An Overview Of Outlier Detection Methods

New Methods in Outlier Detection

An Overview Of Outlier Detection Methods

Utilize this easy-to-follow beginner's guide to understand how deep learning can be applied to the task of anomaly detection. Using Keras and PyTorch in Python, the book focuses on how various deep learning models can be applied to semi-supervised and unsupervised anomaly detection tasks. This book begins with an explanation of what anomaly detection is, what it is used for, and its importance. After covering statistical and traditional machine learning methods for anomaly detection using Scikit-Learn in Python, the book then provides an introduction to deep learning with details on how to build and train a deep learning model in both Keras and PyTorch before shifting the focus to applications of the following deep learning models to anomaly detection: various types of Autoencoders, Restricted Boltzmann Machines, RNNs & LSTMs, and Temporal Convolutional Networks. The book explores unsupervised and semi-supervised anomaly detection along with the basics of time series-based anomaly detection. By the end of the book you will have a thorough understanding of the basic task of anomaly detection as well as an assortment of methods to approach anomaly detection, ranging from traditional methods to deep learning. Additionally, you are introduced to Scikit-Learn and are able to create deep learning models in Keras and PyTorch. What You Will Learn Understand what anomaly detection is and why it is important in today's world Become familiar with statistical and traditional machine learning approaches to anomaly detection using Scikit-Learn Know the basics of deep learning in Python using Keras and PyTorch Be aware of basic data science concepts for measuring a model's performance: understand what AUC is, what precision and recall mean, and more Apply deep learning to semi-supervised and unsupervised anomaly detection Who This Book Is For Data scientists and machine learning engineers interested in learning the basics of deep learning applications in anomaly detection

Anomaly Detection Principles and Algorithms

WILEY-INTERSCIENCE PAPERBACK SERIES The Wiley-Interscience Paperback Series consists of selected books that have been made more accessible to consumers in an effort to increase global appeal and general circulation. With these new unabridged softcover volumes, Wiley hopes to extend the lives of these works by making them available to future generations of statisticians, mathematicians, and scientists. "The writing style is clear and informal, and much of the discussion is oriented to application. In short, the book is a keeper." -Mathematical Geology "I would highly recommend the addition of this book to the libraries of both students and professionals. It is a useful textbook for the graduate student, because it emphasizes both the philosophy and practice of robustness in regression settings, and it provides excellent examples of precise, logical proofs of theorems. Even for those who are familiar with robustness, the book will be a good reference because it consolidates the research in high-breakdown affine equivariant estimators and includes an extensive bibliography in robust regression, outlier diagnostics, and related methods. The aim of this book, the authors tell us, is to make robust regression available for everyday
Outliers are anomalous and interesting objects that are notably different from the rest of the data. The outlier detection task has sometimes been considered as removing noise from the data. However, it is usually the significantly interesting deviations that are of most interest. Different outlier detection techniques work with various data formats. The outlier detection process needs to be sensitive to the nature of the underlying data. Most of the previous work on outlier detection was designed for propositional data. This dissertation focuses on developing outlier detection methods for structured data, more specifically object-relational data. Object-relational data can be viewed as a heterogeneous network with different classes of objects and links. We develop two new approaches to unsupervised outlier detection; both approaches leverage the statistical information obtained from a statistical-relational model. The first method develops a propositionalization approach to summarize information from object-relational data in a single data table. We use Markov Logic Network (MLN) structure learning to construct the features for the single data table and to mitigate the loss of information that usually happens when features are generated by manual aggregation. By using propositionalization as a pipeline, we can apply many previous outlier detection methods that were designed for single-table data. Our second outlier detection method ranks the objects as potential outliers in an object-oriented data model. Our key idea is to compare the feature distribution of a potential outlier object with the feature distribution of the object's class. We introduce a novel distribution divergence concept that is suitable for outlier detection. Our methods are validated on synthetic datasets and on real-world datasets about soccer matches and movies.

Outlier Ensembles

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.
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.

**Advances in Robotics and Automatic Control: Reviews, Vol. 1**

This comprehensive book gives an overview of how cognitive systems and artificial intelligence (AI) can be used in electronic warfare (EW). Readers will learn how EW systems respond more quickly and effectively to battlefield conditions where sophisticated radars and spectrum congestion put a high priority on EW systems that can characterize and classify novel waveforms, discern intent, and devise and test countermeasures. Specific techniques are covered for optimizing a cognitive EW system as well as evaluating its ability to learn new information in real time. The book presents AI for electronic support (ES), including characterization, classification, patterns of life, and intent recognition. Optimization techniques, including temporal tradeoffs and distributed optimization challenges are also discussed. The issues concerning real-time in-mission machine learning and suggests some approaches to address this important challenge are presented and described. The book covers electronic battle management, data management, and knowledge sharing. Evaluation approaches, including how to show that a machine learning system can learn how to handle novel environments, are also discussed. Written by experts with first-hand experience in AI-based EW, this is the first book on in-mission real-time learning and optimization.

**ECAI 2020**

This book provides a readable and elegant presentation of the principles of anomaly detection, providing an easy introduction for newcomers to the field. A large number of algorithms are succinctly described, along with a presentation of their strengths and weaknesses. The authors also cover algorithms that address different kinds of problems of interest with single and multiple time series data and multi-dimensional data. New ensemble anomaly detection algorithms are described, utilizing the benefits provided by diverse algorithms, each of which work well on some kinds of data. With advancements in technology and the extensive use of the internet as a medium for communications and commerce, there has been a tremendous increase in the threats faced by individuals and organizations from attackers and criminal entities. Variations in the observable behaviors of individuals (from others and from their own past behaviors) have been found to be useful in predicting potential problems of various kinds. Hence computer scientists and statisticians have been conducting research on automatically identifying anomalies in large datasets. This book will primarily target practitioners and researchers who are newcomers to the area of modern anomaly detection techniques. Advanced-level students in computer science will also find this book helpful with their studies.

**Network Traffic Anomaly Detection and Prevention**

This book enriches unsupervised outlier detection research by proposing several new distance-based and density-based outlier scores in a k-nearest neighbors' setting. The
respective chapters highlight the latest developments in k-nearest neighbor-based outlier detection research and cover such topics as our present understanding of unsupervised outlier detection in general; distance-based and density-based outlier detection in particular; and the applications of the latest findings to boundary point detection and novel object detection. The book also offers a new perspective on bridging the gap between k-nearest neighbor-based outlier detection and clustering-based outlier detection, laying the groundwork for future advances in unsupervised outlier detection research. The authors hope the algorithms and applications proposed here will serve as valuable resources for outlier detection researchers for years to come.

**Outlier Detection Algorithms**

Heuristic Search is an important sub-discipline of optimization theory and finds applications in a vast variety of fields, including life science and engineering. Search methods have been useful in solving tough engineering-oriented problems that either could not be solved any other way or solutions take a very long time to be computed. This book explores a variety of applications for search methods and techniques in different fields of electrical engineering. By organizing relevant results and applications, this book will serve as a useful resource for students, researchers and practitioners to further exploit the potential of search methods in solving hard optimization problems that arise in advanced engineering technologies, such as image and video processing issues, detection and resource allocation in telecommunication systems, security and harmonic reduction in power generation systems, as well as redundancy optimization problem and search-fuzzy learning mechanisms in industrial applications.

**Business Intelligence: An overview**

This indispensable text/reference presents a comprehensive overview on the detection and prevention of anomalies in computer network traffic, from coverage of the fundamental theoretical concepts to in-depth analysis of systems and methods. Readers will benefit from invaluable practical guidance on how to design an intrusion detection technique and incorporate it into a system, as well as on how to analyze and correlate alerts without prior information. Topics and features: introduces the essentials of traffic management in high speed networks, detailing types of anomalies, network vulnerabilities, and a taxonomy of network attacks; describes a systematic approach to generating large network intrusion datasets, and reviews existing synthetic, benchmark, and real-life datasets; provides a detailed study of network anomaly detection techniques and systems under six different categories: statistical, classification, knowledge-base, cluster and outlier detection, soft computing, and combination learners; examines alert management and anomaly prevention techniques, including alert preprocessing, alert correlation, and alert post-processing; presents a hands-on approach to developing network traffic monitoring and analysis tools, together with a survey of existing tools; discusses various evaluation criteria and metrics, covering issues of accuracy, performance, completeness, timeliness, reliability, and quality; reviews open issues and challenges in network traffic anomaly detection and prevention. This informative work is ideal for graduate and advanced undergraduate students interested in network security and privacy, intrusion detection systems, and data mining in security. Researchers and practitioners specializing in network security will also find the book to be a useful reference.

**Data Mining: Concepts, Methodologies, Tools, and Applications**

**Outlier Detection--techniques and Applications**

**New Developments in Unsupervised Outlier Detection**

This handbook provides comprehensive knowledge and includes an overview of the current state-of-the-art of Big Data Privacy, with chapters written by international world leaders from academia and industry working in this field. The first part of this book offers a review of security challenges in critical infrastructure and offers methods that utilize artificial intelligence (AI) techniques to overcome these issues. It then focuses on big data security and privacy issues in relation to developments in the Industry 4.0. Internet of Things (IoT) devices are becoming a major source of security and privacy concern in big data platforms. Multiple solutions that leverage machine learning for addressing security
and privacy issues in IoT environments are also discussed in this handbook. The second part of 
this handbook is focused on privacy and security issues in different layers of big data 
systems. It discusses about methods for evaluating security and privacy of big data systems 
on network, application and physical layers. This handbook elaborates on existing methods 
to use data analytic and AI techniques at different layers of big data platforms to 
identify privacy and security attacks. The final part of this handbook is focused on 
analyzing cyber threats applicable to the big data environments. It offers an in-depth 
review of attacks applicable to big data platforms in smart grids, smart farming, FinTech, 
and health sectors. Multiple solutions are presented to detect, prevent and analyze cyber- 
attacks and assess the impact of malicious payloads to those environments. This handbook 
provides information for security and privacy experts in most areas of big data including; 
FinTech, Industry 4.0, Internet of Things, Smart Grids, Smart Farming and more. Experts 
working in big data, privacy, security, forensics, malware analysis, machine learning and 
data analysts will find this handbook useful as a reference. Researchers and advanced-level 
computer science students focused on computer systems, Internet of Things, Smart Grid, 
Smart Farming, Industry 4.0 and network analysts will also find this handbook useful as a 
reference.

Search Algorithms for Engineering Optimization

This book constitutes the refereed proceedings of the International Standard Conference on 
Trustworthy Distributed Computing and Services, ISCTCS 2013, held in Beijing, China, in 
November 2013. The 49 revised full papers presented were carefully reviewed and selected 
from 267 papers. The topics covered are trustworthy infrastructure; security, survivability 
and fault tolerance; standards, evaluation and certification; trustworthiness of services.

Data Quality for Analytics Using SAS

This book presents the proceedings of the 24th European Conference on Artificial 
Intelligence (ECAI 2020), held in Santiago de Compostela, Spain, from 29 August to 8 
September 2020. The conference was postponed from June, and much of it conducted online due 
to the COVID-19 restrictions. The conference is one of the principal occasions for 
researchers and practitioners of AI to meet and discuss the latest trends and challenges in 
all fields of AI and to demonstrate innovative applications and uses of advanced AI 
technology. The book also includes the proceedings of the 10th Conference on Prestigious 
Applications of Artificial Intelligence (PAIS 2020) held at the same time. A record number of 
more than 1,700 submissions was received for ECAI 2020, of which 1,443 were reviewed. Of 
these, 361 full-papers and 36 highlight papers were accepted (an acceptance rate of 25% for 
full-papers and 45% for highlight papers). The book is divided into three sections: ECAI 
full papers; ECAI highlight papers; and PAIS papers. The topics of these papers cover all 
aspects of AI, including Agent-based and Multi-agent Systems; Computational Intelligence; 
Constraints and Satisfiability; Games and Virtual Environments; Heuristic Search; Human 
Aspects in AI; Information Retrieval and Filtering; Knowledge Representation and Reasoning; 
Machine Learning; Multidisciplinary Topics and Applications; Natural Language Processing; 
Planning and Scheduling; Robotics; Safe, Explainable, and Trustworthy AI; Semantic 
Technologies; Uncertainty in AI; and Vision. The book will be of interest to all those 
whose work involves the use of AI technology.

Trustworthy Computing and Services

With the rapid rise in the ubiquity and sophistication of Internet technology and the 
accompanying growth in the number of network attacks, network intrusion detection has 
become increasingly important. Anomaly-based network intrusion detection refers to finding 
exceptional or nonconforming patterns in network traffic data compared to normal behavior. 
Finding these anomalies has extensive applications in areas such as cyber security, credit 
card and insurance fraud detection, and military surveillance for enemy activities. Network 
Anomaly Detection: A Machine Learning Perspective presents machine learning techniques in 
depth to help you more effectively detect and counter network intrusion. In this book, 
you’ll learn about: Network anomalies and vulnerabilities at various layers The pros and 
cons of various machine learning techniques and algorithms A taxonomy of attacks based on 
their characteristics and behavior Feature selection algorithms How to assess the accuracy, 
performance, completeness, timeliness, stability, interoperability, reliability, and other 
dynamic aspects of a network anomaly detection system Practical tools for launching 
attacks, capturing packet or flow traffic, extracting features, detecting attacks, and 
evaluating detection performance Important unresolved issues and research challenges that 
need to be overcome to provide better protection for networks Examining numerous attacks in 
detail, the authors look at the tools that intruders use and show how to use this knowledge
Outlier detection is a fundamental task that is used in numerous data analytic applications. It tackles the problem of identifying rare or atypical points that widely diverge from the general behavior or model of the data. The process of detecting outliers and subsequently using them for data analysis relies on the underlying application. For example, outlier detection can be employed as a preprocessing step to clean the data set from erroneous measurements and noisy data points. On the other hand, it can also be used to isolate suspicious or interesting patterns in the data. Examples include fraud detection, customer relationship management, network intrusion, clinical diagnosis, and biological data analysis. Although many successful algorithms have been developed for outlier detection, several challenges have haunted researchers and practitioners for decades. The first one is limited algorithm scalability. Due to the fast evolution of World Wide Web, the collected data can easily reach terabyte- or even petabyte- scale. Most existing approaches, ranging from statistical methods to geometric methods, and from density-based approaches to information theory based approaches, suffer from limited scalability and do not work well on large scale data. The second one is to detect outliers in the irregular, dynamic semi-structured data such as trees and graphs. There have been some research on finding outliers from the graphs. What are the definitions for meaningful outliers in the graph context? How can we detect them accurately and efficiently? The third challenge is to build a unified and modular detection system which provides researchers a complete toolbox for outlier detection tasks. Our research aims at designing the next-generation outlier detection algorithms that tackle the above three challenges. To achieve better scalability, we have done an extensive empirical study on different optimization techniques for distance-based outlier detection. Also, we proposed an ranking scheme driven by the Locality Sensitive Hashing (LSH), which finds all outliers by only visiting a small portion of the data (10%). Find similar points of each point, or all pair similarity search, is the key operation for many distance-based, density-based and cluster-based outliers. We optimized this fundamental kernel in metric space on MapReduce platform, and scaled the algorithm to hundreds of machines and solved the inadequate memory issue. For semi-structured outlier detection, we first designed a clustering-based algorithm, and a generic clustering algorithm for sets/multisets, trees and graphs. We also studied a concrete detection application on the semi-structured knowledge base, and found more than one million anomalies. Finally, we integrated our work seamlessly into a detection framework, which accepts different types of data. Users also enjoy the freedom of choosing and comparing different algorithms.

Emergent Computation

This indispensable text/reference presents a comprehensive overview on the detection and prevention of anomalies in computer network traffic, from coverage of the fundamental theoretical concepts to in-depth analysis of systems and methods. Readers will benefit from invaluable practical guidance on how to design an intrusion detection technique and incorporate it into a system, as well as on how to analyze and correlate alerts without prior information. Topics and features: introduces the essentials of traffic management in high speed networks, detailing types of anomalies, network vulnerabilities, and a taxonomy of network attacks; describes a systematic approach to generating large network intrusion datasets, and reviews existing synthetic, benchmark, and real-life datasets; provides a detailed study of network anomaly detection techniques and systems under six different categories: statistical, classification, knowledge-base, cluster and outlier detection, soft computing, and combination learners; examines alert management and anomaly prevention techniques, including alert preprocessing, alert correlation, and alert post-processing; presents a hands-on approach to developing network traffic monitoring and analysis tools, together with a survey of existing tools; discusses various evaluation criteria and metrics, covering issues of accuracy, performance, completeness, timeliness, reliability, and quality; reviews open issues and challenges in network traffic anomaly detection and prevention. This informative work is ideal for graduate and advanced undergraduate students interested in network security and privacy, intrusion detection systems, and data mining in security. Researchers and practitioners specializing in network security will also find the book to be a useful reference.
Analytics offers many capabilities and options to measure and improve data quality, and SAS is perfectly suited to these tasks. Gerhard Svolba's Data Quality for Analytics Using SAS focuses on selecting the right data sources and ensuring data quantity, relevancy, and completeness. The book is made up of three parts. The first part, which is conceptual, defines data quality and contains text, definitions, explanations, and examples. The second part shows how the data quality status can be profiled and the ways that data quality can be improved with analytical methods. The final part details the consequences of poor data quality for predictive modeling and time series forecasting. With this book you will learn how you can use SAS to perform advanced profiling of data quality status and how SAS can help improve your data quality. This book is part of the SAS Press program.

Outlier detection has been studied extensively in data mining. However, as the emergence of huge data sets in real-life applications nowadays, outlier detection faces a series of new challenges. Many traditional outlier detection techniques do not work well in such an environment. Therefore, developing up-to-date outlier detection methods becomes urgent tasks. In this thesis, we propose several new methods for detecting different kinds of outliers in high-dimensional data sets from two different perspectives, namely, detecting the outlying aspects of a data object and detecting outlying data objects of a data set. Specifically, for detecting the outlying aspects of a data object, we propose the problems of mining outlying aspects and mining contrast subspaces; for detecting outlying data objects of a data set, we propose the problems of mining contextual outliers and mining Markov blanket based outliers. We develop efficient and scalable algorithms to tackle the computational challenges. We also conduct comprehensive empirical studies on real and synthetic data sets to verify the effectiveness of the proposed outlier detection techniques and the efficiency of our algorithms.

Model-based Outlier Detection for Object-Relational Data

This book constitutes the refereed proceedings of the Third International Conference on Information Systems, Technology and Management, ICISTM 2009, held in Ghaziabad, India, in March 2009 The 30 revised full papers presented together with 4 keynote papers were carefully reviewed and selected from 79 submissions. The papers are organized in topical sections on storage and retrieval systems; data mining and classification; managing digital goods and services; scheduling and distributed systems; advances in software engineering; case studies in information management; algorithms and workflows; authentication and detection systems; recommendation and negotiation; secure and multimedia systems; as well as 14 extended poster abstracts.

Cognitive Electronic Warfare: An Artificial Intelligence Approach

This book and its companion volumes, LNCS volumes 9140, 9141 and 9142, constitute the proceedings of the 6th International Conference on Swarm Intelligence, ICSI 2015 held in conjunction with the Second BRICS Congress on Computational Intelligence, CCI 2015, held in Beijing, China in June 2015. The 161 revised full papers presented were carefully reviewed and selected from 294 submissions. The papers are organized in 28 cohesive sections covering all major topics of swarm intelligence and computational intelligence research and development, such as novel swarm-based optimization algorithms and applications; particle swarm optimization; ant colony optimization; artificial bee colony algorithms; evolutionary and genetic algorithms; differential evolution; brain storm optimization algorithm; biogeography based optimization; cuckoo search; hybrid methods; multi-objective optimization; multi-agent systems and swarm robotics; Neural networks and fuzzy methods; data mining approaches; information security; automation control; combinatorial optimization algorithms; scheduling and path planning; machine learning; blind sources separation; swarm interaction behavior; parameters and system optimization; neural networks; evolutionary and genetic algorithms; fuzzy systems; forecasting algorithms; classification; tracking analysis; simulation; image and texture analysis; dimension reduction; system optimization; segmentation and detection system; machine translation; virtual management and disaster analysis.
detection research and cover such topics as our present understanding of unsupervised outlier detection in general; distance-based and density-based outlier detection in particular; and the applications of the latest findings to boundary point detection and novel object detection. The book also offers a new perspective on bridging the gap between k-nearest neighbor-based outlier detection and clustering-based outlier detection, laying the groundwork for future advances in unsupervised outlier detection research. The authors hope the algorithms and applications proposed here will serve as valuable resources for outlier detection researchers for years to come.

Next Generation Outlier Detection

Network Anomaly Detection

As the age of Big Data emerges, it becomes necessary to take the five dimensions of Big Data—volume, variety, velocity, volatility, and veracity—and focus these dimensions towards one critical emphasis—value. The Encyclopedia of Business Analytics and Optimization confronts the challenges of information retrieval in the age of Big Data by exploring recent advances in the areas of knowledge management, data visualization, interdisciplinary communication, and others. Through its critical approach and practical application, this book will be a must-have reference for any professional, leader, analyst, or manager interested in making the most of the knowledge resources at their disposal.

Outlier Detection Methodologies for Alternative Data Sources

Robust Regression and Outlier Detection

This book, drawing on recent literature, highlights several methodologies for the detection of outliers and explains how to apply them to solve several interesting real-life problems. The detection of objects that deviate from the norm in a data set is an essential task in data mining due to its significance in many contemporary applications. More specifically, the detection of fraud in e-commerce transactions and discovering anomalies in network data have become prominent tasks, given recent developments in the field of information and communication technologies and security. Accordingly, the book sheds light on specific state-of-the-art algorithmic approaches such as the community-based analysis of networks and characterization of temporal outliers present in dynamic networks. It offers a valuable resource for young researchers working in data mining, helping them understand the technical depth of the outlier detection problem and devise innovative solutions to address related challenges.

Machine Learning and Intelligent Communications

This book constitutes the refereed conference proceedings of the 16th IFIP WG 6.11 Conference on e-Business, e-Services and e-Society, I3E 2017, held in Delhi, India, in November 2017. The 45 revised full papers presented were carefully reviewed and selected from 92 submissions. They are organized in the following topical sections: Adoption of Smart Services; Assessment of ICT Enabled Smart Initiatives; Analytics for Smart Governance; Social Media and Web 3.0 for Smartness; and Smart Solutions for the Future.

Fundamentals and Applications of AI: An Interdisciplinary Perspective

New Developments in Unsupervised Outlier Detection

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.

**Outlier Detection for Temporal Data**

This book, drawing on recent literature, highlights several methodologies for the detection of outliers and explains how to apply them to solve several interesting real-life problems. The detection of objects that deviate from the norm in a data set is an essential task in data mining due to its significance in many contemporary applications. More specifically, the detection of fraud in e-commerce transactions and discovering anomalies in network data have become prominent tasks, given recent developments in the field of information and communication technologies and security. Accordingly, the book sheds light on specific state-of-the-art algorithmic approaches such as the community-based analysis of networks and characterization of temporal outliers present in dynamic networks. It offers a valuable resource for young researchers working in data mining, helping them understand the technical depth of the outlier detection problem and devise innovative solutions to address related challenges.

**Advances in Swarm and Computational Intelligence**

Business organizations develop strategies and set targets which focus on maximizing profit, reduce cost, improving customer satisfaction & retention and operational performance. In order to achieve the set targets, organizations need to continuously monitor status of organizational performance. Organizations need to collect, store, organize, transform the data to know the current status of set targets. Business Intelligence tools help the organizations to draw meaningful and actionable insights from the raw data in achieving the set targets. Business Intelligence tools help the organizations to answer questions such as where the organization stands in terms of profitability, growth status, brand & market position and market segment. Business intelligence tools focuses mainly on the past or current data and try to explore the hidden insight from the data. Business intelligence tools include querying, reporting, online analytics and data visualization tools which help the business decision makers to arrive at informed decision about the impact and status of their strategies. This book starts with the introduction of business intelligence concepts, components of business intelligence system, business intelligence tools used for querying, reporting and visualization of data. It provides an overview of the data visualization and data mining methods like classification, clustering and regression methods using R open source software. Book also covers some of the basic descriptive and inferential statistical tools. It focuses on both managerial side and technological side of BI. Vinaiitheerthan Renganathan www.vinaiitheerthan.com/book.php

**Outlier Analysis**

This book is dedicated to Professor Selim G. Akl to honour his groundbreaking research achievements in computer science over four decades. The book is an intellectually stimulating excursion into emergent computing paradigms, architectures and implementations. World top experts in computer science, engineering and mathematics overview exciting and intriguing topics of musical rhythms generation algorithms, analyse the computational power of random walks, dispelling a myth of computational universality, computability and complexity at the microscopic level of synchronous computation, descriptional complexity of error detection, quantum cryptography, context-free parallel communicating grammar systems, fault tolerance of hypercubes, finite automata theory of bulk-synchronous parallel computing, dealing with silent data corruptions in high-performance computing, parallel sorting on graphics processing units, mining for functional dependencies in relational databases, cellular automata optimisation of wireless sensors networks, connectivity preserving network transformers, constrained resource networks, vague computing, parallel evolutionary optimisation, emergent behaviour in multi-agent systems, vehicular clouds, epigenetic drug discovery, dimensionality reduction for intrusion detection systems, physical maze solvers, computer chess, parallel algorithms to string alignment, detection of community structure. The book is a unique combination of vibrant essays which inspires scientists and engineers to exploit natural phenomena in designs of computing architectures of the future.
Outlier Detection Methods

Outlier Detection: Techniques and Applications

This two volume set constitutes the refereed post-conference proceedings of the Second International Conference on Machine Learning and Intelligent Communications, MLICOM 2017, held in Weihai, China, in August 2017. The 143 revised full papers were carefully selected from 225 submissions. The papers are organized thematically in machine learning, intelligent positioning and navigation, intelligent multimedia processing and security, intelligent wireless mobile network and security, cognitive radio and intelligent networking, intelligent internet of things, intelligent satellite communications and networking, intelligent remote sensing, visual computing and three-dimensional modeling, green communication and intelligent networking, intelligent ad-hoc and sensor networks, intelligent resource allocation in wireless and cloud networks, intelligent signal processing in wireless and optical communications, intelligent radar signal processing, intelligent cooperative communications and networking.

Beginning Anomaly Detection Using Python-Based Deep Learning

Data mining continues to be an emerging interdisciplinary field that offers the ability to extract information from an existing data set and translate that knowledge for end-users into an understandable way. Data Mining: Concepts, Methodologies, Tools, and Applications is a comprehensive collection of research on the latest advancements and developments of data mining and how it fits into the current technological world.

Detecting Interesting Differences

Outlier, abnormal or unusual observation can be defined as an observation that lies outside the overall pattern of a distribution. Diagnostic methods for identifying a single outlier or influential observation in a linear regression model are relatively simple from both analytical and computational points of view. However, if the data set contains more than one outlier, which is likely to be the case in most data sets, the problem of identifying such observations becomes more difficult because of the masking and swamping effects. A GA was allowed simultaneous detection of outliers in data sets. Thus, this method is to overcome the problems of masking and swamping effects. It is derived additional penalized value of information criteria for Akaike Information Criterion (AIC) and Information Complexity Criterion (ICOMP) and named as AIC' and ICOMP' respectively in this study. The numerical example and simulation results clearly show a much improved performance of the proposed approach in comparison to existing method especially followed by applying the ICOMP' approach in order to accurately (robustly) detect the outliers.

Network Traffic Anomaly Detection and Prevention

This book provides a readable and elegant presentation of the principles of anomaly detection, providing an easy introduction for newcomers to the field. A large number of algorithms are succinctly described, along with a presentation of their strengths and weaknesses. The authors also cover algorithms that address different kinds of problems of interest with single and multiple time series data and multi-dimensional data. New ensemble anomaly detection algorithms are described, utilizing the benefits provided by diverse algorithms, each of which work well on some kinds of data. With advancements in technology and the extensive use of the internet as a medium for communications and commerce, there has been a tremendous increase in the threats faced by individuals and organizations from attackers and criminal entities. Variations in the observable behaviors of individuals (from others and from their own past behaviors) have been found to be useful in predicting potential problems of various kinds. Hence computer scientists and statisticians have been conducting research on automatically identifying anomalies in large datasets. This book will primarily target practitioners and researchers who are newcomers to the area of modern anomaly detection techniques. Advanced-level students in computer science will also find this book helpful with their studies.

Copyright code: feadb3a2995400270d473eb71245e05

Copyright: boombox.pexsoftware.com

Page 10/10