIVACY. IT IS A GREAT BOOK FOR GRADUATE STUDENTS AND RESEARCHERS AS WELL AS PRACTITIONERS." -- PHILIP S. YU, UIC DISTINGUISHED PROFESSOR AND WEXLER CHAIR IN INFORMATION TECHNOLOGY AT UNIVERSITY OF ILLINOIS AT CHICAGO

DATA MINING AND ANALYSIS - MOHAMMED J. ZAKI 2014-05-12

A COMPREHENSIVE OVERVIEW OF DATA MINING FROM AN ALGORITHMIC PERSPECTIVE, INTEGRATING RELATED CONCEPTS FROM MACHINE LEARNING AND STATISTICS.

PRACTICAL TIME SERIES ANALYSIS - AILEEN NIELSEN 2019-09-20

TIME SERIES DATA ANALYSIS IS INCREASINGLY IMPORTANT DUE TO THE MASSIVE PRODUCTION OF SUCH DATA THROUGH THE INTERNET OF THINGS, THE DIGITALIZATION OF HEALTHCARE, AND THE RISE OF SMART CITIES. AS CONTINUOUS MONITORING AND DATA COLLECTION BECOME MORE COMMON, THE NEED FOR COMPETENT TIME SERIES ANALYSIS WITH BOTH STATISTICAL AND MACHINE LEARNING TECHNIQUES WILL INCREASE. COVERING INNOVATIONS IN TIME SERIES DATA ANALYSIS AND USE CASES FROM THE REAL WORLD, THIS PRACTICAL GUIDE WILL HELP YOU SOLVE THE MOST COMMON DATA ENGINEERING AND ANALYSIS CHALLENGES IN TIME SERIES, USING BOTH TRADITIONAL STATISTICAL AND MODERN MACHINE LEARNING TECHNIQUES. AUTHOR AILEEN NIELSEN OFFERS AN ACCESSIBLE, WELL-ROUNDED INTRODUCTION TO TIME SERIES IN BOTH R AND PYTHON THAT WILL HAVE DATA SCIENTISTS, SOFTWARE ENGINEERS, AND RESEARCHERS UP AND RUNNING QUICKLY. YOU'LL GET THE GUIDANCE YOU NEED TO CONFIDENTLY: FIND AND WRANGLE TIME SERIES DATA UNDERTAKE EXPLORATORY TIME SERIES DATA ANALYSIS STORE TEMPORAL DATA SIMULATE TIME SERIES DATA GENERATE AND SELECT FEATURES FOR A TIME SERIES MEASURE ERROR FORECAST AND CLASSIFY TIME SERIES WITH MACHINE OR DEEP LEARNING EVALUATE ACCURACY AND PERFORMANCE

PRIVACY-PRESERVING DATA MINING - CHARU C. AGGARWAL 2008-06-10

ADVANCES IN HARDWARE TECHNOLOGY HAVE INCREASED THE CAPABILITY TO STORE AND

RECORD PERSONAL DATA. THIS HAS CAUSED CONCERNS THAT PERSONAL DATA MAY BE ABUSED. THIS BOOK PROPOSES A NUMBER OF TECHNIQUES TO PERFORM THE DATA MINING TASKS IN A PRIVACY-PRESERVING WAY. THIS EDITED VOLUME CONTAINS SURVEYS BY DISTINGUISHED RESEARCHERS IN THE PRIVACY FIELD. EACH SURVEY INCLUDES THE KEY RESEARCH CONTENT AS WELL AS FUTURE RESEARCH DIRECTIONS OF A PARTICULAR TOPIC IN PRIVACY. THE BOOK IS DESIGNED FOR RESEARCHERS, PROFESSORS, AND ADVANCED-LEVEL STUDENTS IN COMPUTER SCIENCE, BUT IS ALSO SUITABLE FOR PRACTITIONERS IN INDUSTRY.

DATA CLUSTERING - CHARU C. AGGARWAL 2016-04-08

RESEARCH ON THE PROBLEM OF CLUSTERING TENDS TO BE FRAGMENTED ACROSS THE PATTERN RECOGNITION, DATABASE, DATA MINING, AND MACHINE LEARNING COMMUNITIES. ADDRESSING THIS PROBLEM IN A UNIFIED WAY, *DATA CLUSTERING: ALGORITHMS AND APPLICATIONS* PROVIDES COMPLETE COVERAGE OF THE ENTIRE AREA OF CLUSTERING, FROM BASIC METHODS TO MORE REFINED AND COMPLEX DATA CLUSTERING APPROACHES. IT PAYS SPECIAL ATTENTION TO RECENT ISSUES IN GRAPHS, SOCIAL NETWORKS, AND OTHER DOMAINS. THE BOOK FOCUSES ON THREE PRIMARY ASPECTS OF DATA CLUSTERING: METHODS, DESCRIBING KEY TECHNIQUES COMMONLY USED FOR CLUSTERING, SUCH AS FEATURE SELECTION, AGGLOMERATIVE CLUSTERING, PARTITIONAL CLUSTERING, DENSITY-BASED CLUSTERING, PROBABILISTIC CLUSTERING, GRID-BASED CLUSTERING, SPECTRAL CLUSTERING, AND NONNEGATIVE MATRIX FACTORIZATION DOMAINS, COVERING METHODS USED FOR DIFFERENT DOMAINS OF DATA, SUCH AS CATEGORICAL DATA, TEXT DATA, MULTIMEDIA DATA, GRAPH DATA, BIOLOGICAL DATA, STREAM DATA, UNCERTAIN DATA, TIME SERIES CLUSTERING, HIGH-DIMENSIONAL CLUSTERING, AND BIG DATA VARIATIONS AND INSIGHTS, DISCUSSING IMPORTANT VARIATIONS OF THE CLUSTERING PROCESS, SUCH AS SEMISUPERVISED CLUSTERING, INTERACTIVE CLUSTERING, MULTIVIEW CLUSTERING, CLUSTER ENSEMBLES, AND CLUSTER VALIDATION IN THIS BOOK, TOP RESEARCHERS FROM AROUND THE WORLD EXPLORE THE CHARACTERISTICS OF CLUSTERING PROBLEMS IN A VARIETY OF APPLICATION AREAS. THEY ALSO EXPLAIN HOW TO GLEAN DETAILED INSIGHT FROM THE CLUSTERING PROCESS—INCLUDING HOW TO VERIFY THE QUALITY OF THE UNDERLYING CLUSTERS—THROUGH SUPERVISION, HUMAN INTERVENTION, OR THE AUTOMATED GENERATION OF ALTERNATIVE CLUSTERS.

DEEP LEARNING ON GRAPHS - YAO MA 2021-09-23

A COMPREHENSIVE TEXT ON FOUNDATIONS AND TECHNIQUES OF GRAPH NEURAL NETWORKS WITH APPLICATIONS IN NLP, DATA MINING, VISION AND HEALTHCARE.

METHODOLOGIES FOR KNOWLEDGE DISCOVERY AND DATA MINING - NING ZHONG 2003-06-29

THIS BOOK CONSTITUTES THE REFEREED PROCEEDINGS OF THE THIRD PACIFIC-ASIA CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, PAKDD '99, HELD IN BEIJING, CHINA, IN APRIL 1999. THE 29 REVISED FULL PAPERS PRESENTED TOGETHER WITH 37 SHORT PAPERS WERE CAREFULLY SELECTED FROM A TOTAL OF 158 SUBMISSIONS. THE BOOK IS DIVIDED INTO SECTIONS ON EMERGING KDD TECHNOLOGY; ASSOCIATION RULES; FEATURE

SELECTION AND GENERATION; MINING IN SEMI-UNSTRUCTURED DATA; INTERESTINGNESS, SURPRISINGNESS, AND EXCEPTIONS; ROUGH SETS, FUZZY LOGIC, AND NEURAL NETWORKS; INDUCTION, CLASSIFICATION, AND CLUSTERING; VISUALIZATION; CAUSAL MODELS AND GRAPH-BASED METHODS; AGENT-BASED AND DISTRIBUTED DATA MINING; AND ADVANCED TOPICS AND NEW METHODOLOGIES.

MINING GRAPH DATA - DIANE J. COOK 2006-12-18

THIS TEXT TAKES A FOCUSED AND COMPREHENSIVE LOOK AT MINING DATA REPRESENTED AS A GRAPH, WITH THE LATEST FINDINGS AND APPLICATIONS IN BOTH THEORY AND PRACTICE PROVIDED. EVEN IF YOU HAVE MINIMAL BACKGROUND IN ANALYZING GRAPH DATA, WITH THIS BOOK YOU'LL BE ABLE TO REPRESENT DATA AS GRAPHS, EXTRACT PATTERNS AND CONCEPTS FROM THE DATA, AND APPLY THE METHODOLOGIES PRESENTED IN THE TEXT TO REAL DATASETS. THERE IS A MISPRINT WITH THE LINK TO THE ACCOMPANYING WEB PAGE FOR THIS BOOK. FOR THOSE READERS WHO WOULD LIKE TO EXPERIMENT WITH THE TECHNIQUES FOUND IN THIS BOOK OR TEST THEIR OWN IDEAS ON GRAPH DATA, THE WEB PAGE FOR THE BOOK SHOULD BE [HTTP://WWW.EECS.WSU.EDU/MGD](http://www.eecs.wsu.edu/MGD).

OUTLIER DETECTION FOR TEMPORAL DATA - MANISH GUPTA 2014-04-14

OUTLIER (OR ANOMALY) DETECTION IS A VERY BROAD FIELD WHICH HAS BEEN STUDIED IN THE CONTEXT OF A LARGE NUMBER OF RESEARCH AREAS LIKE STATISTICS, DATA MINING, SENSOR NETWORKS, ENVIRONMENTAL SCIENCE, DISTRIBUTED SYSTEMS, SPATIO-TEMPORAL MINING, ETC. INITIAL RESEARCH IN OUTLIER DETECTION FOCUSED ON TIME SERIES-BASED OUTLIERS (IN STATISTICS). SINCE THEN, OUTLIER DETECTION HAS BEEN STUDIED ON A LARGE VARIETY OF DATA TYPES INCLUDING HIGH-DIMENSIONAL DATA, UNCERTAIN DATA, STREAM DATA, NETWORK DATA, TIME SERIES DATA, SPATIAL DATA, AND SPATIO-TEMPORAL DATA. WHILE THERE HAVE BEEN MANY TUTORIALS AND SURVEYS FOR GENERAL OUTLIER DETECTION, WE FOCUS ON OUTLIER DETECTION FOR TEMPORAL DATA IN THIS BOOK. A LARGE NUMBER OF APPLICATIONS GENERATE TEMPORAL DATASETS. FOR EXAMPLE, IN OUR EVERYDAY LIFE, VARIOUS KINDS OF RECORDS LIKE CREDIT, PERSONNEL, FINANCIAL, JUDICIAL, MEDICAL, ETC., ARE ALL TEMPORAL. THIS STRESSES THE NEED FOR AN ORGANIZED AND DETAILED STUDY OF OUTLIERS WITH RESPECT TO SUCH TEMPORAL DATA. IN THE PAST DECADE, THERE HAS BEEN A LOT OF RESEARCH ON VARIOUS FORMS OF TEMPORAL DATA INCLUDING CONSECUTIVE DATA SNAPSHOTS, SERIES OF DATA SNAPSHOTS AND DATA STREAMS. BESIDES THE INITIAL WORK ON TIME SERIES, RESEARCHERS HAVE FOCUSED ON RICH FORMS OF DATA INCLUDING MULTIPLE DATA STREAMS, SPATIO-TEMPORAL DATA, NETWORK DATA, COMMUNITY DISTRIBUTION DATA, ETC. COMPARED TO GENERAL OUTLIER DETECTION, TECHNIQUES FOR TEMPORAL OUTLIER DETECTION ARE VERY DIFFERENT. IN THIS BOOK, WE WILL PRESENT AN ORGANIZED PICTURE OF BOTH RECENT AND PAST RESEARCH IN TEMPORAL OUTLIER DETECTION. WE START WITH THE BASICS AND THEN RAMP UP THE READER TO THE MAIN IDEAS IN STATE-OF-THE-ART OUTLIER DETECTION TECHNIQUES. WE MOTIVATE THE IMPORTANCE OF TEMPORAL OUTLIER DETECTION AND BRIEF THE CHALLENGES BEYOND USUAL OUTLIER DETECTION. THEN, WE LIST DOWN A TAXONOMY OF PROPOSED TECHNIQUES FOR TEMPORAL OUTLIER DETECTION. SUCH

TECHNIQUES BROADLY INCLUDE STATISTICAL TECHNIQUES (LIKE AR MODELS, MARKOV MODELS, HISTOGRAMS, NEURAL NETWORKS), DISTANCE- AND DENSITY-BASED APPROACHES, GROUPING-BASED APPROACHES (CLUSTERING, COMMUNITY DETECTION), NETWORK-BASED APPROACHES, AND SPATIO-TEMPORAL OUTLIER DETECTION APPROACHES. WE SUMMARIZE BY PRESENTING A WIDE COLLECTION OF APPLICATIONS WHERE TEMPORAL OUTLIER DETECTION TECHNIQUES HAVE BEEN APPLIED TO DISCOVER INTERESTING OUTLIERS. TABLE OF CONTENTS: PREFACE / ACKNOWLEDGMENTS / FIGURE CREDITS / INTRODUCTION AND CHALLENGES / OUTLIER DETECTION FOR TIME SERIES AND DATA SEQUENCES / OUTLIER DETECTION FOR DATA STREAMS / OUTLIER DETECTION FOR DISTRIBUTED DATA STREAMS / OUTLIER DETECTION FOR SPATIO-TEMPORAL DATA / OUTLIER DETECTION FOR TEMPORAL NETWORK DATA / APPLICATIONS OF OUTLIER DETECTION FOR TEMPORAL DATA / CONCLUSIONS AND RESEARCH DIRECTIONS / BIBLIOGRAPHY / AUTHORS' BIOGRAPHIES

DATA MINING WITH DECISION TREES: THEORY AND APPLICATIONS (2ND EDITION) - MAIMON ODED Z 2014-09-03

DECISION TREES HAVE BECOME ONE OF THE MOST POWERFUL AND POPULAR APPROACHES IN KNOWLEDGE DISCOVERY AND DATA MINING; IT IS THE SCIENCE OF EXPLORING LARGE AND COMPLEX BODIES OF DATA IN ORDER TO DISCOVER USEFUL PATTERNS. DECISION TREE LEARNING CONTINUES TO EVOLVE OVER TIME. EXISTING METHODS ARE CONSTANTLY BEING IMPROVED AND NEW METHODS INTRODUCED. THIS 2ND EDITION IS DEDICATED ENTIRELY TO THE FIELD OF DECISION TREES IN DATA MINING; TO COVER ALL ASPECTS OF THIS IMPORTANT TECHNIQUE, AS WELL AS IMPROVED OR NEW METHODS AND TECHNIQUES DEVELOPED AFTER THE PUBLICATION OF OUR FIRST EDITION. IN THIS NEW EDITION, ALL CHAPTERS HAVE BEEN REVISED AND NEW TOPICS BROUGHT IN. NEW TOPICS INCLUDE COST-SENSITIVE ACTIVE LEARNING, LEARNING WITH UNCERTAIN AND IMBALANCED DATA, USING DECISION TREES BEYOND CLASSIFICATION TASKS, PRIVACY PRESERVING DECISION TREE LEARNING, LESSONS LEARNED FROM COMPARATIVE STUDIES, AND LEARNING DECISION TREES FOR BIG DATA. A WALK-THROUGH GUIDE TO EXISTING OPEN-SOURCE DATA MINING SOFTWARE IS ALSO INCLUDED IN THIS EDITION. THIS BOOK INVITES READERS TO EXPLORE THE MANY BENEFITS IN DATA MINING THAT DECISION TREES OFFER:

ARTIFICIAL INTELLIGENCE IN ASSET MANAGEMENT - SPENKE M. BARTRAM 2020-08-28

ARTIFICIAL INTELLIGENCE (AI) HAS GROWN IN PRESENCE IN ASSET MANAGEMENT AND HAS REVOLUTIONIZED THE SECTOR IN MANY WAYS. IT HAS IMPROVED PORTFOLIO MANAGEMENT, TRADING, AND RISK MANAGEMENT PRACTICES BY INCREASING EFFICIENCY, ACCURACY, AND COMPLIANCE. IN PARTICULAR, AI TECHNIQUES HELP CONSTRUCT PORTFOLIOS BASED ON MORE ACCURATE RISK AND RETURN FORECASTS AND MORE COMPLEX CONSTRAINTS. TRADING ALGORITHMS USE AI TO DEVISE NOVEL TRADING SIGNALS AND EXECUTE TRADES WITH LOWER TRANSACTION COSTS. AI ALSO IMPROVES RISK MODELING AND FORECASTING BY GENERATING INSIGHTS FROM NEW DATA SOURCES. FINALLY, ROBO-ADVISORS OWE A LARGE PART OF THEIR SUCCESS TO AI TECHNIQUES. YET THE USE OF AI CAN ALSO CREATE NEW RISKS AND CHALLENGES, SUCH AS THOSE RESULTING FROM MODEL OPACITY, COMPLEXITY, AND RELIANCE

ON DATA INTEGRITY.

MACHINE LEARNING AND KNOWLEDGE DISCOVERY FOR ENGINEERING SYSTEMS HEALTH MANAGEMENT - ASHOK N. SRIVASTAVA 2016-04-19

THIS VOLUME PRESENTS STATE-OF-THE-ART TOOLS AND TECHNIQUES FOR AUTOMATICALLY DETECTING, DIAGNOSING, AND PREDICTING THE EFFECTS OF ADVERSE EVENTS IN AN ENGINEERED SYSTEM. IT EMPHASIZES THE IMPORTANCE OF THESE TECHNIQUES IN MANAGING THE INTRICATE INTERACTIONS WITHIN AND BETWEEN ENGINEERING SYSTEMS TO MAINTAIN A HIGH DEGREE OF RELIABILITY. REFLECTING THE INTERDISCIPLINARY NATURE OF THE FIELD, THE BOOK EXPLAINS HOW THE FUNDAMENTAL ALGORITHMS AND METHODS OF BOTH PHYSICS-BASED AND DATA-DRIVEN APPROACHES EFFECTIVELY ADDRESS SYSTEMS HEALTH MANAGEMENT IN APPLICATION AREAS SUCH AS DATA CENTERS, AIRCRAFT, AND SOFTWARE SYSTEMS.

DATA MINING TECHNIQUES - MICHAEL J. A. BERRY 2004-04-09

MANY COMPANIES HAVE INVESTED IN BUILDING LARGE DATABASES AND DATA WAREHOUSES CAPABLE OF STORING VAST AMOUNTS OF INFORMATION. THIS BOOK OFFERS BUSINESS, SALES AND MARKETING MANAGERS A PRACTICAL GUIDE TO ACCESSING SUCH INFORMATION.

DATA CLUSTERING - GUOJUN GAN 2007-01-01

CLUSTER ANALYSIS IS AN UNSUPERVISED PROCESS THAT DIVIDES A SET OF OBJECTS INTO HOMOGENEOUS GROUPS. THIS BOOK STARTS WITH BASIC INFORMATION ON CLUSTER ANALYSIS, INCLUDING THE CLASSIFICATION OF DATA AND THE CORRESPONDING SIMILARITY MEASURES, FOLLOWED BY THE PRESENTATION OF OVER 50 CLUSTERING ALGORITHMS IN GROUPS ACCORDING TO SOME SPECIFIC BASELINE METHODOLOGIES SUCH AS HIERARCHICAL, CENTER-BASED, AND SEARCH-BASED METHODS. AS A RESULT, READERS AND USERS CAN EASILY IDENTIFY AN APPROPRIATE ALGORITHM FOR THEIR APPLICATIONS AND COMPARE NOVEL IDEAS WITH EXISTING RESULTS. THE BOOK ALSO PROVIDES EXAMPLES OF CLUSTERING APPLICATIONS TO ILLUSTRATE THE ADVANTAGES AND SHORTCOMINGS OF DIFFERENT CLUSTERING ARCHITECTURES AND ALGORITHMS. APPLICATION AREAS INCLUDE PATTERN RECOGNITION, ARTIFICIAL INTELLIGENCE, INFORMATION TECHNOLOGY, IMAGE PROCESSING, BIOLOGY, PSYCHOLOGY, AND MARKETING. READERS ALSO LEARN HOW TO PERFORM CLUSTER ANALYSIS WITH THE C/C++ AND MATLAB PROGRAMMING LANGUAGES.

FREQUENT PATTERN MINING - CHARU C. AGGARWAL 2014-08-29

THIS COMPREHENSIVE REFERENCE CONSISTS OF 18 CHAPTERS FROM PROMINENT RESEARCHERS IN THE FIELD. EACH CHAPTER IS SELF-CONTAINED, AND SYNTHESIZES ONE ASPECT OF FREQUENT PATTERN MINING. AN EMPHASIS IS PLACED ON SIMPLIFYING THE CONTENT, SO THAT STUDENTS AND PRACTITIONERS CAN BENEFIT FROM THE BOOK. EACH CHAPTER CONTAINS A SURVEY DESCRIBING KEY RESEARCH ON THE TOPIC, A CASE STUDY AND FUTURE DIRECTIONS. KEY TOPICS INCLUDE: PATTERN GROWTH METHODS, FREQUENT PATTERN MINING IN DATA STREAMS, MINING GRAPH PATTERNS, BIG DATA FREQUENT PATTERN MINING, ALGORITHMS FOR DATA CLUSTERING AND MORE. ADVANCED-LEVEL STUDENTS IN COMPUTER SCIENCE, RESEARCHERS AND PRACTITIONERS FROM INDUSTRY WILL FIND THIS BOOK AN INVALUABLE REFERENCE.

OUTLIER ANALYSIS - CHARU C. AGGARWAL 2016-12-10

THIS BOOK PROVIDES COMPREHENSIVE COVERAGE OF THE FIELD OF OUTLIER ANALYSIS FROM A COMPUTER SCIENCE POINT OF VIEW. IT INTEGRATES METHODS FROM DATA MINING, MACHINE LEARNING, AND STATISTICS WITHIN THE COMPUTATIONAL FRAMEWORK AND THEREFORE APPEALS TO MULTIPLE COMMUNITIES. THE CHAPTERS OF THIS BOOK CAN BE ORGANIZED INTO THREE CATEGORIES: BASIC ALGORITHMS: CHAPTERS 1 THROUGH 7 DISCUSS THE FUNDAMENTAL ALGORITHMS FOR OUTLIER ANALYSIS, INCLUDING PROBABILISTIC AND STATISTICAL METHODS, LINEAR METHODS, PROXIMITY-BASED METHODS, HIGH-DIMENSIONAL (SUBSPACE) METHODS, ENSEMBLE METHODS, AND SUPERVISED METHODS. DOMAIN-SPECIFIC METHODS: CHAPTERS 8 THROUGH 12 DISCUSS OUTLIER DETECTION ALGORITHMS FOR VARIOUS DOMAINS OF DATA, SUCH AS TEXT, CATEGORICAL DATA, TIME-SERIES DATA, DISCRETE SEQUENCE DATA, SPATIAL DATA, AND NETWORK DATA. APPLICATIONS: CHAPTER 13 IS DEVOTED TO VARIOUS APPLICATIONS OF OUTLIER ANALYSIS. SOME GUIDANCE IS ALSO PROVIDED FOR THE PRACTITIONER. THE SECOND EDITION OF THIS BOOK IS MORE DETAILED AND IS WRITTEN TO APPEAL TO BOTH RESEARCHERS AND PRACTITIONERS. SIGNIFICANT NEW MATERIAL HAS BEEN ADDED ON TOPICS SUCH AS KERNEL METHODS, ONE-CLASS SUPPORT-VECTOR MACHINES, MATRIX FACTORIZATION, NEURAL NETWORKS, OUTLIER ENSEMBLES, TIME-SERIES METHODS, AND SUBSPACE METHODS. IT IS WRITTEN AS A TEXTBOOK AND CAN BE USED FOR CLASSROOM TEACHING.

NEURAL NETWORKS AND DEEP LEARNING - CHARU C. AGGARWAL 2018-08-25

THIS BOOK COVERS BOTH CLASSICAL AND MODERN MODELS IN DEEP LEARNING. THE PRIMARY FOCUS IS ON THE THEORY AND ALGORITHMS OF DEEP LEARNING. THE THEORY AND ALGORITHMS OF NEURAL NETWORKS ARE PARTICULARLY IMPORTANT FOR UNDERSTANDING IMPORTANT CONCEPTS, SO THAT ONE CAN UNDERSTAND THE IMPORTANT DESIGN CONCEPTS OF NEURAL ARCHITECTURES IN DIFFERENT APPLICATIONS. WHY DO NEURAL NETWORKS WORK? WHEN DO THEY WORK BETTER THAN OFF-THE-SHELF MACHINE-LEARNING MODELS? WHEN IS DEPTH USEFUL? WHY IS TRAINING NEURAL NETWORKS SO HARD? WHAT ARE THE PITFALLS? THE BOOK IS ALSO RICH IN DISCUSSING DIFFERENT APPLICATIONS IN ORDER TO GIVE THE PRACTITIONER A FLAVOR OF HOW NEURAL ARCHITECTURES ARE DESIGNED FOR DIFFERENT TYPES OF PROBLEMS. APPLICATIONS ASSOCIATED WITH MANY DIFFERENT AREAS LIKE RECOMMENDER SYSTEMS, MACHINE TRANSLATION, IMAGE CAPTIONING, IMAGE CLASSIFICATION, REINFORCEMENT-LEARNING BASED GAMING, AND TEXT ANALYTICS ARE COVERED. THE CHAPTERS OF THIS BOOK SPAN THREE CATEGORIES: THE BASICS OF NEURAL NETWORKS: MANY TRADITIONAL MACHINE LEARNING MODELS CAN BE UNDERSTOOD AS SPECIAL CASES OF NEURAL NETWORKS. AN EMPHASIS IS PLACED IN THE FIRST TWO CHAPTERS ON UNDERSTANDING THE RELATIONSHIP BETWEEN TRADITIONAL MACHINE LEARNING AND NEURAL NETWORKS. SUPPORT VECTOR MACHINES, LINEAR/LOGISTIC REGRESSION, SINGULAR VALUE DECOMPOSITION, MATRIX FACTORIZATION, AND RECOMMENDER SYSTEMS ARE SHOWN TO BE SPECIAL CASES OF NEURAL NETWORKS. THESE METHODS ARE STUDIED TOGETHER WITH RECENT FEATURE ENGINEERING METHODS LIKE WORD2VEC. FUNDAMENTALS OF NEURAL NETWORKS: A DETAILED DISCUSSION

OF TRAINING AND REGULARIZATION IS PROVIDED IN CHAPTERS 3 AND 4. CHAPTERS 5 AND 6 PRESENT RADIAL-BASIS FUNCTION (RBF) NETWORKS AND RESTRICTED BOLTZMANN MACHINES. ADVANCED TOPICS IN NEURAL NETWORKS: CHAPTERS 7 AND 8 DISCUSS RECURRENT NEURAL NETWORKS AND CONVOLUTIONAL NEURAL NETWORKS. SEVERAL ADVANCED TOPICS LIKE DEEP REINFORCEMENT LEARNING, NEURAL TURING MACHINES, KOHONEN SELF-ORGANIZING MAPS, AND GENERATIVE ADVERSARIAL NETWORKS ARE INTRODUCED IN CHAPTERS 9 AND 10. THE BOOK IS WRITTEN FOR GRADUATE STUDENTS, RESEARCHERS, AND PRACTITIONERS. NUMEROUS EXERCISES ARE AVAILABLE ALONG WITH A SOLUTION MANUAL TO AID IN CLASSROOM TEACHING. WHERE POSSIBLE, AN APPLICATION-CENTRIC VIEW IS HIGHLIGHTED IN ORDER TO PROVIDE AN UNDERSTANDING OF THE PRACTICAL USES OF EACH CLASS OF TECHNIQUES.

DATA MINING - CHARU C. AGGARWAL 2015-04-13

THIS TEXTBOOK EXPLORES THE DIFFERENT ASPECTS OF DATA MINING FROM THE FUNDAMENTALS TO THE COMPLEX DATA TYPES AND THEIR APPLICATIONS, CAPTURING THE WIDE DIVERSITY OF PROBLEM DOMAINS FOR DATA MINING ISSUES. IT GOES BEYOND THE TRADITIONAL FOCUS ON DATA MINING PROBLEMS TO INTRODUCE ADVANCED DATA TYPES SUCH AS TEXT, TIME SERIES, DISCRETE SEQUENCES, SPATIAL DATA, GRAPH DATA, AND SOCIAL NETWORKS. UNTIL NOW, NO SINGLE BOOK HAS ADDRESSED ALL THESE TOPICS IN A COMPREHENSIVE AND INTEGRATED WAY. THE CHAPTERS OF THIS BOOK FALL INTO ONE OF THREE CATEGORIES: FUNDAMENTAL CHAPTERS: DATA MINING HAS FOUR MAIN PROBLEMS, WHICH CORRESPOND TO CLUSTERING, CLASSIFICATION, ASSOCIATION PATTERN MINING, AND OUTLIER ANALYSIS. THESE CHAPTERS COMPREHENSIVELY DISCUSS A WIDE VARIETY OF METHODS FOR THESE PROBLEMS. DOMAIN CHAPTERS: THESE CHAPTERS DISCUSS THE SPECIFIC METHODS USED FOR DIFFERENT DOMAINS OF DATA SUCH AS TEXT DATA, TIME-SERIES DATA, SEQUENCE DATA, GRAPH DATA, AND SPATIAL DATA. APPLICATION CHAPTERS: THESE CHAPTERS STUDY IMPORTANT APPLICATIONS SUCH AS STREAM MINING, WEB MINING, RANKING, RECOMMENDATIONS, SOCIAL NETWORKS, AND PRIVACY PRESERVATION. THE DOMAIN CHAPTERS ALSO HAVE AN APPLIED FLAVOR. APPROPRIATE FOR BOTH INTRODUCTORY AND ADVANCED DATA MINING COURSES, DATA MINING: THE TEXTBOOK BALANCES MATHEMATICAL DETAILS AND INTUITION. IT CONTAINS THE NECESSARY MATHEMATICAL DETAILS FOR PROFESSORS AND RESEARCHERS, BUT IT IS PRESENTED IN A SIMPLE AND INTUITIVE STYLE TO IMPROVE ACCESSIBILITY FOR STUDENTS AND INDUSTRIAL PRACTITIONERS (INCLUDING THOSE WITH A LIMITED MATHEMATICAL BACKGROUND). NUMEROUS ILLUSTRATIONS, EXAMPLES, AND EXERCISES ARE INCLUDED, WITH AN EMPHASIS ON SEMANTICALLY INTERPRETABLE EXAMPLES. PRAISE FOR DATA MINING: THE TEXTBOOK - "AS I READ THROUGH THIS BOOK, I HAVE ALREADY DECIDED TO USE IT IN MY CLASSES. THIS IS A BOOK WRITTEN BY AN OUTSTANDING RESEARCHER WHO HAS MADE FUNDAMENTAL CONTRIBUTIONS TO DATA MINING, IN A WAY THAT IS BOTH ACCESSIBLE AND UP TO DATE. THE BOOK IS COMPLETE WITH THEORY AND PRACTICAL USE CASES. IT'S A MUST-HAVE FOR STUDENTS AND PROFESSORS ALIKE!" -- QIANG YANG, CHAIR OF COMPUTER SCIENCE AND ENGINEERING AT HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY "THIS IS THE MOST AMAZING AND COMPREHENSIVE TEXT BOOK

ON DATA MINING. IT COVERS NOT ONLY THE FUNDAMENTAL PROBLEMS, SUCH AS CLUSTERING, CLASSIFICATION, OUTLIERS AND FREQUENT PATTERNS, AND DIFFERENT DATA TYPES, INCLUDING TEXT, TIME SERIES, SEQUENCES, SPATIAL DATA AND GRAPHS, BUT ALSO VARIOUS APPLICATIONS, SUCH AS RECOMMENDERS, WEB, SOCIAL NETWORK AND PRIVACY. IT IS A GREAT BOOK FOR GRADUATE STUDENTS AND RESEARCHERS AS WELL AS PRACTITIONERS." -- PHILIP S. YU, UIC DISTINGUISHED PROFESSOR AND WEXLER CHAIR IN INFORMATION TECHNOLOGY AT UNIVERSITY OF ILLINOIS AT CHICAGO
BIG CRISIS DATA - CARLOS CASTILLO 2016-07-04

SOCIAL MEDIA IS AN INVALUABLE SOURCE OF TIME-CRITICAL INFORMATION DURING A CRISIS. HOWEVER, EMERGENCY RESPONSE AND HUMANITARIAN RELIEF ORGANIZATIONS THAT WOULD LIKE TO USE THIS INFORMATION STRUGGLE WITH AN AVALANCHE OF SOCIAL MEDIA MESSAGES THAT EXCEEDS HUMAN CAPACITY TO PROCESS. EMERGENCY MANAGERS, DECISION MAKERS, AND AFFECTED COMMUNITIES CAN MAKE SENSE OF SOCIAL MEDIA THROUGH A COMBINATION OF MACHINE COMPUTATION AND HUMAN COMPASSION - EXPRESSED BY THOUSANDS OF DIGITAL VOLUNTEERS WHO PUBLISH, PROCESS, AND SUMMARIZE POTENTIALLY LIFE-SAVING INFORMATION. THIS BOOK BRINGS TOGETHER COMPUTATIONAL METHODS FROM MANY DISCIPLINES: NATURAL LANGUAGE PROCESSING, SEMANTIC TECHNOLOGIES, DATA MINING, MACHINE LEARNING, NETWORK ANALYSIS, HUMAN-COMPUTER INTERACTION, AND INFORMATION VISUALIZATION, FOCUSING ON METHODS THAT ARE COMMONLY USED FOR PROCESSING SOCIAL MEDIA MESSAGES UNDER TIME-CRITICAL CONSTRAINTS, AND OFFERING MORE THAN 500 REFERENCES TO IN-DEPTH INFORMATION.

SOCIAL NETWORK DATA ANALYTICS - CHARU C. AGGARWAL 2011-03-18

SOCIAL NETWORK ANALYSIS APPLICATIONS HAVE EXPERIENCED TREMENDOUS ADVANCES WITHIN THE LAST FEW YEARS DUE IN PART TO INCREASING TRENDS TOWARDS USERS INTERACTING WITH EACH OTHER ON THE INTERNET. SOCIAL NETWORKS ARE ORGANIZED AS GRAPHS, AND THE DATA ON SOCIAL NETWORKS TAKES ON THE FORM OF MASSIVE STREAMS, WHICH ARE MINED FOR A VARIETY OF PURPOSES. SOCIAL NETWORK DATA ANALYTICS COVERS AN IMPORTANT NICHE IN THE SOCIAL NETWORK ANALYTICS FIELD. THIS EDITED VOLUME, CONTRIBUTED BY PROMINENT RESEARCHERS IN THIS FIELD, PRESENTS A WIDE SELECTION OF TOPICS ON SOCIAL NETWORK DATA MINING SUCH AS STRUCTURAL PROPERTIES OF SOCIAL NETWORKS, ALGORITHMS FOR STRUCTURAL DISCOVERY OF SOCIAL NETWORKS AND CONTENT ANALYSIS IN SOCIAL NETWORKS. THIS BOOK IS ALSO UNIQUE IN FOCUSING ON THE DATA ANALYTICAL ASPECTS OF SOCIAL NETWORKS IN THE INTERNET SCENARIO, RATHER THAN THE TRADITIONAL SOCIOLOGY-DRIVEN EMPHASIS PREVALENT IN THE EXISTING BOOKS, WHICH DO NOT FOCUS ON THE UNIQUE DATA-INTENSIVE CHARACTERISTICS OF ONLINE SOCIAL NETWORKS. EMPHASIS IS PLACED ON SIMPLIFYING THE CONTENT SO THAT STUDENTS AND PRACTITIONERS BENEFIT FROM THIS BOOK. THIS BOOK TARGETS ADVANCED LEVEL STUDENTS AND RESEARCHERS CONCENTRATING ON COMPUTER SCIENCE AS A SECONDARY TEXT OR REFERENCE BOOK. DATA MINING, DATABASE, INFORMATION SECURITY, ELECTRONIC COMMERCE AND MACHINE LEARNING PROFESSIONALS WILL FIND THIS BOOK A VALUABLE ASSET,

AS WELL AS PRIMARY ASSOCIATIONS SUCH AS ACM, IEEE AND MANAGEMENT SCIENCE.
MINING TEXT DATA - CHARU C. AGGARWAL 2012-02-03

TEXT MINING APPLICATIONS HAVE EXPERIENCED TREMENDOUS ADVANCES BECAUSE OF WEB 2.0 AND SOCIAL NETWORKING APPLICATIONS. RECENT ADVANCES IN HARDWARE AND SOFTWARE TECHNOLOGY HAVE LEAD TO A NUMBER OF UNIQUE SCENARIOS WHERE TEXT MINING ALGORITHMS ARE LEARNED. MINING TEXT DATA INTRODUCES AN IMPORTANT NICHE IN THE TEXT ANALYTICS FIELD, AND IS AN EDITED VOLUME CONTRIBUTED BY LEADING INTERNATIONAL RESEARCHERS AND PRACTITIONERS FOCUSED ON SOCIAL NETWORKS & DATA MINING. THIS BOOK CONTAINS A WIDE SWATH IN TOPICS ACROSS SOCIAL NETWORKS & DATA MINING. EACH CHAPTER CONTAINS A COMPREHENSIVE SURVEY INCLUDING THE KEY RESEARCH CONTENT ON THE TOPIC, AND THE FUTURE DIRECTIONS OF RESEARCH IN THE FIELD. THERE IS A SPECIAL FOCUS ON TEXT EMBEDDED WITH HETEROGENEOUS AND MULTIMEDIA DATA WHICH MAKES THE MINING PROCESS MUCH MORE CHALLENGING. A NUMBER OF METHODS HAVE BEEN DESIGNED SUCH AS TRANSFER LEARNING AND CROSS-LINGUAL MINING FOR SUCH CASES. MINING TEXT DATA SIMPLIFIES THE CONTENT, SO THAT ADVANCED-LEVEL STUDENTS, PRACTITIONERS AND RESEARCHERS IN COMPUTER SCIENCE CAN BENEFIT FROM THIS BOOK. ACADEMIC AND CORPORATE LIBRARIES, AS WELL AS ACM, IEEE, AND MANAGEMENT SCIENCE FOCUSED ON INFORMATION SECURITY, ELECTRONIC COMMERCE, DATABASES, DATA MINING, MACHINE LEARNING, AND STATISTICS ARE THE PRIMARY BUYERS FOR THIS REFERENCE BOOK.

HEALTHCARE DATA ANALYTICS - CHANDAN K. REDDY 2015-06-23

AT THE INTERSECTION OF COMPUTER SCIENCE AND HEALTHCARE, DATA ANALYTICS HAS EMERGED AS A PROMISING TOOL FOR SOLVING PROBLEMS ACROSS MANY HEALTHCARE-RELATED DISCIPLINES. SUPPLYING A COMPREHENSIVE OVERVIEW OF RECENT HEALTHCARE ANALYTICS RESEARCH, HEALTHCARE DATA ANALYTICS PROVIDES A CLEAR UNDERSTANDING OF THE ANALYTICAL TECHNIQUES CURRENTLY AVAILABLE TO SOLVE HEALTHCARE PROBLEMS. THE BOOK DETAILS NOVEL TECHNIQUES FOR ACQUIRING, HANDLING, RETRIEVING, AND MAKING BEST USE OF HEALTHCARE DATA. IT ANALYZES RECENT DEVELOPMENTS IN HEALTHCARE COMPUTING AND DISCUSSES EMERGING TECHNOLOGIES THAT CAN HELP IMPROVE THE HEALTH AND WELL-BEING OF PATIENTS. WRITTEN BY PROMINENT RESEARCHERS AND EXPERTS WORKING IN THE HEALTHCARE DOMAIN, THE BOOK SHEDS LIGHT ON MANY OF THE COMPUTATIONAL CHALLENGES IN THE FIELD OF MEDICAL INFORMATICS. EACH CHAPTER IN THE BOOK IS STRUCTURED AS A "SURVEY-STYLE" ARTICLE DISCUSSING THE PROMINENT RESEARCH ISSUES AND THE ADVANCES MADE ON THAT RESEARCH TOPIC. THE BOOK IS DIVIDED INTO THREE MAJOR CATEGORIES: HEALTHCARE DATA SOURCES AND BASIC ANALYTICS - DETAILS THE VARIOUS HEALTHCARE DATA SOURCES AND ANALYTICAL TECHNIQUES USED IN THE PROCESSING AND ANALYSIS OF SUCH DATA ADVANCED DATA ANALYTICS FOR HEALTHCARE - COVERS ADVANCED ANALYTICAL METHODS, INCLUDING CLINICAL PREDICTION MODELS, TEMPORAL PATTERN MINING METHODS, AND VISUAL ANALYTICS APPLICATIONS AND PRACTICAL SYSTEMS FOR HEALTHCARE - COVERS THE APPLICATIONS OF DATA ANALYTICS TO PERVASIVE HEALTHCARE, FRAUD DETECTION, AND DRUG DISCOVERY ALONG WITH

SYSTEMS FOR MEDICAL IMAGING AND DECISION SUPPORT COMPUTER SCIENTISTS ARE USUALLY NOT TRAINED IN DOMAIN-SPECIFIC MEDICAL CONCEPTS, WHEREAS MEDICAL PRACTITIONERS AND RESEARCHERS HAVE LIMITED EXPOSURE TO THE DATA ANALYTICS AREA. THE CONTENTS OF THIS BOOK WILL HELP TO BRING TOGETHER THESE DIVERSE COMMUNITIES BY CAREFULLY AND COMPREHENSIVELY DISCUSSING THE MOST RELEVANT CONTRIBUTIONS FROM EACH DOMAIN.

DATABASE THEORY - ICDT 2001 - JAN VAN DEN BUSSCHE 2003-06-29

THIS BOOK CONSTITUTES THE REFEREED PROCEEDINGS OF THE 8TH INTERNATIONAL CONFERENCE ON DATABASE THEORY, ICDT 2001, HELD IN LONDON, UK, IN JANUARY 2001. THE 26 REVISED FULL PAPERS PRESENTED TOGETHER WITH TWO INVITED PAPERS WERE CAREFULLY REVIEWED AND SELECTED FROM 75 SUBMISSIONS. ALL CURRENT ISSUES ON DATABASE THEORY AND THE FOUNDATIONS OF DATABASE SYSTEMS ARE ADDRESSED. AMONG THE TOPICS COVERED ARE DATABASE QUERIES, SQL, INFORMATION RETRIEVAL, DATABASE LOGIC, DATABASE MINING, CONSTRAINT DATABASES, TRANSACTIONS, ALGORITHMIC ASPECTS, SEMI-STRUCTURED DATA, DATA ENGINEERING, XML, TERM REWRITING, CLUSTERING, ETC.

DATA MINING: CONCEPTS AND TECHNIQUES - JIAWEI HAN 2011-06-09

DATA MINING: CONCEPTS AND TECHNIQUES PROVIDES THE CONCEPTS AND TECHNIQUES IN PROCESSING GATHERED DATA OR INFORMATION, WHICH WILL BE USED IN VARIOUS APPLICATIONS. SPECIFICALLY, IT EXPLAINS DATA MINING AND THE TOOLS USED IN DISCOVERING KNOWLEDGE FROM THE COLLECTED DATA. THIS BOOK IS REFERRED AS THE KNOWLEDGE DISCOVERY FROM DATA (KDD). IT FOCUSES ON THE FEASIBILITY, USEFULNESS, EFFECTIVENESS, AND SCALABILITY OF TECHNIQUES OF LARGE DATA SETS. AFTER DESCRIBING DATA MINING, THIS EDITION EXPLAINS THE METHODS OF KNOWING, PREPROCESSING, PROCESSING, AND WAREHOUSING DATA. IT THEN PRESENTS INFORMATION ABOUT DATA WAREHOUSES, ONLINE ANALYTICAL PROCESSING (OLAP), AND DATA CUBE TECHNOLOGY. THEN, THE METHODS INVOLVED IN MINING FREQUENT PATTERNS, ASSOCIATIONS, AND CORRELATIONS FOR LARGE DATA SETS ARE DESCRIBED. THE BOOK DETAILS THE METHODS FOR DATA CLASSIFICATION AND INTRODUCES THE CONCEPTS AND METHODS FOR DATA CLUSTERING. THE REMAINING CHAPTERS DISCUSS THE OUTLIER DETECTION AND THE TRENDS, APPLICATIONS, AND RESEARCH FRONTIERS IN DATA MINING. THIS BOOK IS INTENDED FOR COMPUTER SCIENCE STUDENTS, APPLICATION DEVELOPERS, BUSINESS PROFESSIONALS, AND RESEARCHERS WHO SEEK INFORMATION ON DATA MINING. PRESENTS DOZENS OF ALGORITHMS AND IMPLEMENTATION EXAMPLES, ALL IN PSEUDO-CODE AND SUITABLE FOR USE IN REAL-WORLD, LARGE-SCALE DATA MINING PROJECTS ADDRESSES ADVANCED TOPICS SUCH AS MINING OBJECT-RELATIONAL DATABASES, SPATIAL DATABASES, MULTIMEDIA DATABASES, TIME-SERIES DATABASES, TEXT DATABASES, THE WORLD WIDE WEB, AND APPLICATIONS IN SEVERAL FIELDS PROVIDES A COMPREHENSIVE, PRACTICAL LOOK AT THE CONCEPTS AND TECHNIQUES YOU NEED TO GET THE MOST OUT OF YOUR DATA

COMPUTATIONAL METHODS OF FEATURE SELECTION - HUAN LIU 2007-10-29

DUE TO INCREASING DEMANDS FOR DIMENSIONALITY REDUCTION, RESEARCH ON FEATURE

SELECTION HAS DEEPLY AND WIDELY EXPANDED INTO MANY FIELDS, INCLUDING COMPUTATIONAL STATISTICS, PATTERN RECOGNITION, MACHINE LEARNING, DATA MINING, AND KNOWLEDGE DISCOVERY. HIGHLIGHTING CURRENT RESEARCH ISSUES, *COMPUTATIONAL METHODS OF FEATURE SELECTION* INTRODUCES THE BASIC CONCEPTS AND PRINCIPLES, STATE-OF-THE-ART ALGORITHMS, AND NOVEL APPLICATIONS OF THIS TOOL. THE BOOK BEGINS BY EXPLORING UNSUPERVISED, RANDOMIZED, AND CAUSAL FEATURE SELECTION. IT THEN REPORTS ON SOME RECENT RESULTS OF EMPOWERING FEATURE SELECTION, INCLUDING ACTIVE FEATURE SELECTION, DECISION-BORDER ESTIMATE, THE USE OF ENSEMBLES WITH INDEPENDENT PROBES, AND INCREMENTAL FEATURE SELECTION. THIS IS FOLLOWED BY DISCUSSIONS OF WEIGHTING AND LOCAL METHODS, SUCH AS THE RELIEFF FAMILY, K-MEANS CLUSTERING, LOCAL FEATURE RELEVANCE, AND A NEW INTERPRETATION OF RELIEF. THE BOOK SUBSEQUENTLY COVERS TEXT CLASSIFICATION, A NEW FEATURE SELECTION SCORE, AND BOTH CONSTRAINT-GUIDED AND AGGRESSIVE FEATURE SELECTION. THE FINAL SECTION EXAMINES APPLICATIONS OF FEATURE SELECTION IN BIOINFORMATICS, INCLUDING FEATURE CONSTRUCTION AS WELL AS REDUNDANCY-, ENSEMBLE-, AND PENALTY-BASED FEATURE SELECTION. THROUGH A CLEAR, CONCISE, AND COHERENT PRESENTATION OF TOPICS, THIS VOLUME SYSTEMATICALLY COVERS THE KEY CONCEPTS, UNDERLYING PRINCIPLES, AND INVENTIVE APPLICATIONS OF FEATURE SELECTION, ILLUSTRATING HOW THIS POWERFUL TOOL CAN EFFICIENTLY HARNESS MASSIVE, HIGH-DIMENSIONAL DATA AND TURN IT INTO VALUABLE, RELIABLE INFORMATION.

MACHINE LEARNING FOR TEXT - CHARU C. AGGARWAL 2018-03-19

TEXT ANALYTICS IS A FIELD THAT LIES ON THE INTERFACE OF INFORMATION RETRIEVAL, MACHINE LEARNING, AND NATURAL LANGUAGE PROCESSING, AND THIS TEXTBOOK CAREFULLY COVERS A COHERENTLY ORGANIZED FRAMEWORK DRAWN FROM THESE INTERSECTING TOPICS. THE CHAPTERS OF THIS TEXTBOOK IS ORGANIZED INTO THREE CATEGORIES: - BASIC ALGORITHMS: CHAPTERS 1 THROUGH 7 DISCUSS THE CLASSICAL ALGORITHMS FOR MACHINE LEARNING FROM TEXT SUCH AS PREPROCESSING, SIMILARITY COMPUTATION, TOPIC MODELING, MATRIX FACTORIZATION, CLUSTERING, CLASSIFICATION, REGRESSION, AND ENSEMBLE ANALYSIS. - DOMAIN-SENSITIVE MINING: CHAPTERS 8 AND 9 DISCUSS THE LEARNING METHODS FROM TEXT WHEN COMBINED WITH DIFFERENT DOMAINS SUCH AS MULTIMEDIA AND THE WEB. THE PROBLEM OF INFORMATION RETRIEVAL AND WEB SEARCH IS ALSO DISCUSSED IN THE CONTEXT OF ITS RELATIONSHIP WITH RANKING AND MACHINE LEARNING METHODS. - SEQUENCE-CENTRIC MINING: CHAPTERS 10 THROUGH 14 DISCUSS VARIOUS SEQUENCE-CENTRIC AND NATURAL LANGUAGE APPLICATIONS, SUCH AS FEATURE ENGINEERING, NEURAL LANGUAGE MODELS, DEEP LEARNING, TEXT SUMMARIZATION, INFORMATION EXTRACTION, OPINION MINING, TEXT SEGMENTATION, AND EVENT DETECTION. THIS TEXTBOOK COVERS MACHINE LEARNING TOPICS FOR TEXT IN DETAIL. SINCE THE COVERAGE IS EXTENSIVE, MULTIPLE COURSES CAN BE OFFERED FROM THE SAME BOOK, DEPENDING ON COURSE LEVEL. EVEN THOUGH THE PRESENTATION IS TEXT-CENTRIC, CHAPTERS 3 TO 7 COVER MACHINE LEARNING ALGORITHMS THAT ARE OFTEN USED IN DOMAINS BEYOND

TEXT DATA. THEREFORE, THE BOOK CAN BE USED TO OFFER COURSES NOT JUST IN TEXT ANALYTICS BUT ALSO FROM THE BROADER PERSPECTIVE OF MACHINE LEARNING (WITH TEXT AS A BACKDROP). THIS TEXTBOOK TARGETS GRADUATE STUDENTS IN COMPUTER SCIENCE, AS WELL AS RESEARCHERS, PROFESSORS, AND INDUSTRIAL PRACTITIONERS WORKING IN THESE RELATED FIELDS. THIS TEXTBOOK IS ACCOMPANIED WITH A SOLUTION MANUAL FOR CLASSROOM TEACHING.

DATA CLUSTERING - CHARU C. AGGARWAL 2013-08-21

RESEARCH ON THE PROBLEM OF CLUSTERING TENDS TO BE FRAGMENTED ACROSS THE PATTERN RECOGNITION, DATABASE, DATA MINING, AND MACHINE LEARNING COMMUNITIES. ADDRESSING THIS PROBLEM IN A UNIFIED WAY, *DATA CLUSTERING: ALGORITHMS AND APPLICATIONS* PROVIDES COMPLETE COVERAGE OF THE ENTIRE AREA OF CLUSTERING, FROM BASIC METHODS TO MORE REFINED AND COMPLEX DATA CLUSTERING APPROACHES. IT PAYS SPECIAL ATTENTION TO RECENT ISSUES IN GRAPHS, SOCIAL NETWORKS, AND OTHER DOMAINS. THE BOOK FOCUSES ON THREE PRIMARY ASPECTS OF DATA CLUSTERING: METHODS, DESCRIBING KEY TECHNIQUES COMMONLY USED FOR CLUSTERING, SUCH AS FEATURE SELECTION, AGGLOMERATIVE CLUSTERING, PARTITIONAL CLUSTERING, DENSITY-BASED CLUSTERING, PROBABILISTIC CLUSTERING, GRID-BASED CLUSTERING, SPECTRAL CLUSTERING, AND NONNEGATIVE MATRIX FACTORIZATION DOMAINS, COVERING METHODS USED FOR DIFFERENT DOMAINS OF DATA, SUCH AS CATEGORICAL DATA, TEXT DATA, MULTIMEDIA DATA, GRAPH DATA, BIOLOGICAL DATA, STREAM DATA, UNCERTAIN DATA, TIME SERIES CLUSTERING, HIGH-DIMENSIONAL CLUSTERING, AND BIG DATA VARIATIONS AND INSIGHTS, DISCUSSING IMPORTANT VARIATIONS OF THE CLUSTERING PROCESS, SUCH AS SEMISUPERVISED CLUSTERING, INTERACTIVE CLUSTERING, MULTIVIEW CLUSTERING, CLUSTER ENSEMBLES, AND CLUSTER VALIDATION IN THIS BOOK, TOP RESEARCHERS FROM AROUND THE WORLD EXPLORE THE CHARACTERISTICS OF CLUSTERING PROBLEMS IN A VARIETY OF APPLICATION AREAS. THEY ALSO EXPLAIN HOW TO GLEAN DETAILED INSIGHT FROM THE CLUSTERING PROCESS—INCLUDING HOW TO VERIFY THE QUALITY OF THE UNDERLYING CLUSTERS—THROUGH SUPERVISION, HUMAN INTERVENTION, OR THE AUTOMATED GENERATION OF ALTERNATIVE CLUSTERS.

CLUSTERING AND INFORMATION RETRIEVAL - WEILI WU 2013-12-01

CLUSTERING IS AN IMPORTANT TECHNIQUE FOR DISCOVERING RELATIVELY DENSE SUB-REGIONS OR SUB-SPACES OF A MULTI-DIMENSION DATA DISTRIBUTION. CLUSTERING HAS BEEN USED IN INFORMATION RETRIEVAL FOR MANY DIFFERENT PURPOSES, SUCH AS QUERY EXPANSION, DOCUMENT GROUPING, DOCUMENT INDEXING, AND VISUALIZATION OF SEARCH RESULTS. IN THIS BOOK, WE ADDRESS ISSUES OF CLUSTERING ALGORITHMS, EVALUATION METHODOLOGIES, APPLICATIONS, AND ARCHITECTURES FOR INFORMATION RETRIEVAL. THE FIRST TWO CHAPTERS DISCUSS CLUSTERING ALGORITHMS. THE CHAPTER FROM BAEZA-YATES ET AL. DESCRIBES A CLUSTERING METHOD FOR A GENERAL METRIC SPACE WHICH IS A COMMON MODEL OF DATA RELEVANT TO INFORMATION RETRIEVAL. THE CHAPTER BY GUHA, RASTOGI, AND SHIM PRESENTS A SURVEY AS WELL AS DETAILED DISCUSSION OF TWO CLUSTERING

ALGORITHMS: CURE AND ROCK FOR NUMERIC DATA AND CATEGORICAL DATA RESPECTIVELY. EVALUATION METHODOLOGIES ARE ADDRESSED IN THE NEXT TWO CHAPTERS. ERTOZ ET AL. DEMONSTRATE THE USE OF TEXT RETRIEVAL BENCHMARKS, SUCH AS TRECS, TO EVALUATE CLUSTERING ALGORITHMS. HE ET AL. PROVIDE OBJECTIVE MEASURES OF CLUSTERING QUALITY IN THEIR CHAPTER. APPLICATIONS OF CLUSTERING METHODS TO INFORMATION RETRIEVAL IS ADDRESSED IN THE NEXT FOUR CHAPTERS. CHU ET AL. AND NOEL ET AL. EXPLORE FEATURE SELECTION USING WORD STEMS, PHRASES, AND LINK ASSOCIATIONS FOR DOCUMENT CLUSTERING AND INDEXING. WEN ET AL. AND SUNG ET AL. DISCUSS APPLICATIONS OF CLUSTERING TO USER QUERIES AND DATA CLEANSING. FINALLY, WE CONSIDER THE PROBLEM OF DESIGNING ARCHITECTURES FOR INFORMATION RETRIEVAL. CRICHTON, HUGHES, AND KELLY ELABORATE ON THE DEVELOPMENT OF A SCIENTIFIC DATA SYSTEM ARCHITECTURE FOR INFORMATION RETRIEVAL.

DATA CLASSIFICATION - CHARU C. AGGARWAL 2014-07-25

RESEARCH ON THE PROBLEM OF CLASSIFICATION TENDS TO BE FRAGMENTED ACROSS SUCH AREAS AS PATTERN RECOGNITION, DATABASE, DATA MINING, AND MACHINE LEARNING. ADDRESSING THE WORK OF THESE DIFFERENT COMMUNITIES IN A UNIFIED WAY, *DATA CLASSIFICATION: ALGORITHMS AND APPLICATIONS* EXPLORES THE UNDERLYING

ARTIFICIAL INTELLIGENCE - CHARU C. AGGARWAL 2022-07-31

THIS TEXTBOOK COVERS THE BROADER FIELD OF ARTIFICIAL INTELLIGENCE. THE CHAPTERS FOR THIS TEXTBOOK SPAN WITHIN THREE CATEGORIES: DEDUCTIVE REASONING METHODS: THESE METHODS START WITH PRE-DEFINED HYPOTHESES AND REASON WITH THEM IN ORDER TO ARRIVE AT LOGICALLY SOUND CONCLUSIONS. THE UNDERLYING METHODS INCLUDE SEARCH AND LOGIC-BASED METHODS. THESE METHODS ARE DISCUSSED IN CHAPTERS 1 THROUGH 5. INDUCTIVE LEARNING METHODS: THESE METHODS START WITH EXAMPLES AND USE STATISTICAL METHODS IN ORDER TO ARRIVE AT HYPOTHESES. EXAMPLES INCLUDE REGRESSION MODELING, SUPPORT VECTOR MACHINES, NEURAL NETWORKS, REINFORCEMENT LEARNING, UNSUPERVISED LEARNING, AND PROBABILISTIC GRAPHICAL MODELS. THESE METHODS ARE DISCUSSED IN CHAPTERS 6 THROUGH 11. INTEGRATING REASONING AND LEARNING: CHAPTERS 11 AND 12 DISCUSS TECHNIQUES FOR INTEGRATING REASONING AND LEARNING. EXAMPLES INCLUDE THE USE OF KNOWLEDGE GRAPHS AND NEURO-SYMBOLIC ARTIFICIAL INTELLIGENCE. THE PRIMARY AUDIENCE FOR THIS TEXTBOOK ARE PROFESSORS AND ADVANCED-LEVEL STUDENTS IN COMPUTER SCIENCE. IT IS ALSO POSSIBLE TO USE THIS TEXTBOOK FOR THE MATHEMATICS REQUIREMENTS FOR AN UNDERGRADUATE DATA SCIENCE COURSE. PROFESSIONALS WORKING IN THIS RELATED FIELD MANY ALSO FIND THIS TEXTBOOK USEFUL AS A REFERENCE.

DATA STREAMS - CHARU C. AGGARWAL 2007-04-03

THIS BOOK PRIMARILY DISCUSSES ISSUES RELATED TO THE MINING ASPECTS OF DATA STREAMS AND IT IS UNIQUE IN ITS PRIMARY FOCUS ON THE SUBJECT. THIS VOLUME COVERS MINING ASPECTS OF DATA STREAMS COMPREHENSIVELY: EACH CONTRIBUTED CHAPTER

CONTAINS A SURVEY ON THE TOPIC, THE KEY IDEAS IN THE FIELD FOR THAT PARTICULAR TOPIC, AND FUTURE RESEARCH DIRECTIONS. THE BOOK IS INTENDED FOR A PROFESSIONAL AUDIENCE COMPOSED OF RESEARCHERS AND PRACTITIONERS IN INDUSTRY. THIS BOOK IS ALSO APPROPRIATE FOR ADVANCED-LEVEL STUDENTS IN COMPUTER SCIENCE.

OUTLIER ENSEMBLES - CHARU C. AGGARWAL 2017-04-06

THIS BOOK DISCUSSES A VARIETY OF METHODS FOR OUTLIER ENSEMBLES AND ORGANIZES THEM BY THE SPECIFIC PRINCIPLES WITH WHICH ACCURACY IMPROVEMENTS ARE ACHIEVED. IN ADDITION, IT COVERS THE TECHNIQUES WITH WHICH SUCH METHODS CAN BE MADE MORE EFFECTIVE. A FORMAL CLASSIFICATION OF THESE METHODS IS PROVIDED, AND THE CIRCUMSTANCES IN WHICH THEY WORK WELL ARE EXAMINED. THE AUTHORS COVER HOW OUTLIER ENSEMBLES RELATE (BOTH THEORETICALLY AND PRACTICALLY) TO THE ENSEMBLE TECHNIQUES USED COMMONLY FOR OTHER DATA MINING PROBLEMS LIKE CLASSIFICATION. THE SIMILARITIES AND (SUBTLE) DIFFERENCES IN THE ENSEMBLE TECHNIQUES FOR THE CLASSIFICATION AND OUTLIER DETECTION PROBLEMS ARE EXPLORED. THESE SUBTLE DIFFERENCES DO IMPACT THE DESIGN OF ENSEMBLE ALGORITHMS FOR THE LATTER PROBLEM. THIS BOOK CAN BE USED FOR COURSES IN DATA MINING AND RELATED CURRICULA. MANY ILLUSTRATIVE EXAMPLES AND EXERCISES ARE PROVIDED IN ORDER TO FACILITATE CLASSROOM TEACHING. A FAMILIARITY IS ASSUMED TO THE OUTLIER DETECTION PROBLEM AND ALSO TO GENERIC PROBLEM OF ENSEMBLE ANALYSIS IN CLASSIFICATION. THIS IS BECAUSE MANY OF THE ENSEMBLE METHODS DISCUSSED IN THIS BOOK ARE ADAPTATIONS FROM THEIR COUNTERPARTS IN THE CLASSIFICATION DOMAIN. SOME TECHNIQUES EXPLAINED IN THIS BOOK, SUCH AS WAGGING, RANDOMIZED FEATURE WEIGHTING, AND GEOMETRIC SUBSAMPLING, PROVIDE NEW INSIGHTS THAT ARE NOT AVAILABLE ELSEWHERE. ALSO INCLUDED IS AN ANALYSIS OF THE PERFORMANCE OF VARIOUS TYPES OF BASE DETECTORS AND THEIR RELATIVE EFFECTIVENESS. THE BOOK IS VALUABLE FOR RESEARCHERS AND PRACTITIONERS FOR LEVERAGING ENSEMBLE METHODS INTO OPTIMAL ALGORITHMIC DESIGN.

ENCYCLOPEDIA OF MACHINE LEARNING - CLAUDE SAMMUT 2011-03-28

THIS COMPREHENSIVE ENCYCLOPEDIA, IN A-Z FORMAT, PROVIDES EASY ACCESS TO RELEVANT INFORMATION FOR THOSE SEEKING ENTRY INTO ANY ASPECT WITHIN THE BROAD FIELD OF MACHINE LEARNING. MOST OF THE ENTRIES IN THIS PREEMINENT WORK INCLUDE USEFUL LITERATURE REFERENCES.

MANAGING AND MINING SENSOR DATA - CHARU C. AGGARWAL 2013-01-15

ADVANCES IN HARDWARE TECHNOLOGY HAVE LEAD TO AN ABILITY TO COLLECT DATA WITH THE USE OF A VARIETY OF SENSOR TECHNOLOGIES. IN PARTICULAR SENSOR NOTES HAVE BECOME CHEAPER AND MORE EFFICIENT, AND HAVE EVEN BEEN INTEGRATED INTO DAY-TO-DAY DEVICES OF USE, SUCH AS MOBILE PHONES. THIS HAS LEAD TO A MUCH LARGER SCALE OF APPLICABILITY AND MINING OF SENSOR DATA SETS. THE HUMAN-CENTRIC ASPECT OF SENSOR DATA HAS CREATED TREMENDOUS OPPORTUNITIES IN INTEGRATING SOCIAL ASPECTS OF SENSOR DATA COLLECTION INTO THE MINING PROCESS. MANAGING AND MINING SENSOR DATA IS A CONTRIBUTED VOLUME BY PROMINENT LEADERS IN THIS FIELD, TARGETING ADVANCED-

LEVEL STUDENTS IN COMPUTER SCIENCE AS A SECONDARY TEXT BOOK OR REFERENCE. PRACTITIONERS AND RESEARCHERS WORKING IN THIS FIELD WILL ALSO FIND THIS BOOK USEFUL.

RECOMMENDER SYSTEMS - CHARU C. AGGARWAL 2016-03-28

THIS BOOK COMPREHENSIVELY COVERS THE TOPIC OF RECOMMENDER SYSTEMS, WHICH PROVIDE PERSONALIZED RECOMMENDATIONS OF PRODUCTS OR SERVICES TO USERS BASED ON THEIR PREVIOUS SEARCHES OR PURCHASES. RECOMMENDER SYSTEM METHODS HAVE BEEN ADAPTED TO DIVERSE APPLICATIONS INCLUDING QUERY LOG MINING, SOCIAL NETWORKING, NEWS RECOMMENDATIONS, AND COMPUTATIONAL ADVERTISING. THIS BOOK SYNTHESIZES BOTH FUNDAMENTAL AND ADVANCED TOPICS OF A RESEARCH AREA THAT HAS NOW REACHED MATURITY. THE CHAPTERS OF THIS BOOK ARE ORGANIZED INTO THREE CATEGORIES: ALGORITHMS AND EVALUATION: THESE CHAPTERS DISCUSS THE FUNDAMENTAL ALGORITHMS IN RECOMMENDER SYSTEMS, INCLUDING COLLABORATIVE FILTERING METHODS, CONTENT-BASED METHODS, KNOWLEDGE-BASED METHODS, ENSEMBLE-BASED METHODS, AND EVALUATION. RECOMMENDATIONS IN SPECIFIC DOMAINS AND CONTEXTS: THE CONTEXT OF A RECOMMENDATION CAN BE VIEWED AS IMPORTANT SIDE INFORMATION THAT AFFECTS THE RECOMMENDATION GOALS. DIFFERENT TYPES OF CONTEXT SUCH AS TEMPORAL DATA, SPATIAL DATA, SOCIAL DATA, TAGGING DATA, AND TRUSTWORTHINESS ARE EXPLORED. ADVANCED TOPICS AND APPLICATIONS: VARIOUS ROBUSTNESS ASPECTS OF RECOMMENDER SYSTEMS, SUCH AS SHILLING SYSTEMS, ATTACK MODELS, AND THEIR DEFENSES ARE DISCUSSED. IN ADDITION, RECENT TOPICS, SUCH AS LEARNING TO RANK, MULTI-ARMED BANDITS, GROUP SYSTEMS, MULTI-CRITERIA SYSTEMS, AND ACTIVE LEARNING SYSTEMS, ARE INTRODUCED TOGETHER WITH APPLICATIONS. ALTHOUGH THIS BOOK PRIMARILY SERVES AS A TEXTBOOK, IT WILL ALSO APPEAL TO INDUSTRIAL PRACTITIONERS AND RESEARCHERS DUE TO ITS FOCUS ON APPLICATIONS AND REFERENCES. NUMEROUS EXAMPLES AND EXERCISES HAVE BEEN PROVIDED, AND A SOLUTION MANUAL IS AVAILABLE FOR INSTRUCTORS.

WEB DATA MINING - BING LIU 2011-06-25

LIU HAS WRITTEN A COMPREHENSIVE TEXT ON WEB MINING, WHICH CONSISTS OF TWO PARTS. THE FIRST PART COVERS THE DATA MINING AND MACHINE LEARNING FOUNDATIONS, WHERE ALL THE ESSENTIAL CONCEPTS AND ALGORITHMS OF DATA MINING AND MACHINE LEARNING ARE PRESENTED. THE SECOND PART COVERS THE KEY TOPICS OF WEB MINING, WHERE WEB CRAWLING, SEARCH, SOCIAL NETWORK ANALYSIS, STRUCTURED DATA EXTRACTION, INFORMATION INTEGRATION, OPINION MINING AND SENTIMENT ANALYSIS, WEB USAGE MINING, QUERY LOG MINING, COMPUTATIONAL ADVERTISING, AND RECOMMENDER SYSTEMS ARE ALL TREATED BOTH IN BREADTH AND IN DEPTH. HIS BOOK THUS BRINGS ALL THE RELATED CONCEPTS AND ALGORITHMS TOGETHER TO FORM AN AUTHORITATIVE AND COHERENT TEXT. THE BOOK OFFERS A RICH BLEND OF THEORY AND PRACTICE. IT IS SUITABLE FOR STUDENTS, RESEARCHERS AND PRACTITIONERS INTERESTED IN WEB MINING AND DATA MINING BOTH AS A LEARNING TEXT AND AS A REFERENCE BOOK. PROFESSORS CAN READILY USE IT FOR CLASSES ON DATA MINING, WEB MINING, AND TEXT MINING. ADDITIONAL TEACHING MATERIALS SUCH AS

LECTURE SLIDES, DATASETS, AND IMPLEMENTED ALGORITHMS ARE AVAILABLE ONLINE.