Kindle File Format K Nearest Neighbor Algorithm For Classification

Right here, we have countless book k nearest neighbor algorithm for classification and collections to check out. We additionally have the funds for variant types and then type of the books to browse. The normal book, fiction, history, novel, scientific research, as well as various further sorts of books are readily to hand here.

As this k nearest neighbor algorithm for classification, it ends up bodily one of the favored book k nearest neighbor algorithm for classification collections that we have. This is why you remain in the best website to look the incredible book to have.

basic design patterns, optimization techniques, and data mining and

Dimensionality Reduction with Unsupervised Nearest Neighbors - Oliver Kramer - 2013-05-30
This book is devoted to a novel approach for dimensionality reduction based on the famous nearest neighbor method that is a powerful classification and regression approach. It starts with an introduction to machine learning concepts and a real-world application from the energy domain. Then, unsupervised nearest neighbors (UNN) is introduced as efficient iterative method for dimensionality reduction. Various UNN models are developed step by step, reaching from a simple iterative strategy for discrete latent spaces to a stochastic kernel-based algorithm for learning submanifolds with independent parameterizations. Extensions that allow the embedding of incomplete and noisy patterns are introduced. Various optimization approaches are compared, from evolutionary to swarm-based heuristics. Experimental comparisons to related methodologies taking into account artificial test data sets and also real-world data demonstrate the behavior of UNN in practical scenarios. The book contains numerous color figures to illustrate the introduced concepts and to highlight the experimental results.

Dimensionality Reduction with Unsupervised Nearest Neighbors - Oliver Kramer - 2013-05-30
This book is devoted to a novel approach for dimensionality reduction based on the famous nearest neighbor method that is a powerful classification and regression approach. It starts with an introduction to machine learning concepts and a real-world application from the energy domain. Then, unsupervised nearest neighbors (UNN) is introduced as efficient iterative method for dimensionality reduction. Various UNN models are developed step by step, reaching from a simple iterative strategy for discrete latent spaces to a stochastic kernel-based algorithm for learning submanifolds with independent parameterizations. Extensions that allow the embedding of incomplete and noisy patterns are introduced. Various optimization approaches are compared, from evolutionary to swarm-based heuristics. Experimental comparisons to related methodologies taking into account artificial test data sets and also real-world data demonstrate the behavior of UNN in practical scenarios. The book contains numerous color figures to illustrate the introduced concepts and to highlight the experimental results.

Data Algorithms - Mahmoud Parsian - 2015-07-13
If you are ready to dive into the MapReduce framework for processing large datasets, this practical book takes you step by step through the algorithms and tools you need to build distributed MapReduce applications with Apache Hadoop or Apache Spark. Each chapter provides a recipe for solving a massive computational problem, such as building a recommendation system. You’ll learn how to implement the appropriate MapReduce solution with code that you can use in your projects. Dr. Mahmoud Parsian covers basic design patterns, optimization techniques, and data mining and machine learning solutions for problems in bioinformatics, genomics, statistics, and social network analysis. This book also includes an overview of MapReduce, Hadoop, and Spark. Topics include: Market basket analysis for a large set of transactions Data mining algorithms (K-means, KNN, and Naive Bayes) Using huge genomic data to sequence DNA and RNA Naive Bayes theorem and Markov chains for data and market prediction Recommendation algorithms and pairwise document similarity Linear regression, Cox regression, and Pearson correlation Allelic frequency and mining DNA Social network analysis (recommendation systems, counting triangles, sentiment analysis)

Data Algorithms - Mahmoud Parsian - 2015-07-13
If you are ready to dive into the MapReduce framework for processing large datasets, this practical book takes you step by step through the algorithms and tools you need to build distributed MapReduce applications with Apache Hadoop or Apache Spark. Each chapter provides a recipe for solving a massive computational problem, such as building a recommendation system. You’ll learn how to implement the appropriate MapReduce solution with code that you can use in your projects. Dr. Mahmoud Parsian covers basic design patterns, optimization techniques, and data mining and machine learning solutions for problems in bioinformatics, genomics, statistics, and social network analysis. This book also includes an overview of MapReduce, Hadoop, and Spark. Topics include: Market basket analysis for a large set of transactions Data mining algorithms (K-means, KNN, and Naive Bayes) Using huge genomic data to sequence DNA and RNA Naive Bayes theorem and Markov chains for data and market prediction Recommendation algorithms and pairwise document similarity Linear regression, Cox regression, and Pearson correlation Allelic frequency and mining DNA Social network analysis (recommendation systems, counting triangles, sentiment analysis)

Data Mining in Agriculture - Antonio Mucherino - 2009-09-22
Data Mining in Agriculture represents a comprehensive effort to provide graduate students and researchers with an analytical text on data mining techniques applied to agriculture and environmental related fields. This book presents both theoretical and practical insights with a focus on presenting the context of each data mining technique rather intuitively with ample concrete examples represented graphically and with algorithms written in MATLAB®.

Data Mining in Agriculture - Antonio Mucherino - 2009-09-22
Data Mining in Agriculture represents a comprehensive effort to provide graduate students and researchers with an analytical text on data mining techniques applied to agriculture and environmental related fields. This book presents both theoretical and practical insights with a focus on presenting the context of each data mining technique rather intuitively with ample concrete examples represented graphically and with algorithms written in MATLAB®.

Advances in Databases and Information Systems - Yannis Ioannidis - 2007-09-17
This book constitutes the refereed proceedings of the 11th East European Conference on Advances in Databases and Information Systems, ADBIS 2007, held in Varna, Bulgaria, in September/October 2007. The 23 revised papers presented together with three invited lectures were carefully reviewed and selected from 77 submissions. The papers address current research on database theory, development of advanced DBMS technologies, and their advanced applications.

Advances in Databases and Information Systems - Yannis Ioannidis - 2007-09-17
This book constitutes the refereed proceedings of the 11th East European Conference on Advances in Databases and Information Systems, ADBIS 2007, held in Varna, Bulgaria, in September/October 2007. The 23 revised papers presented together with three invited lectures were carefully reviewed and selected from 77 submissions. The papers address current research on database theory, development of advanced DBMS technologies, and their advanced applications.

Advanced Data Mining and Applications - Longbiao Cao - 2010-11-18
With the ever-growing power of generating, transmitting, and collecting huge amounts of data, information overloads now an imminent problem to mankind. The overwhelming demand for information processing is not just about a better understanding of data, but also a better usage of data in a timely fashion. Data mining, or knowledge discovery from databases, is proposed to gain insight into aspects of data and to help people make informed, sensible, and better decisions. At present, growing attention has been paid to the study, development, and application of data mining. As a result there is an urgent need for sophisticated techniques and tools that can handle new big data of mining, e. g., spatiotemporal mining, biomedical data mining, and mining on high-speed and time-variant data streams. The knowledge of data mining should also be expanded to new applications. The 6th International Conference on Advanced Data Mining and Applications (ADMA2010) aimed to bring together the experts and domain mining through the world. It provided a leading international forum for the dissemination of
predict a quantitative outcome value using linear regression and non-linear algorithms, software and systems, and different applied disciplines. The conference attracted over 34 different countries and areas. All full papers were peer reviewed by at least three members of the Program Committee composed of international experts in data mining and AI.

A total number of 118 papers were accepted for the conference. Amongst them, 63 papers were selected as regular papers and 55 papers were selected as short papers.

Advanced Data Mining and Applications - Longbing Cao - 2010-11-18

With the ever-growing power of generating, transmitting, and collecting huge amounts of data, information overload now menaces prominent mankind. The overwhelming demand for information processing is not just about a better understanding of data, but also a better usage of data in a timely fashion. Data mining, or knowledge discovery from databases, is proposed to gain insight into aspects of data and to help people make informed, sensible, and better decisions. At present, growing attention has been paid to the study, development, and application of data mining. As a result, there is an urgent need for a sophisticated theoretical framework and toolkit that can handle new questions of data mining, e.g., spatial data mining, biometrical data mining, and mining on high-speed and time-variant data streams. The knowledge of data mining should also be expanded to new applications. The 6th International Conference on Advanced Data Mining and Applications (ADMA2010) aimed to bring together experts on data mining throughout the world. It provided a leading international forum for the dissemination of original research results in advanced data mining techniques, applications, algorithms, software and systems, and different applied disciplines. The conference attracted 361 online submissions from 34 different countries and areas. All full papers were peer reviewed by at least three members of the Program Committee composed of international experts in data mining and AI. A total number of 118 papers were accepted for the conference. Amongst them, 63 papers were selected as regular papers and 55 papers were selected as short papers.

Integration Challenges for Analytics, Business Intelligence, and Data Mining - Azevedo, Ana - 2020-12-11

As technology continues to advance, it is critical for businesses to implement systems that can support the transformation of data into information that is crucial for the success of the company. Without the integration of data (both structured and unstructured) mining in business intelligence systems, invaluable knowledge is lost. However, there are currently many different models and approaches that must be explored to determine the best method of integration. Integration Challenges for Analytics, Business Intelligence, and Data Mining is a relevant academic book that provides empirical research findings on increasing the understanding of using data mining in the context of business intelligence and analytics systems. Covering topics that include big data, artificial intelligence, and decision making, this book is an ideal reference source for professionals working in the areas of data mining, business intelligence, and analytics; data scientists; IT specialists; managers; researchers; academicians; practitioners; and graduate students.

Integration Challenges for Analytics, Business Intelligence, and Data Mining - Azevedo, Ana - 2020-12-11

As technology continues to advance, it is critical for businesses to implement systems that can support the transformation of data into information that is crucial for the success of the company. Without the integration of data (both structured and unstructured) mining in business intelligence systems, invaluable knowledge is lost. However, there are currently many different models and approaches that must be explored to determine the best method of integration. Integration Challenges for Analytics, Business Intelligence, and Data Mining is a relevant academic book that provides empirical research findings on increasing the understanding of using data mining in the context of business intelligence and analytics systems. Covering topics that include big data, artificial intelligence, and decision making, this book is an ideal reference source for professionals working in the areas of data mining, business intelligence, and analytics; data scientists; IT specialists; managers; researchers; academicians; practitioners; and graduate students.

Machine Learning Essentials - Aloukakel Kassambura - 2018-03-10

Discovering knowledge from big multivariate data, recorded every days, requires specialized machine learning techniques. This book presents an easy to use practical guide in R to compute the most popular machine learning methods for exploring real word data sets, as well as, for building predictive models. The main parts of the book include: A) Unsupervised learning methods, to explore and discover knowledge from a large multivariate data set using clustering and principal component methods. You will learn hierarchical clustering, k-means, principal component analysis and correspondence analysis methods. B) Regression analysis, to predict a quantitative outcome value using linear regression and non-linear prediction strategies. C) Classification techniques, to predict a qualitative outcome value using logistic regression, discriminant analysis, naive Bayes classifier and support vector machines. D) Advanced machine learning methods, to build robust regression and classification models using k-nearest neighbors methods, decision tree models, ensemble methods (bagging, random forest and boosting). E) Model selection methods, to select automatically the best combination of predictor variables for building an optimal predictive model. These include, best subsets selection methods, stepwise regression and penalized regression (ridge, lasso and elastic net regression models). We also present principal component-based regression methods, which are useful when the data contain multiple correlated predictor variables. F) Model validation and evaluation techniques for measuring the performance of a predictive model. G) Model diagnostics for detecting and fixing a potential problems in a predictive model. The book presents the basic principles of these tasks and provide many examples in R. This book offers solid guidance in data mining for students and academics; practitioners; and graduate students. Key features: - Covers machine learning algorithm and implementation - Key mathematical concepts are presented - Short, self-contained chapters with practical examples.


missions in fact also treat an envisaged mutual impact among them. As for the 2002 edition in Irvine, the organizers wanted to stimulate this cross-pollination with a program of shared famous keynote speakers (this year we got Sycara, - ble, Soley and Mylopoulos!), and encouraged multiple attendance by providing authors with free access to another conference or workshop of their choice. We received an even larger number of submissions than last year for the three conferences (360 in total) and the workshops (170 in total). Not only can we therefore again claim a measurable success in attracting a representative volume of scientists’ papers, but such a high degree of interest also enabled the program committee to compose a high-quality cross-section of worldwide research in the areas covered. In spite of the increased number of submissions, the Program Chairs of the three main conferences decided to accept only approximately the same number of papers for presentation and publication as in 2002 (i.e., around 1 paper out of every 4–5 submitted). For the workshops, the acceptance rate was about 1 in 2. Also for this reason, we decided to separate the proceedings into two volumes with their own titles, and we are grateful to Springer-Verlag for their collaboration in producing these two books. The reviewing process by the respective program committees was very professional and each paper in the main conferences was reviewed by at least three referees.


missions in fact also treat an envisaged mutual impact among them. As for the 2002 edition in Irvine, the organizers wanted to stimulate this cross-pollination with a program of shared famous keynote speakers (this year we got Sycara, - ble, Soley and Mylopoulos!), and encouraged multiple attendance by providing authors with free access to another conference or workshop of their choice. We received an even larger number of submissions than last year for the three conferences (360 in total) and the workshops (170 in total). Not only can we therefore again claim a measurable success in attracting a representative volume of scientists’ papers, but such a high degree of interest also enabled the program committee to compose a high-quality cross-section of worldwide research in the areas covered. In spite of the increased number of submissions, the Program Chairs of the three main conferences decided to accept only approximately the same number of papers for presentation and publication as in 2002 (i.e., around 1 paper out of every 4–5 submitted). For the workshops, the acceptance rate was about 1 in 2. Also for this reason, we decided to separate the proceedings into two volumes with their own titles, and we are grateful to Springer-Verlag for their collaboration in producing these two books. The reviewing process by the respective program committees was very professional and each paper in the main conferences was reviewed by at least three referees.
Neural networks, Markov chains, tools from statistical physics and wavelet
the analysis of big data coming from biology, medicine and economics.
This graduate text covers a variety of mathematical and statistical tools for
analysis are used to develop efficient computational algorithms, which are
good professional and each paper in the main conferences was reviewed by
at least three referees.
Intelligent Data Engineering and Automated Learning - IDEAL 2004 -
Zhen Rong Yang - 2004-10-29
This book constitutes the refereed proceedings of the 5th International
Conference on Intelligent Data Engineering and Automated Learning,
IDEAL 2004, held in Exeter, UK, in August 2004. The 124 revised full
papers presented were carefully reviewed and selected from 272
submissions. The papers are organized in topical sections on bioinformatics,
data mining and knowledge engineering, learning algorithms and systems,
financial engineering, and agent technologies.
Intelligent Data Engineering and Automated Learning - IDEAL 2004 -
Zhen Rong Yang - 2004-10-29
This book constitutes the refereed proceedings of the 5th International
Conference on Intelligent Data Engineering and Automated Learning,
IDEAL 2004, held in Exeter, UK, in August 2004. The 124 revised full
papers presented were carefully reviewed and selected from 272
submissions. The papers are organized in topical sections on bioinformatics,
data mining and knowledge engineering, learning algorithms and systems,
financial engineering, and agent technologies.
Knowledge Discovery in Databases: PKDD 2007 - Joost N. Kok -
2007-08-30
This book constitutes the refereed proceedings of the 11th European
Conference on Principles and Practice of Knowledge Discovery in
Databases, PKDD 2007, held in Warsaw, Poland, co-located with ECML
2007, the 18th European Conference on Machine Learning. The 28 revised
full papers and 35 revised short papers present original results on leading-
edge subjects of discovery from conventional and complex data and
address all current issues in the area.
Knowledge Discovery in Databases: PKDD 2007 - Joost N. Kok -
2007-08-30
This book constitutes the refereed proceedings of the 11th European
Conference on Principles and Practice of Knowledge Discovery in
Databases, PKDD 2007, held in Warsaw, Poland, co-located with ECML
2007, the 18th European Conference on Machine Learning. The 28 revised
full papers and 35 revised short papers present original results on leading-
edge subjects of discovery from conventional and complex data and
address all current issues in the area.
Explaining the Success of Nearest Neighbor Methods in Prediction -
George H. Chen - 2018-05-30
Explains the success of Nearest Neighbor Methods in Prediction, both in
theory and in practice.
Explaining the Success of Nearest Neighbor Methods in Prediction -
George H. Chen - 2018-05-30
Explains the success of Nearest Neighbor Methods in Prediction, both in
theory and in practice.
Computational Methods for Data Analysis - Yeliz Karaca - 2018-12-17
This graduate text covers a variety of mathematical and statistical tools for
the analysis of big data coming from biology, medicine and economics.
Neural networks, Markov chains, tools from statistical physics and wavelet
analysis are used to develop efficient computational algorithms, which are
then used for the processing of real-life data using Matlab.
Computational Methods for Data Analysis - Yeliz Karaca - 2018-12-17
This graduate text covers a variety of mathematical and statistical tools for
the analysis of big data coming from biology, medicine and economics.
Neural networks, Markov chains, tools from statistical physics and wavelet
analysis are used to develop efficient computational algorithms, which are
then used for the processing of real-life data using Matlab.
Encyclopedia of GIS - Shashi Shekhar - 2007-12-12
The Encyclopedia of GIS provides a comprehensive and authoritative guide,
contributed by experts and peer-reviewed for accuracy, and alphabetically
arranged for convenient access. The entries explain key software and
processes used by geographers and computational scientists. Major
overviews are provided for nearly 200 topics: Geoinformatics, Spatial
Cognition, and Location-Based Services and more. Shorter entries define
specific terms and concepts. The reference will be published as a print
volume with abundant black and white art, and simultaneously as an XML
online reference with hyperlinked citations, cross-references, four-color art,
links to web-based maps, and other interactive features.
Encyclopedia of GIS - Shashi Shekhar - 2007-12-12
The Encyclopedia of GIS provides a comprehensive and authoritative guide,
contributed by experts and peer-reviewed for accuracy, and alphabetically
arranged for convenient access. The entries explain key software and
processes used by geographers and computational scientists. Major
overviews are provided for nearly 200 topics: Geoinformatics, Spatial
Cognition, and Location-Based Services and more. Shorter entries define
specific terms and concepts. The reference will be published as a print
volume with abundant black and white art, and simultaneously as an XML
online reference with hyperlinked citations, cross-references, four-color art,
links to web-based maps, and other interactive features.
Hands-On Machine Learning with R provides a practical and applied
approach to learning and developing intuition into today’s most popular
machine learning methods. This book serves as a practitioner’s guide to the
machine learning process and is meant to help the reader learn to apply the
machine learning stack within R, which includes using various R packages
such as glmnet, h2o, ranger, xgboost, keras, and others to effectively model
and gain insight from their data. The book favors a hands-on approach,
providing an intuitive understanding of machine learning concepts through
concrete examples and just a little bit of theory. Throughout this book, the
reader will be exposed to the entire machine learning process including feature
engineering, resampling, hyperparameter tuning, model evaluation,
and interpretation. The reader will be exposed to powerful algorithms such as
regularized regression, random forests, gradient boosting machines,
deep learning, generalized low rank models, and more! By favoring a hands-
on approach and using real world data, the reader will gain an intuitive
understanding of the architectures and engines that drive these algorithms
and packages, understand when and how to tune the various
hyperparameters, and be able to interpret model results. By the end of this
book, the reader should have a firm grasp of R’s machine learning stack and
be able to implement a systematic approach for producing high quality
modeling results. Features: · Offers a practical and applied introduction to
the most popular machine learning methods. · Topics covered include
feature engineering, resampling, deep learning and more. · Uses a hands-on
approach and real world data.
Hands-On Machine Learning with R provides a practical and applied
approach to learning and developing intuition into today’s most popular
machine learning methods. This book serves as a practitioner’s guide to the
machine learning process and is meant to help the reader learn to apply the
machine learning stack within R, which includes using various R packages
such as glmnet, h2o, ranger, xgboost, keras, and others to effectively model
and gain insight from their data. The book favors a hands-on approach,
providing an intuitive understanding of machine learning concepts through
concrete examples and just a little bit of theory. Throughout this book, the
reader will be exposed to the entire machine learning process including feature
engineering, resampling, hyperparameter tuning, model evaluation,
and interpretation. The reader will be exposed to powerful algorithms such as
regularized regression, random forests, gradient boosting machines,
deep learning, generalized low rank models, and more! By favoring a hands-
on approach and using real world data, the reader will gain an intuitive
understanding of the architectures and engines that drive these algorithms
and packages, understand when and how to tune the various
hyperparameters, and be able to interpret model results. By the end of this
book, the reader should have a firm grasp of R’s machine learning stack and
be able to implement a systematic approach for producing high quality
modeling results. Features: · Offers a practical and applied introduction to
the most popular machine learning methods. · Topics covered include
feature engineering, resampling, deep learning and more. · Uses a hands-on
approach and real world data.
Advanced Data Mining and Applications - Xudong Luo - 2014-12-17
This book constitutes the Proceedings of the 10th International Conference
on Advanced Data Mining and Applications, ADMA 2014, held in Guilin,
China during December 2014. The 48 regular papers and 10 workshop
papers are organized in topical sections on bioinformatics, feature
engineering, resampling, hyperparameter tuning, model evaluation,
and interpretation. The reader will be exposed to powerful algorithms such as
regularized regression, random forests, gradient boosting machines,
deep learning, generalized low rank models, and more! By favoring a hands-
on approach and using real world data, the reader will gain an intuitive
understanding of the architectures and engines that drive these algorithms
and packages, understand when and how to tune the various
hyperparameters, and be able to interpret model results. By the end of this
book, the reader should have a firm grasp of R’s machine learning stack and
be able to implement a systematic approach for producing high quality
modeling results. Features: · Offers a practical and applied introduction to
the most popular machine learning methods. · Topics covered include
feature engineering, resampling, deep learning and more. · Uses a hands-on
approach and real world data.
and sequential data; entity resolution and topic modeling; itemset and high-
90 submissions. They deal with the following topics: data mining, social
network and social media, database, dimensionality reduction, advance machine learning techniques, classification, big data and
applications, clustering methods, machine learning, and data mining and
database.

**Advanced Data Mining and Applications** - Xuong Luo - 2014-12-17
This book constitutes the proceedings of the 10th International Conference
on Advanced Data Mining and Applications, ADMA 2014, held in Guilin,
China during December 2014. The 48 regular papers and 10 workshop
papers presented in this volume were carefully reviewed and selected from
90 submissions. They deal with the following topics: data mining, social
network and social media, recommend systems, database, dimensionality
reduction, advance machine learning techniques, classification, big data and
applications, clustering methods, machine learning, and data mining and
database.

**Machine Learning for Hackers** - Drew Conway - 2012-02-13
If you're an experienced programmer interested in crunching data, this
book will get you started with machine learning—a toolkit of algorithms that
enables computers to train themselves to automate useful tasks. Authors
Drew Conway and John Myles White help you understand machine learning and
statistics tools through stories of hands-on case studies, instead of a
traditional math-heavy presentation. Each chapter focuses on a specific
problem in machine learning, such as classification, prediction,
optimization, and recommendation. Using the R programming language,
you'll learn how to analyze sample datasets and write simple machine
learning algorithms. Machine Learning for Hackers is ideal for
programmers from any background, including business, government, and
academic research. Develop a naive Bayesian classifier to determine if an
email is spam, based only on its text Use linear regression to predict the
number of page views for the top 1,000 websites Learn optimization
techniques by attempting to break a simple letter cipher Compare and
contrast U.S. Senators statistically, based on their voting records Build a
"whom to follow" recommendation system from Twitter data

**Advances in Knowledge Discovery and Data Mining** - Thu Cao - 2015
This two-volume set, LNAI 9077 + 9078, constitutes the refereed
proceedings of the 19th Pacific-Asia Conference on Advances in Knowledge
Discovery and Data Mining, PAKDD 2015, held in Ho Chi Minh City,
Vietnam, in May 2015. The proceedings contain 117 papers carefully
reviewed and selected from 405 submissions. They have been organized in
topical sections named: social networks and social media; classification;
machine learning; applications; novel methods and algorithms; opinion
mining and sentiment analysis; clustering; outlier and anomaly detection;
mining uncertain and imprecise data; mining temporal and spatial data;
feature extraction and selection; mining heterogeneous, high-dimensional,
and sequential data; entity resolution and topic modeling; itemset and high-
performance data mining; and recommendations.

**Handbook of Neural Computation** - Pijush Samui - 2017-07-18
Handbook of Neural Computation explores neural computation applications,
ranging from conventional fields of mechanical and civil engineering, to
electronics, electrical engineering and computer science. This book covers
the numerous applications of artificial and deep neural networks and their
uses in learning machines, including image and speech recognition, natural
language processing and risk analysis. Edited by renowned authorities in
this field, this work is comprised of articles from reputable industry and
academic scholars and experts from around the world. Each contributor
presents a specific research issue with its recent and future trends. As the
demand rises in the engineering and medical industries for neural networks
and other machine learning methods to solve different types of operations,
such as data prediction, classification of images, analysis of big data, and
intelligent decision-making, this book provides readers with the latest
cutting-edge research in one comprehensive text. Features high-quality
research articles on multivariate adaptive regression splines, the minimax
probability machine, and more Discusses machine learning techniques,
including classification, clustering, regression, web mining, information
retrieval and natural language processing Covers supervised, unsupervised,
reinforced, ensemble, and nature-inspired learning methods

**Pattern Classification** - Richard O. Duda - 2012-11-09
The first edition, published in 1973, has become a classic in the field.
Now with the second edition, readers will find information on key new
topics such as neural networks and statistical pattern recognition, the
theory of machine learning, and the theory of invariances. Also included are worked
examples, comparisons between different methods, extensive graphics,
expanded exercises and computer project topics. An instructor's manual
presenting detailed solutions to all the problems in the book is available from
the Wiley editorial department.

**Handbook of Metabonomics** - John C. Lindon - 2011-08-11
Metabolomic biology operates at three levels – genes, proteins and
metabolites. This book is unique in that it provides a comprehensive
description of an approach (metabonomics) to characterise the endogenous
metabolites in a living system, complementing gene and protein studies
(genomics and proteomics). These "omics" methods form the basis for
understanding biology at a systems level. The Handbook of Metabonomics
and Metabolomics and the definitive work to be published on the rapidly
developing subjects of metabolic profiling, metabolite and biomarker identification,
comprehensive metabolites and metabolomics. It covers the
principles of the subject, the analytical and statistical techniques used and
the wide variety of applications. * comprehensive description of an
approach (metabonomics) to characterise the endogenous metabolites in a
living system, complementing gene and protein studies * aims to be the
definitive work on the rapidly expanding subjects of metabolic profiling,
subject, the analytical and statistical techniques used and the wide variety of applications.

The Handbook of Metabolomics and Metabolomics - John C. Lindon - 2011-08-11
Molecular biology operates at three levels – genes, proteins and metabolites. This book is unique in that it provides a comprehensive description of an approach (metabolomics) to characterise the endogenous metabolites in a living system, complementing gene and protein studies (genomics and proteomics). These "omics" methods form the basis for understanding biology at a systems level. The Handbook of Metabolomics and Metabolomics aims to be the definitive work on the rapidly expanding subjects of metabolic profiling, metabolite and biomarker identification, encompassing the fields of metabolomics and metabonomics. It covers the principles of the subject, the analytical and statistical techniques used and the wide variety of applications. * comprehensive description of an approach (metabolomics) to characterise the endogenous metabolites in a living system, complementing gene and protein studies * aims to be the definitive work on the rapidly expanding subjects of metabolic profiling, metabolite and biomarker identification * covers the principles of the subject, the analytical and statistical techniques used and the wide variety of applications.

Master Machine Learning Algorithms - Jason Brownlee - 2016-03-04
You must understand the algorithms to get good (and be recognized as being good) at machine learning. In this Ebook, finally cut through the math and learn exactly how machine learning algorithms work, then implement them from scratch, step-by-step.

Master Machine Learning Algorithms - Jason Brownlee - 2016-03-04
You must understand the algorithms to get good (and be recognized as being good) at machine learning. In this Ebook, finally cut through the math and learn exactly how machine learning algorithms work, then implement them from scratch, step-by-step.

Machine Learning with R - Brett Lantz - 2013-10-25
Written as a tutorial to explore and understand the power of R for machine learning. This practical guide that covers all of the need to know topics in a very systematic way. For each machine learning approach, each step in the process is detailed, from preparing the data for analysis to evaluating the results. These steps will build the knowledge you need to apply them to your own data science tasks. Intended for those who want to learn how to use R's machine learning capabilities and gain insight from your data. Perhaps you already know a bit about machine learning, but have never used R; or perhaps you know a little R but are new to machine learning. In either case, this book will get you up and running quickly. It would be helpful to have a bit of familiarity with basic programming concepts, but no prior experience is required.

Machine Learning with R - Brett Lantz - 2013-10-25
Written as a tutorial to explore and understand the power of R for machine learning. This practical guide that covers all of the need to know topics in a very systematic way. For each machine learning approach, each step in the process is detailed, from preparing the data for analysis to evaluating the results. These steps will build the knowledge you need to apply them to your own data science tasks. Intended for those who want to learn how to use R's machine learning capabilities and gain insight from your data. Perhaps you already know a bit about machine learning, but have never used R; or perhaps you know a little R but are new to machine learning. In either case, this book will get you up and running quickly. It would be helpful to have a bit of familiarity with basic programming concepts, but no prior experience is required.

Machine Learning Algorithms From Scratch with Python - Jason Brownlee - 2016-11-16
You must understand algorithms to get good at machine learning. The problem is that they are only ever explained using Math. No longer. In this Ebook, finally cut through the math and learn exactly how machine learning algorithms work. Using clear explanations, simple pure Python code (no libraries!) and step-by-step tutorials you will discover how to load and prepare data, evaluate model skill, and implement a suite of linear, nonlinear and ensemble machine learning algorithms from scratch.

Machine Learning Algorithms From Scratch with Python - Jason Brownlee - 2016-11-16
You must understand algorithms to get good at machine learning. The problem is that they are only ever explained using Math. No longer. In this Ebook, finally cut through the math and learn exactly how machine learning algorithms work. Using clear explanations, simple pure Python code (no libraries!) and step-by-step tutorials you will discover how to load and prepare data, evaluate model skill, and implement a suite of linear, nonlinear and ensemble machine learning algorithms from scratch.

Database Systems for Advanced Applications - Sourav S. Bhomick - 2014-04-01
These two volumes set LNCS 8421 and LNCS 8422 constitutes the refereed proceedings of the 19th International Conference on Database Systems for Advanced Applications, DASFAA 2014, held in Bali, Indonesia, in April 2014. The 62 revised full papers presented together with 1 extended abstract paper, 4 industrial papers, 6 demo presentations, 3 tutorials and 1 panel paper were carefully reviewed and selected from a total of 257 submissions. The papers cover the following topics: big data management, indexing and query processing, graph data management, spatio-temporal data management, database for emerging hardware, data mining, probabilistic and uncertain data management, web and social data management, security, privacy and trust, keyword search, data stream management and data quality.

Database Systems for Advanced Applications - Sourav S. Bhomick - 2014-04-01
These two volumes set LNCS 8421 and LNCS 8422 constitutes the refereed proceedings of the 19th International Conference on Database Systems for Advanced Applications, DASFAA 2014, held in Bali, Indonesia, in April 2014. The 62 revised full papers presented together with 1 extended abstract paper, 4 industrial papers, 6 demo presentations, 3 tutorials and 1 panel paper were carefully reviewed and selected from a total of 257 submissions. The papers cover the following topics: big data management, indexing and query processing, graph data management, spatio-temporal data management, database for emerging hardware, data mining, probabilistic and uncertain data management, web and social data management, security, privacy and trust, keyword search, data stream management and data quality.

Artificial Intelligence in Precision Health - Debmalya Barh - 2020-03-04
Artificial Intelligence in Precision Health: From Concept to Applications provides a readily available resource to understand artificial intelligence and its real time applications in precision medicine in practice. Written by experts from different countries and with diverse background, the content encompasses accessible knowledge easily understandable for non-specialists in computer sciences. The book discusses topics such as cognitive computing and emotional intelligence, big data analysis, clinical decision support systems, deep learning, personal omics, digital health, predictive models, prediction of epidemics, drug discovery, precision nutrition and fitness. Additionally, there is a section dedicated to discuss and analyze AI products related to precision healthcare already available. This book is a valuable source for clinicians, healthcare workers, and researchers from diverse areas of biomedical field who may or may not have computational background, and want to learn more about the innovative field of artificial intelligence for precision health. Provides computational approaches used in artificial intelligence easily understandable for non-computer specialists Gives know-how and real successful cases of artificial intelligence approaches in predictive models, modeling disease physiology, and public health surveillance Discusses the applicability of AI on multiple areas, such as drug discovery, clinical trials, radiology, surgery, patient care and clinical decision support.

Artificial Intelligence in Precision Health - Debmalya Barh - 2020-03-04
Artificial Intelligence in Precision Health: From Concept to Applications provides a readily available resource to understand artificial intelligence and its real time applications in precision medicine in practice. Written by experts from different countries and with diverse background, the content encompasses accessible knowledge easily understandable for non-specialists in computer sciences. The book discusses topics such as cognitive computing and emotional intelligence, big data analysis, clinical decision support systems, deep learning, personal omics, digital health, predictive models, prediction of epidemics, drug discovery, precision nutrition and fitness. Additionally, there is a section dedicated to discuss and analyze AI products related to precision healthcare already available. This book is a valuable source for clinicians, healthcare workers, and researchers from diverse areas of biomedical field who may or may not have computational background, and want to learn more about the innovative field of artificial intelligence for precision health. Provides computational approaches used in artificial intelligence easily understandable for non-computer specialists Gives know-how and real successful cases of artificial intelligence approaches in predictive models, modeling disease physiology, and public health surveillance Discusses the applicability of AI on multiple areas, such as drug discovery, clinical trials, radiology, surgery, patient care and clinical decision support.


of heterogeneous data over a long time span. This requirement demands effective ontology-based information retrieval approaches for clinical information systems so that the pertinent information can be mined from large amount of distributed data. This unique and groundbreaking book highlights the key advances in ontology-based information retrieval techniques being applied in the healthcare domain and covers the following areas: Semantic data integration in e-health care systems Keyword-based medical information retrieval Ontology-based query retrieval support for e-health implementation Ontology-based multi-agent system for matchmaking patient healthcare monitoring A multi-agent system for querying heterogeneous data sources with ontologies for reducing cost of customized healthcare systems A methodology for ontology based multi agent systems development Ontology based systems for clinical systems: validity, ethics and regulation

Ontology-Based Information Retrieval for Healthcare Systems - Vishal Jain - 2020-07-29

With the advancements of semantic web, ontology has become the crucial mechanism for representing concepts in various domains. For research and dispersal of customized healthcare services, a major challenge is to efficiently retrieve and analyze individual patient data from a large volume of heterogeneous data over a long time span. This requirement demands effective ontology-based information retrieval approaches for clinical information systems so that the pertinent information can be mined from large amount of distributed data. This unique and groundbreaking book highlights the key advances in ontology-based information retrieval techniques being applied in the healthcare domain and covers the following areas: Semantic data integration in e-health care systems Keyword-based medical information retrieval Ontology-based query retrieval support for e-health implementation Ontology-based multi-agent system for matchmaking patient healthcare monitoring A multi-agent system for querying heterogeneous data sources with ontologies for reducing cost of customized healthcare systems A methodology for ontology based multi agent systems development Ontology based systems for clinical systems: validity, ethics and regulation

Ontology-Based Information Retrieval for Healthcare Systems - Vishal Jain - 2020-07-29

With the advancements of semantic web, ontology has become the crucial mechanism for representing concepts in various domains. For research and dispersal of customized healthcare services, a major challenge is to efficiently retrieve and analyze individual patient data from a large volume of heterogeneous data over a long time span. This requirement demands effective ontology-based information retrieval approaches for clinical information systems so that the pertinent information can be mined from large amount of distributed data. This unique and groundbreaking book highlights the key advances in ontology-based information retrieval techniques being applied in the healthcare domain and covers the following areas: Semantic data integration in e-health care systems Keyword-based medical information retrieval Ontology-based query retrieval support for e-health implementation Ontology-based multi-agent system for matchmaking patient healthcare monitoring A multi-agent system for querying heterogeneous data sources with ontologies for reducing cost of customized healthcare systems A methodology for ontology based multi agent systems development Ontology based systems for clinical systems: validity, ethics and regulation

Advances in Knowledge Discovery and Data Mining - Hady W. Laww - 2020

The two-volume set LNAI 12084 and 12085 constitutes the thoroughly refereed proceedings of the 24th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2020, which was due to be held in Singapore, in May 2020. The conference was held virtually due to the COVID-19 pandemic. The 135 full papers presented were carefully reviewed and selected from 628 submissions. The papers present new ideas, original research results, and practical development experiences from all KDD related areas, including data mining, data warehousing, machine learning, artificial intelligence, databases, statistics, knowledge engineering, visualization, decision-making systems, and the emerging applications. They are organized in the following topical sections: recommender systems; classification; clustering; mining social networks; representation learning and embedding; mining behavioral data; deep learning; feature extraction and selection; human, domain, organizational and social factors in data mining; mining sequential data; mining imbalanced data; association; privacy and security; supervised learning; novel algorithms; mining multimedia/multi-temporal data; application; mining graph and network data; anomaly detection and mining; mining spatial, temporal, unstructured and semi-structured data; sentiment analysis; statistical/algorithmic models; multi-source/distributed/parallel/cloud computing.

Advances in Knowledge Discovery and Data Mining - Hady W. Laww - 2020

The two-volume set LNAI 12084 and 12085 constitutes the thoroughly refereed proceedings of the 24th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2020, which was due to be held in Singapore, in May 2020. The conference was held virtually due to the COVID-19 pandemic. The 135 full papers presented were carefully reviewed and selected from 628 submissions. The papers present new ideas, original research results, and practical development experiences from all KDD related areas, including data mining, data warehousing, machine learning, artificial intelligence, databases, statistics, knowledge engineering, visualization, decision-making systems, and the emerging applications. They are organized in the following topical sections: recommender systems; classification; clustering; mining social networks; representation learning and embedding; mining behavioral data; deep learning; feature extraction and selection; human, domain, organizational and social factors in data mining; mining sequential data; mining imbalanced data; association; privacy and security; supervised learning; novel algorithms; mining multimedia/multi-temporal data; application; mining graph and network data; anomaly detection and mining; mining spatial, temporal, unstructured and semi-structured data; sentiment analysis; statistical/algorithmic models; multi-source/distributed/parallel/cloud computing.

Advances in Knowledge Discovery and Data Mining - Hady W. Laww - 2020

The two-volume set LNAI 12084 and 12085 constitutes the thoroughly refereed proceedings of the 24th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2020, which was due to be held in Singapore, in May 2020. The conference was held virtually due to the COVID-19 pandemic. The 135 full papers presented were carefully reviewed and selected from 628 submissions. The papers present new ideas, original research results, and practical development experiences from all KDD related areas, including data mining, data warehousing, machine learning, artificial intelligence, databases, statistics, knowledge engineering, visualization, decision-making systems, and the emerging applications. They are organized in the following topical sections: recommender systems; classification; clustering; mining social networks; representation learning and embedding; mining behavioral data; deep learning; feature extraction and selection; human, domain, organizational and social factors in data mining; mining sequential data; mining imbalanced data; association; privacy and security; supervised learning; novel algorithms; mining multimedia/multi-temporal data; application; mining graph and network data; anomaly detection and mining; mining spatial, temporal, unstructured and semi-structured data; sentiment analysis; statistical/algorithmic models; multi-source/distributed/parallel/cloud computing.

Advances in Knowledge Discovery and Data Mining - Hady W. Laww - 2020

The two-volume set LNAI 12084 and 12085 constitutes the thoroughly refereed proceedings of the 24th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2020, which was due to be held in Singapore, in May 2020. The conference was held virtually due to the COVID-19 pandemic. The 135 full papers presented were carefully reviewed and selected from 628 submissions. The papers present new ideas, original research results, and practical development experiences from all KDD related areas, including data mining, data warehousing, machine learning, artificial intelligence, databases, statistics, knowledge engineering, visualization, decision-making systems, and the emerging applications. They are organized in the following topical sections: recommender systems; classification; clustering; mining social networks; representation learning and embedding; mining behavioral data; deep learning; feature extraction and selection; human, domain, organizational and social factors in data mining; mining sequential data; mining imbalanced data; association; privacy and security; supervised learning; novel algorithms; mining multimedia/multi-temporal data; application; mining graph and network data; anomaly detection and mining; mining spatial, temporal, unstructured and semi-structured data; sentiment analysis; statistical/algorithmic models; multi-source/distributed/parallel/cloud computing.
combined to produce tremendous value. If you want to pick-up the skills you and selection; human, domain, organizational and social factors in data mining; mining semi-structured data; mining imbalanced data; association; privacy and security; supervised learning; novel algorithms; mining multi-
media/multi-dimensional data; application; mining graph and network data; anomaly detection and analytics; mining spatial, temporal, unstructured and semi-structured data; sentiment analysis; statistical/graphical model; multi-
source/distributed/parallel/cloud computing.

Time Efficient and Quality Effective K Nearest Neighbor Search in High Dimension Space - Renwei Yu - 2011

K-Nearest-Neighbors (KNN) search is a fundamental problem in many application domains such as database and data mining, information retrieval, machine learning, pattern recognition and plagiarism detection. Locality sensitive hash (LSH) is so far the most practical approximate KNN search algorithm for high dimensional data. Algorithms such as Multi-Probe LSH and LSH-Forest improve upon the basic LSH algorithm by varying hash bucket size dynamically at query time, so these two algorithms can answer different KNN queries adaptively. However, these two algorithms need a data access post-processing step after candidates’ collection in order to get the final answer to the KNN query. In this thesis, Multi-Probe LSH with data access post-processing (Multi-Probe LSH with DAPP) algorithm and LSH-
Forest with data access post-processing (LSH-Forest with DAPP) algorithm are improved by replacing the costly data access post-processing (DAPP) step with a much faster histogram-based post-processing (HBPP). Two HBPP algorithms: LSH-Forest with HBPP and Multi-
Probe LSH with HBPP are presented in this thesis, both of them achieve the three goals for KNN search in large scale high dimensional data set: high search quality, high time efficiency, high space efficiency. None of the previous KNN algorithms can achieve all three goals. More specifically, it is shown that HBPP algorithms can achieve higher search quality (as good as LSH-Forest with DAPP and Multi-Probe LSH with DAPP) with much less time cost (one to several orders of magnitude speedup) and same memory usage. It is also shown that with almost same time cost and memory usage, HBPP algorithms can always achieve better search quality than LSH-Forest with random pick (LSH-Forest with RP) and Multi-Probe LSH with random pick (Multi-Probe LSH with RP). Moreover, to achieve a very high search quality, Multi-Probe with HBPP is always a better choice than LSH-Forest with HBPP, regardless of the distribution, size and dimension number of the data set.

Data Science For Dummies - Lillian Pierson - 2017-03-06

Discover how data science can help you gain in-depth insight into your business - the easy way! Jobs in data science abound, but few people have the data science skills needed to fill these increasingly important roles. Data Science For Dummies is perfect for you if you want a quick primer on all areas of the expansive data science space. With a focus on business cases, the book explores topics in big data, data science, and data engineering, and how these three areas are need to begin a new career or initiate a new project, reading this book will help you understand what technologies, programming languages, and mathematical methods on which to focus. While this book serves as a wildly fantastic guide through the broad, sometimes intimidating field of big data and data science, it is not an instruction manual for hands-on implementation. Here’s what to expect: Provides a background in big data and data engineering before moving on to data science and how it’s applied to generate value Includes a guide to the most recent advances in information and communication technologies that can be used to showcase, summarize, and communicate the data insights you generate It’s a big, big world out there—let Data Science For Dummies help you harness its power and gain a competitive edge for your organization.

Innovation in Health Informatics - Militsiadis Lytras - 2019-11-13

Innovation in Health Informatics: A Smart Healthcare Primer explains how the most recent advances in information and communication technologies have paved the way for new breakthroughs in healthcare. The book showcases current and prospective applications in a context defined by an imperative to deliver efficient, patient-centered and sustainable healthcare systems. Topics discussed include big data, medical data analytics, artificial intelligence, machine learning, virtual and augmented reality, 5g and sensors, Internet of Things, nanotechnologies and biotechnologies. Additionally, there is a discussion on social issues and policy-making for the implementation of smart healthcare. This book is a valuable resource for undergraduate and graduate students, practitioners, researchers, clinicians and data scientists who are interested in how to explore the intersections between bioinformatics and health informatics. Provides a holistic discussion on the new landscape of medical technologies, including big data, analytics, artificial intelligence, machine learning, virtual and augmented reality, 5g and sensors, Internet of Things, nanotechnologies and biotechnologies Presents a case study driven approach, with references to real-world applications and systems Discusses topics with a research-oriented approach that aims to promote research skills and competencies of readers

Innovation in Health Informatics - Militsiadis Lytras - 2019-11-13

Innovation in Health Informatics: A Smart Healthcare Primer explains how the most recent advances in information and communication technologies have paved the way for new breakthroughs in healthcare. The book showcases current and prospective applications in a context defined by an imperative to deliver efficient, patient-centered and sustainable healthcare systems. Topics discussed include big data, medical data analytics, artificial intelligence, machine learning, virtual and augmented reality, 5g and sensors, Internet of Things, nanotechnologies and biotechnologies. Additionally, there is a discussion on social issues and policy-making for the implementation of smart healthcare. This book is a valuable resource for undergraduate and graduate students, practitioners, researchers, clinicians and data scientists who are interested in how to explore the intersections between bioinformatics and health informatics. Provides a holistic discussion on the new landscape of medical technologies, including big data, analytics, artificial intelligence, machine learning, virtual and augmented reality, 5g and sensors, Internet of Things, nanotechnologies and biotechnologies.
Clinical Research Computing - Prakash Nadkarni - 2016-04-29
Clinical Research Computing: A Practitioner’s Handbook deals with the nuts-and-bolts of providing informatics and computing support for clinical research. The subjects that the practitioner must be aware of are not only technological and scientific, but also organizational and managerial. Therefore, the author offers case studies based on real life experiences in order to prepare the readers for the challenges they may face during their experiences either supporting clinical research or supporting electronic record systems. Clinical research computing is the application of computational methods to the broad field of clinical research. With the advent of modern digital computing, and the powerful data collection, storage, and analysis that is possible with it, it becomes more relevant to understand the technical details in order to fully seize its opportunities. Offers case studies, based on real-life examples where possible, to engage the readers with more complex examples Provides studies backed by technical details, e.g., schema diagrams, code snippets or algorithms illustrating particular techniques, to give the readers confidence to employ the techniques described in their own settings Offers didactic content organization and an increasing complexity through the chapters

Clinical Research Computing - Prakash Nadkarni - 2016-04-29
Clinical Research Computing: A Practitioner’s Handbook deals with the nuts-and-bolts of providing informatics and computing support for clinical research. The subjects that the practitioner must be aware of are not only technological and scientific, but also organizational and managerial. Therefore, the author offers case studies based on real life experiences in order to prepare the readers for the challenges they may face during their experiences either supporting clinical research or supporting electronic record systems. Clinical research computing is the application of computational methods to the broad field of clinical research. With the advent of modern digital computing, and the powerful data collection, storage, and analysis that is possible with it, it becomes more relevant to understand the technical details in order to fully seize its opportunities. Offers case studies, based on real-life examples where possible, to engage the readers with more complex examples Provides studies backed by technical details, e.g., schema diagrams, code snippets or algorithms illustrating particular techniques, to give the readers confidence to employ the techniques described in their own settings Offers didactic content organization and an increasing complexity through the chapters

Weighted K-nearest Neighbor Algorithm as an Object Localization Technique Using Passive RFID Tags - Akshay Shetty - 2010
Technologies using identification by radio frequencies (RFID) are experiencing rapid development and healthcare is a major application area benefiting from it. Highly pervasive RFID enables remote identification, tracking and localization of the medical staff, patients, medications and equipment, thus increasing safety, optimizing in real-time management and providing support for new ambient-intelligent services. This thesis describes and evaluates an algorithm that enables object localization and tracking using passive RFID tags. This thesis also describes scenarios of how this technology can be used as a part of building a smart trauma resuscitation room by tracking the equipments. The main contribution of this thesis is the adaptation of the Weighted K-Nearest Neighbor Algorithm as a localization technique to track objects in a confined and crowded space by using passive RFID tags. The input parameter to the algorithm is the received signal strength indicator (RSSI), which gives a measure of back-scattered radio frequencies from passive tags. While using RFID technology special attention has to be given to the placement of antennas to get the optimum result. Therefore, we analyzed various antenna placement configurations with mean error and error consistency as the two performance parameters. The detection of multiple tags and human occlusion are two major concerns while tracking tags in a confined space with many team members collaborating on solving a problem. The RF signal can be interrupted by people walking around randomly and holding multiple (tagged) instruments at the same time. While the algorithm worked fine when tracking multiple tags, we had to modify the experimental set-up and attach an antenna onto the ceiling (which we call a vertical antenna), so that even if all the wall antennas are blocked we get at least one input parameter to base our localization decision on. We evaluated the algorithm for different human occlusion scenarios i.e. blocking 1 or 2 wall antennas, standing in random positions and then roaming in the field area randomly. The mean error rate for the standing scenario was measured as 20% of the maximum possible error and 18% in the case of roaming. The results were consistent within our defined maximum error for 100% of the recorded readings. The results obtained were found to be satisfactory for our application where, more than the exact location of the object, knowing whether the object is within a particular region is good enough for the users to know what task is being carried out in the trauma bay. Also the algorithm holds good in an indoor environment having a lot of factors and materials which affect the RF signal disrupting accurate calculation of the location co-ordinates. The algorithm does not require extensive data collection prior to implementation which makes it easily deployable in any environment. Apart from the problems mentioned there are some other factors like materials on which the tags are attached and orientation of tags which were found to be potential hindrances for accurate localization. Acceptable solutions to these problems form a part of our future work.

Weighted K-nearest Neighbor Algorithm as an Object Localization Technique Using Passive RFID Tags - Akshay Shetty - 2010
Technologies using identification by radio frequencies (RFID) are experiencing rapid development and healthcare is a major application area benefiting from it. Highly pervasive RFID enables remote identification, tracking and localization of the medical staff, patients, medications and equipment, thus increasing safety, optimizing in real-time management and providing support for new ambient-intelligent services. This thesis describes and evaluates an algorithm that enables object localization and tracking using passive RFID tags. This thesis also describes scenarios of how this technology can be used as a part of building a smart trauma resuscitation room by tracking the equipments. The main contribution of this thesis is the adaptation of the Weighted K-Nearest Neighbor Algorithm as a localization technique to track objects in a confined and crowded space by using passive RFID tags. The input parameter to the algorithm is the received signal strength indicator (RSSI), which gives a measure of back-scattered radio frequencies from passive tags. While using RFID technology special attention has to be given to the placement of antennas to get the optimum result. Therefore, we analyzed various antenna placement configurations with mean error and error consistency as the two performance parameters. The detection of multiple tags and human occlusion are two major concerns while tracking tags in a confined space with many team members collaborating on solving a problem. The RF signal can be interrupted by people walking around randomly and holding multiple (tagged) instruments at the same time. While the algorithm worked fine when tracking multiple tags, we had to modify the experimental set-up and attach an antenna onto the ceiling (which we call a vertical antenna), so that even if all the wall antennas are blocked we get at least one input parameter to base our localization decision on. We evaluated the algorithm for different human occlusion scenarios i.e. blocking 1 or 2 wall antennas, standing in random positions and then roaming in the field area randomly. The mean error rate for the standing scenario was measured as 20% of the maximum possible error and 18% in the case of roaming. The results were consistent within our defined maximum error for 100% of the recorded readings. The results obtained were found to be satisfactory for our application where, more than the exact location of the object, knowing whether the object is within a particular region is good enough for the users to know what task is being carried out in the trauma bay. Also the algorithm holds good in an indoor environment having a lot of factors and materials which affect the RF signal disrupting accurate calculation of the location co-ordinates. The algorithm does not require extensive data collection prior to implementation which makes it easily deployable in any environment. Apart from the problems mentioned there are some other factors like materials on which the tags are attached and orientation of tags which were found to be potential hindrances for accurate localization. Acceptable solutions to these problems form a part of our future work.

A guide to using the power of S-PLUS to perform statistical analyses, providing both an introduction to the program and a course in modern statistical methods. Readers are assumed to have a basic grounding in statistics, thus the book is intended for would-be users, as well as students and researchers using statistics. Throughout, the emphasis is on presenting practical problems and full analyses of real data sets, with many of the methods discussed being modern approaches to topics such as linear and non-linear regression models, robust and smooth regression methods, survival analysis, multivariate analysis, tree-based methods, time series, spatial statistics, and classification. This second edition is intended for users of S-PLUS 3.3, or later, and covers both Windows and UNIX. It treats the

k-nearest-neighbor-algorithm-for-classification
practical graph algorithms that cover optimal pathfinding, importance bootstrapping, mixed effects linear and non-linear models, factor analysis, and regression with autocorrelated errors. The authors have written several software libraries which enhance S-PLUS, and these, plus all the datasets used, are available on the Internet.


A guide to using the power of S-PLUS to perform statistical analyses, providing both an introduction to the program and a course in modern statistical methods. Readers are assumed to have a basic grounding in statistics, thus the book is intended for would-be users, as well as students and researchers using statistics. Throughout, the emphasis is on presenting practical problems and full analyses of real data sets, with many of the methods discussed being modern approaches to topics such as linear and non-linear regression models, robust and smooth regression methods, survival analysis, multivariate analysis, tree-based methods, time series, spatial statistics, and classification. This second edition is intended for users of S-PLUS 3.3, or later, and covers both Windows and UNIX. It treats the recent developments in graphics and new statistical functionality, including bootstrapping, mixed effects linear and non-linear models, factor analysis, and regression with autocorrelated errors. The authors have written several software libraries which enhance S-PLUS, and these, plus all the datasets used, are available on the Internet.

Interpretable Machine Learning - Christoph Molnar - 2019

Interpretable Machine Learning - Christoph Molnar - 2019

Biological Distance Analysis - Marin A. Pilloud - 2016-08-19

Biological Distance Analysis: Forensic and Bioarchaeological Perspectives synthesizes research within the realm of biological distance analysis, highlighting current work within the field and discussing future directions. The book is divided into three main sections. The first section clearly outlines datasets and methods within biological distance analysis, beginning with a brief history of the field and how it has progressed to its current state. The second section focuses on approaches using the individual within a forensic context, including ancestry estimation and case studies. The final section concentrates on population-based bioarchaeological approaches, providing key techniques and examples from archaeological samples. The volume also includes an appendix with additional resources available to those interested in biological distance analysis. Defines datasets and how they are used within biodistance analysis Applies methodology to individual and population studies Bridges the sub-fields of forensic anthropology and bioarchaeology Highlights current research and future directions of biological distance analysis Identifies statistical programs and datasets for use in biodistance analysis Contains cases studies and thorough index for those interested in biological distance analyses

Biological Distance Analysis - Marin A. Pilloud - 2016-08-19

Biological Distance Analysis: Forensic and Bioarchaeological Perspectives synthesizes research within the realm of biological distance analysis, highlighting current work within the field and discussing future directions. The book is divided into three main sections. The first section clearly outlines datasets and methods within biological distance analysis, beginning with a brief history of the field and how it has progressed to its current state. The second section focuses on approaches using the individual within a forensic context, including ancestry estimation and case studies. The final section concentrates on population-based bioarchaeological approaches, providing key techniques and examples from archaeological samples. The volume also includes an appendix with additional resources available to those interested in biological distance analysis. Defines datasets and how they are used within biodistance analysis Applies methodology to individual and population studies Bridges the sub-fields of forensic anthropology and bioarchaeology Highlights current research and future directions of biological distance analysis Identifies statistical programs and datasets for use in biodistance analysis Contains cases studies and thorough index for those interested in biological distance analyses

Graph Algorithms - Mark Needham - 2019-05-16

Discover how graph algorithms can help you leverage the relationships within your data to develop more intelligent solutions and enhance your machine learning models. You’ll learn how graph analytics are uniquely suited to unfold complex structures and reveal difficult-to-find patterns lurking in your data. Whether you are trying to build dynamic network models or forecast real-world behavior, this book illustrates how graph algorithms deliver value—from finding vulnerabilities and bottlenecks to detecting communities and improving machine learning predictions. This practical book walks you through hands-on examples of how to use graph algorithms in Apache Spark and Neo4j—two of the most common choices for graph analytics. Also included: sample code and tips for over 20 through centrality, and community detection. Learn how graph analytics vary from conventional statistical analysis Understand how classic graph algorithms work, and how they are applied Get guidance on which algorithms to use for different types of questions Explore algorithm examples with working code and sample datasets from Spark and Neo4j See how connected feature extraction can increase machine learning accuracy and precision Walk through creating an ML workflow for link prediction combining Neo4j and Spark

Graph Algorithms - Mark Needham - 2019-05-16

Discover how graph algorithms can help you leverage the relationships within your data to develop more intelligent solutions and enhance your machine learning models. You’ll learn how graph analytics are uniquely suited to unfold complex structures and reveal difficult-to-find patterns lurking in your data. Whether you are trying to build dynamic network models or forecast real-world behavior, this book illustrates how graph algorithms deliver value—from finding vulnerabilities and bottlenecks to detecting communities and improving machine learning predictions. This practical book walks you through hands-on examples of how to use graph algorithms in Apache Spark and Neo4j—two of the most common choices for graph analytics. Also included: sample code and tips for over 20 through centrality, and community detection. Learn how graph analytics vary from conventional statistical analysis Understand how classic graph algorithms work, and how they are applied Get guidance on which algorithms to use for different types of questions Explore algorithm examples with working code and sample datasets from Spark and Neo4j See how connected feature extraction can increase machine learning accuracy and precision Walk through creating an ML workflow for link prediction combining Neo4j and Spark

Mastering Machine Learning with scikit-learn - Gavin Hackeling - 2017-07-24

Use scikit-learn to apply machine learning to real-world problems About This Book Master popular machine learning models including k-nearest neighbors, random forests, logistic regression, k-means, naive Bayes, and artificial neural networks Learn how to build and evaluate performance of efficient models using scikit-learn Practical guide to master your basics and learn from real life applications of machine learning Who This Book Is For This book is intended for software engineers who want to learn about common machine learning algorithms work and develop an intuition for how to use them, and for data scientists who want to learn about the scikit-learn API. Familiarity with machine learning fundamentals and Python are helpful, but not required. What You Will Learn Review fundamental concepts such as bias and variance Extract features from categorical variables, text, and images Predict the values of continuous variables using linear regression and K Nearest Neighbors Classify documents and images using logistic regression and support vector machines Create ensembles of estimators using bagging and boosting techniques Discover hidden structures in data using K-Means clustering Evaluate the performance of machine learning systems in common tasks In Detail Machine learning is the buzzword bringing computer science and statistics together to build smart and efficient models. Using powerful algorithms and techniques offered by machine learning you can automate any analytical model. This book examines a variety of machine learning models including popular machine learning algorithms such as k-nearest neighbors, logistic regression, naive Bayes, k-means, decision trees, and artificial neural networks. It discusses data preprocessing, hyperparameter optimization, and ensemble methods. You will build systems that classify documents, recognize images, detect ads, and more. You will learn to use scikit-learn’s API to extract features from categorical variables, text and images; evaluate model performance, and develop an intuition for how to improve your models performance. For this book, the author uses best-practice concepts of scikit-learn to build efficient models at work to carry out advanced tasks with the practical approach. Style and approach This book is motivated by the belief that you do not understand something until you can describe it simply. Work through toy problems to develop your understanding of the learning algorithms and models, then apply your learnings to real-life problems.

Mastering Machine Learning with scikit-learn - Gavin Hackeling - 2017-07-24

Use scikit-learn to apply machine learning to real-world problems About This Book Master popular machine learning models including k-nearest neighbors, random forests, logistic regression, k-means, naive Bayes, and artificial neural networks Learn how to build and evaluate performance of efficient models using scikit-learn Practical guide to master your basics and learn from real life applications of machine learning Who This Book Is For This book is intended for software engineers who want to learn about common machine learning algorithms work and develop an intuition for how to use them, and for data scientists who want to learn about the scikit-learn API. Familiarity with machine learning fundamentals and Python are helpful, but not required. What You Will Learn Review fundamental concepts such as bias and variance Extract features from categorical variables, text, and images Predict the values of continuous variables using linear regression and K Nearest Neighbors Classify documents and images using logistic regression and support vector machines Create ensembles of estimators using bagging and boosting techniques Discover hidden structures in data using K-Means clustering Evaluate the performance of machine learning systems in common tasks In Detail Machine learning is the buzzword bringing computer science and statistics together to build smart and efficient models. Using powerful algorithms and techniques offered by machine learning you can automate any analytical model. This book examines a variety of machine learning models including popular machine learning algorithms such as k-nearest neighbors, logistic regression, naive Bayes, k-means, decision trees, and artificial neural networks. It discusses data preprocessing, hyperparameter optimization, and ensemble methods. You will build systems that classify documents, recognize images, detect ads, and more. You will learn to use scikit-learn’s API to extract features from categorical variables, text and images; evaluate model performance, and develop an intuition for how to improve your models performance. For this book, the author uses best-practice concepts of scikit-learn to build efficient models at work to carry out advanced tasks with the practical approach. Style and approach This book is motivated by the belief that you do not understand something until you can describe it simply. Work through toy problems to develop your understanding of the learning algorithms and models, then apply your learnings to real-life problems.
API. Familiarity with machine learning fundamentals and Python are helpful, but not required. What You Will Learn Review fundamental concepts such as bias and variance Extract features from categorical variables, text, and images Predict the values of continuous variables using linear regression and K Nearest Neighbors Classify documents and images using logistic regression and support vector machines Create ensembles of estimators using bagging and boosting techniques Discover hidden structures in data using K-Means clustering Evaluate the performance of machine learning systems in common tasks In Detail Machine learning is the buzzword bringing computer science and statistics together to build smart and efficient models. Using powerful algorithms and techniques offered by machine learning you can automate any analytical model. This book examines a variety of machine learning models including popular machine learning algorithms such as k-nearest neighbors, logistic regression, naïve Bayes, k-means, decision trees, and artificial neural networks. It discusses data preprocessing, hyperparameter optimization, and ensemble methods. You will build systems that classify documents, recognize images, detect ads, and more. You will learn to use scikit-learn’s API to extract features from categorical variables, text and images; evaluate model performance, and develop an intuition for how to improve your model’s performance. By the end of this book, you will master all required concepts of scikit-learn to build efficient models at work to carry out advanced tasks with the practical approach. Style and approach This book is motivated by the belief that you do not understand something until you can describe it simply. Work through toy problems to develop your understanding of the learning algorithms and models, then apply your learnings to real-life problems.