



Review

A Review on Business Analytics: Definitions, Techniques, Applications and Challenges

Shiyu Liu ¹, Ou Liu ^{2,*} and Junyang Chen ³¹ School of Economics and Management, Beihang University, Beijing 100191, China² Wenzhou Institute, University of Chinese Academy of Sciences, Wenzhou 325001, China³ College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China

* Correspondence: oliu@ucas.ac.cn

Abstract: Over the past few decades, business analytics has been widely used in various business sectors and has been effective in increasing enterprise value. With the advancement of science and technology in the Big Data era, business analytics techniques have been changing and evolving rapidly. Therefore, this paper reviews the latest techniques and applications of business analytics based on the existing literature. Meanwhile, many problems and challenges are inevitable in the progress of business analytics. Therefore, this review also presents the current challenges faced by business analytics and open research directions that need further consideration. All the research papers were obtained from the Web of Science and Google Scholar databases and were filtered with several selection rules. This paper will help to provide important insights for researchers in the field of business analytics, as it presents the latest techniques, various applications and several directions for future research.

Keywords: business analytics; descriptive analytics; predictive analytics; perspective analytics; machine learning and artificial intelligence

MSC: 62-07; 68U35



Citation: Liu, S.; Liu, O.; Chen, J. A Review on Business Analytics: Definitions, Techniques, Applications and Challenges. *Mathematics* **2023**, *11*, 899. <https://doi.org/10.3390/math11040899>

Academic Editor: Amir Mosavi

Received: 31 December 2022

Revised: 5 February 2023

Accepted: 7 February 2023

Published: 10 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent decades, data have been rapidly changing the world. Especially in the era of Big Data, data are cheap and ubiquitous, but what makes datum a valuable asset is how it is used to obtain useful information. Since there are many different types of business objectives, different analytics techniques are needed to achieve them. These techniques have many applications in the business area and “business analytics” enables the business application of Big Data. Since the emergence of the term business analytics, it is growing by leaps and bounds, reflecting the increasing importance of data in terms of volume, variety and velocity [1]. Although there is no uniform definition of business analytics, the existing definitions can be summarized into several dimensions, such as a movement, a transformation process, a capacity set and so on [2].

Interest in analytics and data science is growing as business organizations are using business analytics extensively to improve their business value. Business analytics has evolved into an important part of the business decision-making process, using data to drive decisions and support decision-makers in making strategic, operational and tactical decisions [3]. Specifically, business analytics can help companies to leverage the value of historical data by harnessing the power of statistical and mathematical models and advanced techniques such as artificial intelligence algorithms. Through these models and algorithms, enterprises can integrate disparate data sources for trend prediction, decision optimization and more. As business analytics continues to evolve, its applications continue to broaden. It is adapted in some functional departments within the enterprise and some non-business areas.

Judging from the volume of literature in the database, there are many kinds of literature to study business analytics, including its techniques, impact, applications in some areas and so on. Among them, several scholars have systematically summarized the many aspects of business analytics. However, the techniques and applications of business analytics have changed significantly as technology has evolved rapidly in recent years. Thus, in order to organize the latest knowledge about business analytics, we present four main research questions:

RQ1: What is business analytics?

RQ2: How to achieve business analytics?

RQ3: Where is business analytics used?

RQ4: What are the challenges for business analytics?

This article is structured as follows. Section 2 introduces the methodology used in this review and conducts a simple bibliometric analysis of the literature. Section 3 concludes the definitions of business analytics in four categories to answer RQ1. Techniques used in business analytics are presented in Section 4 to answer RQ2. Section 5 describes applications of business analytics in several business areas and industry sectors to answer RQ3. RQ4 is responded to in Section 6 to reveal the challenges faced by business analytics. Finally, Section 7 concludes this paper.

2. Methodology and Literature Analysis

2.1. Methodology

To understand the research trends in business analytics, we collected related academic literature from Web of Science and Google Scholar databases since they are widely recognized and cover a large number of high-quality publications in peer-reviewed journals [4]. Then, we conducted a bibliometric analysis of the existing literature regarding the number of publications per year and their research directions in Section 2.2. Since there were plenty of materials on the research of business analytics, we designed several selection rules to filter the literature for further review. First, 'business analytics' should be contained in the title or abstract of publications. Second, we only focused on English publications. Third, we considered various publication types, including research articles, reviews and book chapters. What is more, to consider both the newness and impact of the articles, at least ten citations were required for the publications before 2020, while at least two citations were required for those after 2020. Based on the selection rules, high-impact academic pieces of literature were selected. Furthermore, it was feasible to read all the selected papers entirely. We then read the abstract of each piece of literature to decide whether it fit the goal of our further review.

Based on the methodology, we conducted the process of literature selection. Figure 1 shows the flowchart of the selection process. We researched on the Web of Science with the keyword 'business analytics' in the title or abstract, and without other selection rules, and the number of results was 821. After filtering language (English) and publication types (research articles, reviews and book chapters), there were 365 papers left. Then, we constrained the number of citations before and after 2020 and excluded 193 results. Finally, we read the abstract of the selected papers to further filter for relevant articles and there were 76 papers ready for in-depth review.

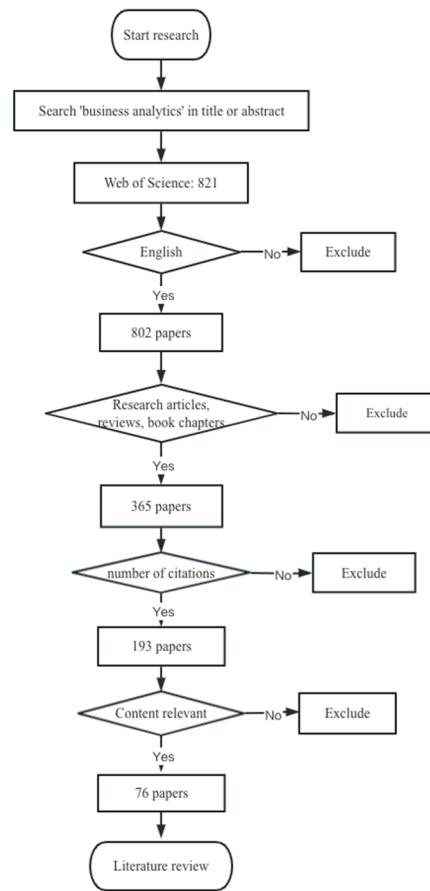


Figure 1. Process of selection.

2.2. Literature Analysis

Firstly, we conducted a quantitative analysis on business analytics literature in terms of the publication number per year from 2012 to 2022, which is shown in Figure 2. From 2012 to 2017, the number of publications per year showed a significant upward trend and peaked in 2017. After 2017, the number decreased slightly but still remained at a high level compared to 2012, which means that the research on business analytics continues to attract many scholars now.

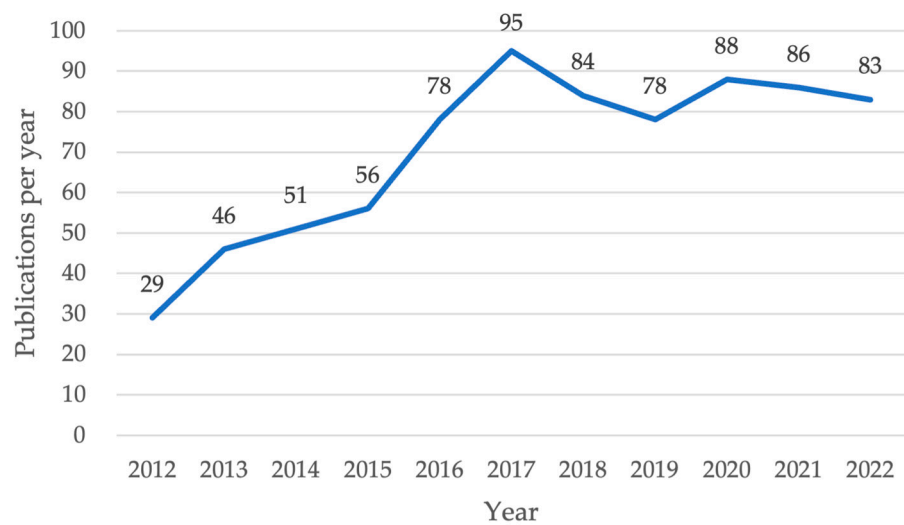


Figure 2. Publications per year.

Secondly, we conducted an analysis of the top ten research directions of academic literature on business analytics in Figure 3. It is clear that computer science is the most popular research direction among published literature about business analytics. It is because computer science is an essential part of business analytics and drives the development of business analytics applications. The second most popular research direction is engineering which implies the application area of business analytics, whereas the third one is business economics showing the value of business analytics on economics. The remaining research directions also all reflect the techniques and applications of business analytics, respectively.

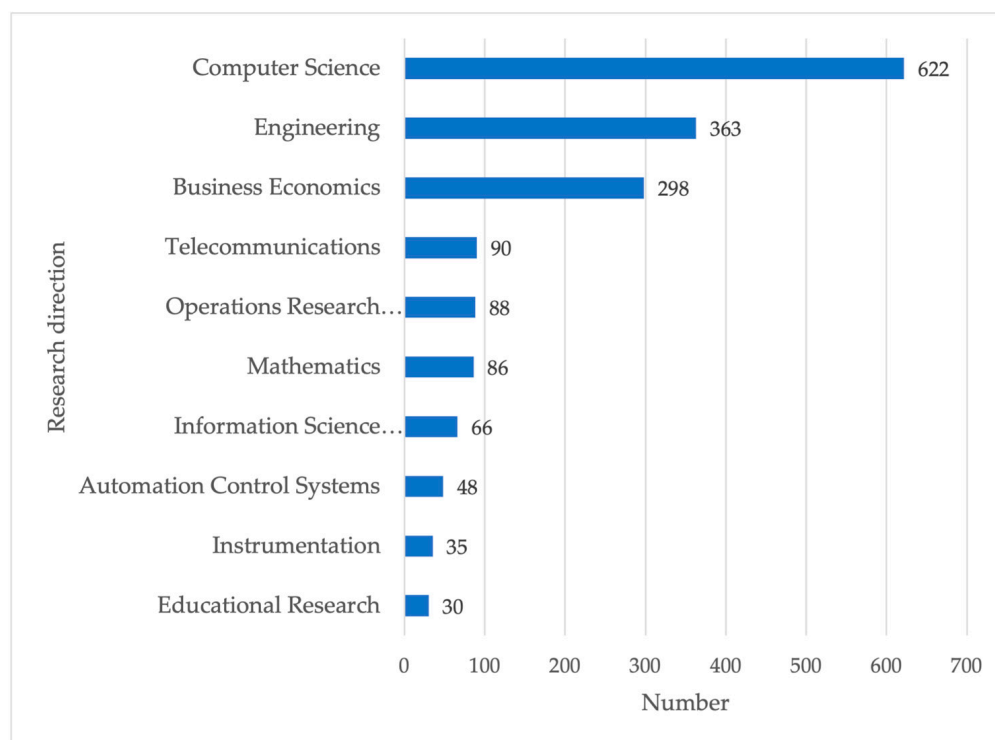


Figure 3. Distribution of the articles over research directions.

3. Definitions of Business Analytics

At present, there is still no uniform definition of business analytics. Scholars in different fields have defined the term business analytics from several perspectives. Holsapple concluded 18 definitions of analytics in 6 dimensions [2]. Referring to the dimensions, this article organizes recent definitions of business analytics into four categories in Table 1.

First, from the perspective of techniques, business analytics is considered an application of any data analytics [5] or data science [6] in business fields, which uses tools and techniques statistically and quantitatively to analyze a huge collection of data sources to support decisions for business [7]. More specifically, business analytics can be viewed as ‘a broad category of applications, technologies, and processes for gathering, storing, accessing, and analyzing data to help business users make better decisions’ [8]. With the continuous emergence of new technologies, business analytics can also be viewed as a combination of operation research, artificial intelligence (machine learning) and information systems [1].

Second, from the process perspective, business analytics is an encapsulation of tools to convert data into actionable insights through a scientific/mathematical/intelligent process [9]. The Institute for Operations Research and the Management Sciences (INFORMS) defined it as ‘a scientific process of transforming data into insight for making better decisions’ [10].

Third, from the practice perspective, business analytics is defined as ‘an ability of firms and organizations to collect, manage, and analyze data from a variety of sources to enhance the understanding of business processes, operations, and systems’ [11]. Business analytics

refers to ‘the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions’ [12].

Finally, from the perspective of the management, business analytics is a qualitative methodology to derive valuable meanings based on data [13] and is ‘a paradigm shifter of models, technologies, opportunities, and capabilities used to scrutinize a corporation’s data and performance to transpire data-driven decision-making analytics for the corporation’s future direction and investment plans’ [3].

Overall, regardless of the perspective from which it is defined, we can conclude that the implementers of business analytics are enterprises; the approaches to achieve business analytics are various techniques; and the ultima goal of business analytics is to improve enterprise values.

Table 1. Summary of definitions.

Category	Definition	Reference
Techniques	the general term for any data analytics in business problems	[5]
	data science in business	[6]
	a broad category of applications, technologies and processes for gathering, storing, accessing and analyzing data to help business users make better decisions	[8]
	the intersection of OR, artificial intelligence (machine learning) and information systems	[1]
Process	the encapsulation of all mechanisms that help convert data into actionable insight for better and faster decision-making	[9]
	a scientific process of transforming data into insight for making better decisions	[10]
Practice	an ability of firms and organizations to collect, manage and analyze data from a variety of sources in order to enhance the understanding of business processes, operations and systems	[11]
	the extensive use of data, statistical and quantitative analysis, explanatory and predictive models and fact-based management to drive decisions and actions	[12]
Management	one of the qualitative methodologies to derive valuable meanings based on data	[13]
	a paradigm shifter of models, technologies, opportunities and capabilities used to scrutinize a corporation’s data and performance to transpire data-driven decision-making analytics for the corporation’s future direction and investment plans	[3]

4. Techniques of Business Analytics

Business analytics is generally identified into three types: descriptive analytics, predictive analytics and prescriptive analytics [9]. Descriptive analytics is used to provide a summary of descriptive statistics as a straightforward presentation of facts. Predictive analytics is used to discover what is likely to happen in the future based on current data. Prescriptive analytics focuses on identifying optimal actions in the decision-making process. In 2013, the famous Gartner Group added more diagnostic analytics into the process of business analytics, which aims to answer why did it happen [14]. Since it is difficult to distinguish descriptive analytics and diagnostic analytics, which are both used to deal with historical data, we adopt the three-stages ideas used by [5,9] which consider descriptive analytics, predictive analytics and prescriptive analytics as three stages of business analytics. Figure 4 shows the summary of techniques mentioned in this section.

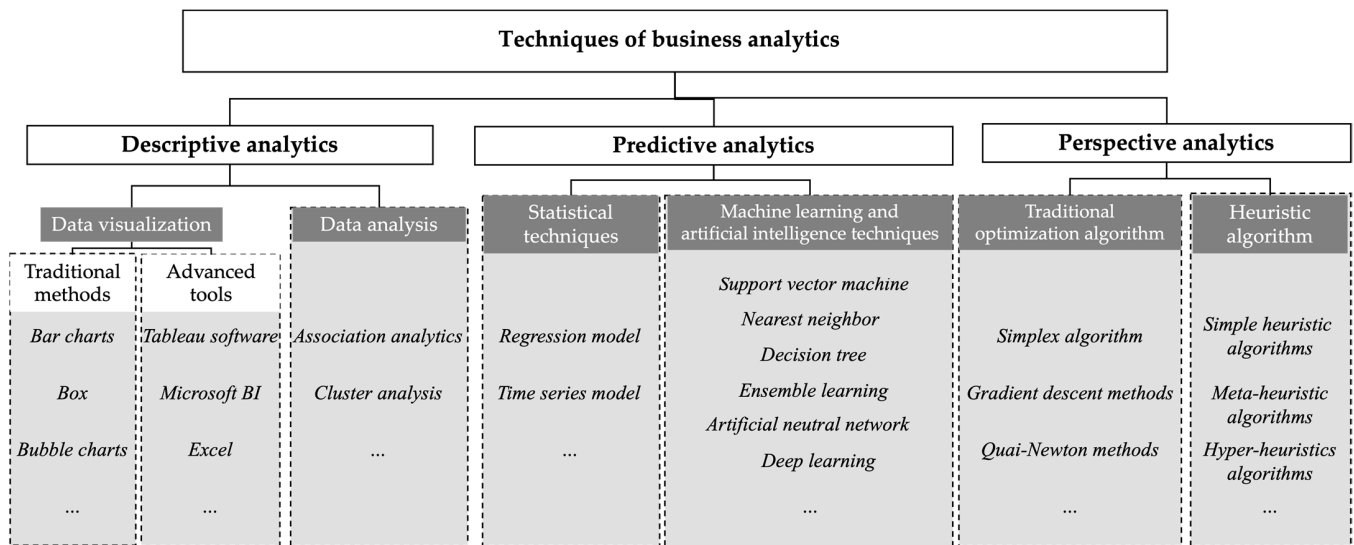


Figure 4. Summary of techniques mentioned in Section 4.

4.1. Descriptive Analytics Techniques

Descriptive statistics is a process of characterizing historical data. There are two core techniques of descriptive statistics: data visualization and data analysis. Data visualization produces graphical images of data or concepts, which helps decision making [15]. Data analysis consists of common statistical techniques, including mean, median, standard deviation, range, stem, histogram and advanced data mining techniques used to describe hidden patterns in the data.

4.1.1. Data Visualization

Over the years, many data visualization techniques have been developed to represent large amounts of information and examine them. These methods include bar charts, box, and whiskers, bubble charts, choropleth maps, dot distribution maps, histograms, line graphs, pie charts, population pyramids, proportional symbol maps, scatter plots, stacked bar charts and tree maps.

When working with data sets that include big data points, automation of the data visualization process makes the process much easier. Therefore, a large variety of data visualization tools are developed to create visual representations of large data sets, including Tableau Software 2022.4 [16], Microsoft Power BI [17], Excel, FusionCharts, Sisense, etc. In addition to the visualization tools as software, there are many online visualization tools such as Infogram, RAWGraphs, Sovit, etc.

Data visualization is a necessity in any data-driven business. It transforms data into visuals that are easier to understand, digest and make important business decisions from. Data visualization creates actionable insights that we may have not discovered. With the continuous development of technology, data visualization tools are gradually improving and developing towards a user-friendly and easy-to-use interface.

4.1.2. Data Analysis

Data analysis is to analyze the collected data and derive various quantitative characteristics reflecting objective phenomena. In addition to the traditional statistical methods of data concentration trend analysis, data dispersion analysis and data frequency distribution analysis, advanced data mining techniques probe more deeply into the underlying characteristics of data. Association and cluster analysis are two typical data analysis methods used in descriptive analytics.

- Association analysis

Association analysis, also called association rule mining, is an unsupervised algorithm that is used to mine potential association relationships from data. There are two classical algorithms in association analysis: Apriori Algorithm and Frequent Pattern tree (FP-tree). Apriori Algorithm uses an iterative method of searching the database level by level to find the relationships of item sets to form rules. Its process consists of concatenation and pruning. To improve the Apriori algorithm, many improvement methods are proposed including the Direct Hashing and Pruning (DHP) algorithm [18], Dynamic Itemset Counting (DIC) [19], Parallel Apriori algorithms based on various frameworks such as MapReduce [20–22], Spark [23,24] and Flink [25] and adaptive Apriori algorithms [26]. Compared to the Apriori algorithm, the FP-tree algorithm only requires two scans of the database when performing frequent pattern mining and does not generate candidate item sets. There are various improvement algorithms based on FP-tree, such as QFP-growth [27], fuzzy FP-tree [28], PFP [29], balanced parallel FP-tree (BPPF) [30] and tree partition based parallel FP-tree [31].

- Cluster analysis

Cluster analysis is a multivariate statistical analysis method for classifying samples or indicators. The clustering algorithms can be divided into five categories: partitioning-based, hierarchical-based, density-based, grid-based and model-based.

Partitioning-based algorithms include K-means [32], Fuzzy C-means (FCM) [33], K-medoids [34], CLARA (Clustering Large Applications) [34], K-modes [35], and CLARANS (Clustering Large Applications based on a RANdomized Search) [36]. Hierarchical clustering algorithms include BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) [37], (CURE) [38], ROCK (Robust Clustering using Links) [39] and Chameleon (clustering using interconnectivity) [40]. DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is the first density-based clustering algorithm [41]. In addition, DENCLU (DENSITY-based CLUstEring) [42] and OPTICS (Ordering Points to Identify the Clustering Structure) [43] are both widely used in cluster analysis. The typical grid-based algorithms include STING (Statistical Information Grid) [44], CLIQU (Clustering in Quest) [45] and WaveCluster [46]. There are usually two attempted ideas in the model-based algorithm: statistical methods and neural network methods. Among them, statistical methods are the COBWEB algorithm [47], GMM (Gaussian Mixture Model) [48] and the neural network algorithm is the SOM (Self Organized Maps) algorithm [49].

In conclusion, description analysis shows the overall picture and characteristics of the data and lays the foundation for subsequent analysis of business analytics. Descriptive analysis is a well-researched field with many classical algorithms and visualization tools available. With the advent of the big data era, there are greater challenges for the processing speed and capacity of algorithms and visualization platforms, and based on the continuous efforts of researchers, many effective improvements to classical algorithms and new algorithms have been proposed.

4.2. Predictive Analytics Techniques

In general, predictive analytics techniques can be divided into statistical techniques and machine learning techniques. Statistical methods to predict mainly refer to building suitable forecasting models and estimating model parameters, listing forecasting formulas and thus making extrapolated forecasts. In machine learning techniques, systems are trained to use specialized algorithms to study, learn and make predictions and recommendations based on large amounts of data.

4.2.1. Statistical Techniques

In statistical predictive techniques, statistical theories and methods are used for prediction by building statistical models and fitting the model parameters with past data. There are two groups of methods in statistical predictive techniques: regression models and time series models.

- Regression model

Regression model is one of the most famous statistical techniques used to predict. The linear regression model is the basic model which is represented as an equation that finds specific weights for the input variables, which in turn describe a straight line that best fits the relationship between the input variables and the output variables [50]. When the output variable is a categorical variable, a classification model such as the logistic regression model [51] is needed. Meanwhile, polynomial regression models are used to fit nonlinear relationships between variables.

- Time series model

Time series models can be divided into two groups: exponential smoothing models and ARIMA series models [52]. The exponential smoothing model decomposes the time series into components and uses an additive or multiplicative structure to reassemble the smoothed components to predict future values [53]. Typical exponential smoothing models include simple exponential smoothing, Holt's exponential smoothing and Holt-Winters' seasonal exponential smoothing [54]. The ARIMA series model mainly includes AR (AutoRegressive) model, MA (Moving Average) model, ARMA (AutoRegressive Moving Average) model, ARIMA (AutoRegressive Integrated Moving Average) model and SARIMA (seasonal ARIMA) model [55].

4.2.2. Machine Learning and Artificial Intelligence Techniques

With the advent of the big data era, machine learning to guide predictive analytics has become a widely used approach. There are many classic machine learning prediction algorithms, such as support vector machine, nearest neighbor, decision tree, ensemble learning and artificial neural network and more advanced deep learning techniques.

- Support vector machine

In machine learning, the support vector machine (SVM) is a supervised learning model for analyzing data in classification and regression analysis with associated learning algorithms [56]. Given a set of training instances, each labeled as belonging to one or the other of two classes, the SVM training algorithm builds a model that assigns new instances to one of the two classes. Thus, it is usually used to predict binary classification problems.

- Nearest neighbor

The k-nearest neighbor algorithm, known as KNN, is a non-parametric supervised learning classifier [57]. It can be applied to classification problems or regression problems. In a classification problem, the output is a member of a category, and in a regression problem, the output is the value of an object's attributes. The nearest neighbor is considered the simplest type of machine learning algorithm [58].

- Decision tree

Decision tree is a non-parametric supervised learning algorithm for classification and regression tasks. It is a hierarchical tree structure consisting of a root node, branches, internal nodes and leaf nodes. There are three typical decision tree algorithms: ID3, C4.5 and CART (Classification and Regression Tree). Iterative Dichotomiser 3 (ID3) uses information entropy and information gain as metrics to evaluate candidate splitting [59]. C4.5 is an improved version of ID3, which does not use the information gain directly but introduces the information gain ratio metric as the basis for feature selection [60]. The CART algorithm uses the Gini coefficient instead of the information entropy model, which can be used for classification and regression problems [61].

- Ensemble learning

The basic idea of the Ensemble Learning algorithm is combining multiple classifiers to achieve an integrated classifier with better prediction. Ensemble learning includes the bagging method, boosting method and stacking method.

Bagging is a technique to reduce the generalization error by combining several models. The main idea is to train several different models separately and then let all models vote on the output of the test samples. Random Forests is a bagging algorithm using decision trees as the base learner [62].

Boosting is a framework algorithm that mainly obtains a subset of samples by manipulating the sample set and then generates a series of base classifiers by training on the sample subset with a weak classification algorithm [63]. Adaboost [64] and GBDT (Gradient Boosting Decision Tree) [65] are representative boosting algorithms. Based on GBDT, XGBoost [66], LightGBM [67] and CatBoost [68] are proposed to improve and achieve higher accuracy predictions.

Stacking approach is a hierarchical ensemble framework that takes the output of a series of models and inputs them as new features into other models [69]. The base classifier of stacking can be any predictive algorithm mentioned above.

Predictive analytics is an essential part of the business analytics process. It predicts the future by analyzing past data, and the results of its predictions are an important foundation for prescriptive analytics. Predictive analytics algorithms have evolved rapidly in recent years, gradually expanding from traditional statistical predictive methods to machine learning algorithms. In particular, among the machine learning algorithms, deep learning prediction algorithms are currently a hotly researched and fast-developing area.

- Artificial Neural network

Artificial Neural network (ANN) is a model that mimics the structure and function of biological neural networks, especially in the brain [70]. According to the connectionism of networks, ANN can be divided into feed-forward neural networks and feedback neural networks. Feedforward neural networks (FNN) divide each neuron into different groups according to the order of receiving information, and each group can be considered as a neural layer [71]. The neurons in each layer receive the output of the neurons in the previous layer and output to the neurons in the next layer. FNN has two categories depending on the number of layers: single-layer and multi-layer networks [72]. Single-layer FNN is also known as fully connected feedforward neural networks (FC), and a typical multi-layer network is the convolutional neural network (CNN). In feedback neural networks, neurons can receive signals from other neurons and their own feedback signals. Compared with feedforward neural networks, the neurons in feedback neural networks have a memory function and have different states at different moments. Common feedback neural networks include recurrent neural networks (RNN) [73], Hopfield networks [74] and Boltzmann machines [75].

- Deep learning

The concept of deep learning originates from the study of artificial neural networks, and a multilayer perceptron with multiple hidden layers is a deep learning structure. Recently, deep learning has been widely used in predictive analytics, including RNN, CNN, Transformer and Nbeats. LSTM is a well-known RNN algorithm used in prediction [76]. DeepAR employs a classical RNN model to solve the time series forecasting problem [77], and Deep state space model is proposed to improve DeepAR limitations [78]. Since DeepAR and Deep state space model are both one-horizon forecast models, MQRNN (multi-horizon forecast model) is designed to simultaneously predict for multiple future time steps [79]. The CNN-LSTM algorithm, which combines CNN and LSTM, has been applied in many predictive analyses [80–82]. Transformer model was first proposed by Google in 2017 [83] and was improved to deal with time-series data prediction in 2019 [84]. Since then, there have been various transformer-based algorithms proposed [85–87]. N-BEATS is a different deep learning algorithm proposed in 2020 [88]. In N-BEATS, there is no RNN or CNN in the internal structure, and the network is composed of all fully connected structures.

4.3. Prescriptive Analytics

Prescriptive analytics is the final step of business analytics. Prescriptive analytics mainly refers to the use of operations research methods such as mathematical programming models and intelligent optimization algorithms to give recommendations on the optimal actions that an enterprise should take. Compared to the traditional decision methods which rely too much on human experience, prescriptive analytics gives more reliable and reasonable decisions through scientific approaches including traditional optimization algorithms and heuristic algorithms.

4.3.1. Traditional Optimization Algorithm

Based on the features of the objective function, constraints and decision variables, mathematical programs can be divided into linear programming, nonlinear programming, integer programming, stochastic programming, dynamic programming and so on [89]. In order to solve these problems, many traditional optimization algorithms are proposed. For constrained programming, Simplex algorithm is a well-known linear programming algorithm [90], and penalty-series methods are proposed for nonlinear programming. Gradient Descent Method [91], Quasi-Newton Method [92] and Conjugate gradient method [93] are classical iteration algorithms for unconstrained optimizations.

4.3.2. Heuristic Algorithm

In the face of complex optimization problems, traditional optimization methods require traversing the entire search space, which cannot be completed in a short time. Inspired by human intelligence, the social nature of biological groups, or the laws of natural phenomena, many heuristic algorithms have been invented to solve these complex optimization problems.

- Simple Heuristic Algorithms

Simple heuristic algorithms mainly contain greedy algorithms, local search algorithms and hill-climbing algorithms. The greedy algorithm is an algorithm that takes the optimal choice in the current state at each step of the selection process, thereby hopefully leading to the best or optimal outcome [94]. The local search algorithm is based on the greedy idea of starting with a candidate solution and continuously searching in its neighborhood until there are no better solutions in the neighborhood [95]. The hill-climbing algorithm is a simple greedy search algorithm that selects one optimal solution at a time as the current solution from the proximity solution space of the current solution until a local optimal solution is reached [96].

- Meta-heuristic algorithms

Meta-Heuristic algorithms are improvements of simple heuristic algorithms, usually using randomized search techniques, and can be applied to a wide range of problems. Meta-heuristic algorithms include Evolutionary Algorithms, Swarm Intelligence algorithms, Simulated Annealing algorithms and Tabu Search algorithms. Evolutionary algorithms are inspired by the evolutionary mechanisms of living organisms and simulate the evolutionary processes to conduct evolutionary calculations on the candidate solutions of optimization problems. Typical evolutionary algorithms are Genetic Algorithm (GA), Differential Evolution (DE) and Immune Algorithm (IM). Swarm intelligence refers to the property of unintelligent subjects to exhibit intelligent behavior through cooperation and is a computational technique based on the behavioral laws of biological groups. Two representative swarm intelligence algorithms are Particle Swarm Optimization (PSO) [97] and ACO (ant colony optimization) [98]. Simulated Annealing is an algorithm that solves the global optimum by finding states with relatively small objective values in the neighborhood [99]. Tabu search algorithm searches for the optimal solution of the target by searching for a better solution in the solution neighborhood and puts the search history into a Tabu List during the search process to avoid duplicate searches [100].

Furthermore, the artificial neural network is also used in the optimization field. Google’s DeepMind used a neural network to solve mixed integer programs (MIP) [101]. The well-known Pointer Network (PN) solved some classical combinatorial optimization problems, such as the Traveling Traders Problem (TSP) and Knapsack problem [102]. Graph Neural Networks are also used to deal with combinatorial optimization problems [103]. Furthermore, deep reinforcement learning (DRL) algorithms have been widely applied in solving optimization problems in recent years [104,105].

- Hyper-Heuristic algorithms

Hyper-Heuristic algorithms provide a high-level heuristic by managing or manipulating a set of Low-Level Heuristics (LLH) to generate new heuristics. These new heuristics are used to solve various combinatorial optimization problems.

4.4. Summary of Techniques

The three types of data analytics mentioned above represent the process of business analytics. Descriptive analytics is the first stage used to understand the relationship and pattern of historical data. Then, predictive analytics is the next stage, which uses historical data and predictive techniques to forecast future trends and events. The final stage is perspective analytics, which uses optimization or machine learning methods to make the best decisions and actions. The running of each stage is based on the precise analytics of the previous stage. Following the three stages in order, it is easy to realize business analytics and achieve the goal of improving enterprise value.

Furthermore, reviewing all the techniques in the three parts of business analytics, there are two main categories: traditional statistical techniques and more advanced techniques (e.g., machine learning). There is no doubt that statistical techniques are the most classical and are the foundation of advanced techniques. However, the limitations of classical techniques are unavoidable, especially when dealing with complex and big data problems. Machine learning techniques are widely used in business analytics nowadays and still improve and update at a rapid speed. Especially, the emergence of deep learning and reinforcement learning brings a large shock to predictive analytics and perspective analytics techniques. However, there are worries about the interpretability of machine learning because machine learning models remain mostly as black boxes. Since the three steps of business analysis have different purposes, we conduct an internal comparison of the techniques involved in each stage and summarize the advantages and disadvantages of all techniques in Table 2. Choosing the most suitable techniques is the key point to realizing effective business analytics.

Table 2. Advantages and disadvantages of techniques.

Technique		Advantages	Disadvantages
Descriptive analytics techniques	Data visualization	Traditional method	<ul style="list-style-type: none"> • Easy to use • Inability to uncover deep relationships between data
		Visualization tools	<ul style="list-style-type: none"> • Can deal with large data set • Inability to uncover deep relationships between data
	Data analysis	Association analysis	<ul style="list-style-type: none"> • Mine potential association relationships • Support indirect data mining • When the problem becomes larger, the computational volume grows larger
		Cluster analysis	<ul style="list-style-type: none"> • Can handle big data • Fast and identifiable noise points • Poor clustering of high-dimensional data features

Table 2. Cont.

Technique		Advantages	Disadvantages	
Statistical techniques	Regression model	<ul style="list-style-type: none"> • Easy to use and explain 	<ul style="list-style-type: none"> • Cannot handle a large number of multi-class features or variables well • Prone to underfitting 	
	Time series model	<ul style="list-style-type: none"> • Low complexity • Fast calculation speed 	<ul style="list-style-type: none"> • Low accuracy rate in complex scenes 	
Predictive analytics techniques	Machine learning and artificial intelligence techniques	support vector machine	<ul style="list-style-type: none"> • Can map to higher dimensional spaces • Can solve the classification of nonlinear 	<ul style="list-style-type: none"> • Large-scale training samples are difficult to implement • Difficulties in solving multi-classification problems • Sensitive to missing data • Sensitive to the choice of parameters and kernel functions
		Nearest neighbor	<ul style="list-style-type: none"> • Simple and no need to estimate parameters • Can be used for non-linear classification • Short training time • High accuracy and insensitivity to outliers 	<ul style="list-style-type: none"> • Huge computation • Poor interpretability • Low prediction accuracy for rare categories when the sample is unbalanced • Poor fault tolerance for training data
	Machine learning and artificial intelligence techniques	Decision tree	<ul style="list-style-type: none"> • High interpretability • Fast speed • Insensitive to missing values • Can handle both continuous and discrete data • Suitable for high-dimensional data 	<ul style="list-style-type: none"> • No support for online learning • Prone to overfitting
		Ensemble learning	<ul style="list-style-type: none"> • Preventing underfitting • Preventing overfitting 	<ul style="list-style-type: none"> • Insufficient amount of data will lead to poor generalization of the trained model
		Artificial Neural Network	<ul style="list-style-type: none"> • Less formal statistical training to develop • Implicitly detect complex nonlinear relationships between independent and dependent variables • The ability to detect all possible interactions between predictor variables • Being developed using multiple different training algorithms 	<ul style="list-style-type: none"> • Limited ability to explicitly identify possible causal relationships • Requirement of greater computational resources • Prone to overfitting
		Deep learning	<ul style="list-style-type: none"> • High learning ability • Wide coverage and good adaptability 	<ul style="list-style-type: none"> • High computational volume and poor portability • High hardware requirements • Complex model design • Data-dependent and not highly interpretable
Prescriptive analytics techniques	Traditional optimization algorithm	Simplex algorithm	<ul style="list-style-type: none"> • No requirement for the analyticity of the objective function • Fast convergence speed 	<ul style="list-style-type: none"> • Difficulties in dealing with large scale high latitude problems
		Gradient Descent Method	<ul style="list-style-type: none"> • Simple implementation 	<ul style="list-style-type: none"> • Easy to fall into local minima • Prone to overfitting
		Quasi-Newton Method	<ul style="list-style-type: none"> • Fast convergence speed 	<ul style="list-style-type: none"> • Strict requirements for the objective function • Large computing and storage capacity
	Heuristic algorithm	<ul style="list-style-type: none"> • Able to give a better solution in an acceptable amount of time 	<ul style="list-style-type: none"> • Not guaranteed to be globally optimal • Algorithm instability 	

5. Business Analytics Applications

In the course of the literature review, we found that business analytics has been very comprehensively applied. Business analytics applications can be classified into two dimensions: functional areas and industry sectors. From the perspective of functional areas, applications include supply chain management, marketing management, risk management, strategic management, management accounting and human resources management. From the perspective of industry sectors, business analytics is mainly used in healthcare, circular economy, retail, financial and professional sports organizations.

5.1. Applications in Functional Areas

Supply chain management is a representative application of business analytics in the business area. Business analytics has a strong impact on the supply chain performance in the plan, source, make and deliver area [106–108]. For example, descriptive analysis helps to identify demand patterns and predict analysis forecasts customer demand in the future through statistical and machine learning algorithms. Based on the predictions, optimization algorithms are used to make pricing and inventory management decisions to maximize retailers' profit.

In the area of marketing management, business analytics integrates market and customer-related data and uses analysis algorithms to provide managers with a variety of relevant perspectives for better optimization decisions. Among the various areas of marketing, customer relationship management (CRM) is a key area that uses business analytics to analyze, integrate and utilize information resources and customer feedback to support CRM technology, such as acquiring and retaining customers [109]. Furthermore, recommending systems that suggest what products to buy based on personal preferences and past behaviors is also an important application of business analytics in marketing, especially in the e-commerce field [110].

Risk management is an essential area of company management, and business analytics techniques are widely used in the process of risk management. Predict analysis techniques such as artificial neural networks and support vector machines are applied to establish the early warning system [111,112] and risk evaluation [113,114]. Optimization tools of perspective analysis are used to make better risk-based decisions [115].

Strategic management plays an important role in the business area to create or sustain competitive advantages of an enterprise, which consists of analyses, decisions and actions undertaken. Business analytics helps firms to reveal their strengths and weaknesses by identifying business units, activities and processes [116].

Business analytics can be helpful in various tasks of managerial accounting. Several papers discuss the impact of business analytics on managerial accounting [117–119]. For financial reporting, descriptive analytics helps summarize and describe the financial position of an organization. In the tasks of performance measurement, management accountants can utilize predictive analytics techniques such as machine learning algorithms with inputs from descriptive analytics to forecast future performance. With the results of cost accounting and performance measurement, prescriptive analytics are incorporated into planning and decision-making to provide decision-makers with information about optimal solutions.

The emergency of business analytics drives the development of data-driven human resources (HR) management [120]. Human resources management is progressively increasing its adoption of advanced data analytics, visualization models and techniques to strengthen strategic decision-making and serve the needs of decision-makers. Descriptive analytics uses internal and external organizational data and HR administrative information to generate ratios, metrics, dashboards and reports on HR. Predictive analytics can analyze process data and make predictions. Based on predictive analytics and the large and diverse HR data available, HR departments gain decision options to optimize performance and completely reshape the decision-making process [121]. For example, Pape describes a framework for prioritizing data items for business analytics and applying it to human resources [122].

5.2. Applications in Industry Sectors

Business analytics is widely used in the healthcare sector. Data visualization tools such as dashboards and control charts are used to monitor outcomes and look for variations in process [123]. Descriptive analytics techniques are used to mine genetic data to identify the relationships between human genes, diseases, variants, proteins, cells and biological pathways [124]. Predictive analytics methods help to forecast the emergency and development of diseases [125]. The application of perspective algorithms can increase efficiency and reduce costs in the healthcare industry [126].

The circular economy is defined as ways to improve economic performance without depleting resources at a rate that exceeds the Earth's capacity [127]. Several researchers have recognized the positive relationship between business analytics and the circular economy. BA can connect the required material and information flow to help understand and enact circular material flows, enhance and expand the use of products and components and recycle waste materials [128]. Thus, business analytics capability can improve an enterprise's ability to operate a circular strategy and overall circular economy implementation [128–130].

The retail industry has various applications of business analytics. Retailers can collect customer demographics and behavior data to analyze customer preferences and shopping features through business analytics. The classical one is the market basket analysis using data mining methods to examine large transaction databases and determine which items are most frequently purchased [131,132]. Customer visit segments can be mined by data mining rules [133]. Business analytics techniques are also used in the establishment of recommend systems, especially in the e-commerce fields [134,135].

Except for the industry sector mentioned above, business analytics is also applied to the financial industry and professional sports organizations. Business analytics helps the financial industry to build effective corporate financial distress prediction models to measure and manage financial sustainability [136]. In the management of professional sports organizations, business analytics can be used to drive improvements in ticket pricing, customer retention, lead scoring, sponsorship, premium sales, digital marketing, food and beverage, merchandising and fan/game experience [137].

6. Challenges in Business Analytics

Business analytics is rapidly evolving and has been used across a wide range of industrial sectors and business areas, and there is a large body of research that demonstrates the effectiveness of business analytics in increasing enterprise value. However, there are still some challenges and opportunities in business analytics, such as the research for data quality, data security and privacy problems.

6.1. Data Quality

With the advent of the Big Data era, the accessibility of data and the volume of data available have increased significantly compared to the past. However, the problem that arises is how to select useful and accurate data for analytics from the vast amount of information. Machine learning plays an important role in business analytics, which relies on data. Thus, business analytics can be considered a data-driven analytics process; so, data quality is very important for subsequent analysis and guidance. In business analytics, data quality challenges mainly include data completeness, consistency and accuracy.

Data completeness refers to the presence or absence of missing data information. Missing data can be the absence of an entire data record or the absence of a record of a field of information in the data. The value that can be drawn from incomplete data is greatly reduced. For raw data containing missing data, we can choose to fill them with specific values or just delete them. If deleted, a part of the sample information will be lost. If the padding is inappropriate, it will add noise to the sample. Thus, efficient missing value handling approaches are a challenge for business analytics.

Data consistency refers to whether the data follows a uniform structure. It is a vital factor in business analytics as is heterogeneity of data. Heterogeneity of data means that raw data contain structured, semi-structured and unstructured data, such as text data, graph data and time-series data [138]. There are two types of efforts to address these challenges. The first type is the transfer of unstructured data to a structured format. Then, all the classical methods can be used for the transferred data. The second type is to develop new methods to deal with unstructured data, such as unsupervised learning algorithms.

Data accuracy refers to anomalies or errors in the information recorded in the data. Common data accuracy errors include garbled data and abnormally large or small data. There are various outlier detection algorithms, each with its advantages, disadvantages

and scope of application, and it is difficult to directly determine which one is the best. In practical applications, an appropriate outlier detection algorithm is selected according to the characteristics of business operations, such as the requirements for computational volume and tolerance for outliers.

6.2. Data Security and Privacy

The emergency of Big Data has made data analysis and application more complex and difficult to manage. The increase in data makes data security and privacy protection increasingly important.

There is no completely secure data infrastructure unless it is isolated and disconnected from all other networks. However, this is impossible for business analytics, especially when cloud computing emerges [9]. Throughout the data lifecycle, enterprises need to comply with stricter security standards and confidentiality regulations; therefore, the security requirements for data storage and use are increasingly high. Traditional data protection methods often cannot meet the new changes in the network and digital life. As a result, there are more criminal means that cannot be easily traced and prevented, while the existing laws and regulations and technical means can hardly overcome such problems.

Meanwhile, the security needs of data are changing, and a new complete chain has been formed from data collection, data integration, data refinement, data mining, security analysis, security posture determination and security detection to threat discovery. In this chain, data may be lost, leaked, accessed by unauthorized access, tampered with, or even involved in user privacy and corporate secrets. Therefore, data security protection in the big data environment is a significant challenge for business analytics. From the perspective of customers, there are concerns about the privacy of individuals. The use of the personal data of customers, even within the limits of the law, should be avoided or scrutinized to prevent the organization from adverse effects and public condemnation.

7. Conclusions

This review summarizes the existing literature on business analytics in terms of definitions, techniques, applications and challenges to answer the four questions proposed in Section 1. For RQ1, in terms of definition, business analytics has been defined in four main types: as an integration of technology, as a process of transforming data to results, as a capability of an enterprise or people and as a paradigm of management. Although there is no agreed definition of business analytics, there is a basic agreement that the purpose of business analytics is to improve the accuracy and efficiency of decision-making. As for RQ2, business analytics can be divided into three steps: descriptive analytics, predictive analytics and perspective analytics. For each step, there are various techniques concluded in this paper. According to the latest literature, artificial intelligence techniques are gradually becoming a trend for the future development of business analytics, especially in the era of big data. Compared to traditional statistical methods, artificial intelligence algorithms can process data and produce results more efficiently.

For RQ3, based on various techniques, business analytics has been applied in a large number of business areas and industry sectors. It can help supply chain management, marketing management, risk management, strategic management, management accounting and human resources management and is widely used in healthcare, circular economy, retail and professional sports organizations. With regard to RQ4, although business analytics has been studied for several years, some challenges still need to be solved. Since business analytics is a data-driven process, data quality plays an important role in the success of decision-making. Meanwhile, the issue of data security and privacy is a big challenge for both enterprises and customers in the usage of business analytics.

In conclusion, business analytics is a common approach for enterprises to use historical data to drive optimal decisions and to create large business value. We believe that business analytics has a great future ahead, especially with the rapid development of technology, enabling more dimensional data analytics to support final decision making. However, the

challenges it faces will limit the further development of business analytics to a certain extent or even have a negative impact if they are not effectively addressed. Therefore, we believe that there are still many elements and opportunities worth exploring in the field of business analytics in the future, such as new methods to deal with data completeness, data consistency and data accuracy, or new techniques to preserve privacy in the use and integration of data.

Author Contributions: Conceptualization, methodology, S.L. and O.L.; writing—original draft preparation, S.L.; writing—review and editing, O.L. and J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Mortenson, M.J.; Doherty, N.F.; Robinson, S. Operational Research from Taylorism to Terabytes: A Research Agenda for the Analytics Age. *Eur. J. Oper. Res.* **2015**, *241*, 583–595. [\[CrossRef\]](#)
- Holsapple, C.; Lee-Post, A.; Pakath, R. A Unified Foundation for Business Analytics. *Decis. Support Syst.* **2014**, *64*, 130–141. [\[CrossRef\]](#)
- Bayrak, T. A Review of Business Analytics: A Business Enabler or Another Passing Fad. *Procedia-Soc. Behav. Sci.* **2015**, *195*, 230–239. [\[CrossRef\]](#)
- Harzing, A.-W.; Alakangas, S. Google Scholar, Scopus and the Web of Science: A Longitudinal and Cross-Disciplinary Comparison. *Scientometrics* **2016**, *106*, 787–804. [\[CrossRef\]](#)
- Duan, L.; Xiong, Y. Big Data Analytics and Business Analytics. *J. Manag. Anal.* **2015**, *2*, 1–21. [\[CrossRef\]](#)
- Chen, H.; Chiang, R.H.; Storey, V.C. Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Q.* **2012**, *36*, 1165–1188. [\[CrossRef\]](#)
- Delen, D.; Zolbanin, H.M. The Analytics Paradigm in Business Research. *J. Bus. Res.* **2018**, *90*, 186–195. [\[CrossRef\]](#)
- Watson, H.J. Tutorial: Business Intelligence—Past, Present, and Future. *CAIS* **2009**, *25*, 39. [\[CrossRef\]](#)
- Delen, D.; Ram, S. Research Challenges and Opportunities in Business Analytics. *J. Bus. Anal.* **2018**, *1*, 2–12. [\[CrossRef\]](#)
- INFORMS. *Certified Analytics Professional Handbook*; INFORMS: Catonsville, MD, USA, 2016.
- Kraus, M.; Feuerriegel, S.; Oztekin, A. Deep Learning in Business Analytics and Operations Research: Models, Applications and Managerial Implications. *Eur. J. Oper. Res.* **2020**, *281*, 628–641. [\[CrossRef\]](#)
- Davenport, T.H.; Harris, J.G. Competing on Analytics: The New Science of Winning. *Language* **2007**, *15*, 24.
- Lee, C.S.; Cheang, P.Y.S.; Moslehpour, M. Predictive Analytics in Business Analytics: Decision Tree. *Adv. Decis. Sci.* **2022**, *26*, 1–29.
- Silva, A.J.; Cortez, P.; Pereira, C.; Pilastri, A. Business Analytics in Industry 4.0: A Systematic Review. *Expert Syst.* **2021**, *38*, e12741. [\[CrossRef\]](#)
- Ware, C. *Information Visualization: Perception for Design*; Morgan Kaufmann: Burlington, MA, USA, 2019; ISBN 0-12-812876-3.
- Batt, S.; Grealis, T.; Harmon, O.; Tomolonis, P. Learning Tableau: A Data Visualization Tool. *J. Econ. Educ.* **2020**, *51*, 317–328. [\[CrossRef\]](#)
- Becker, L.T.; Gould, E.M. Microsoft Power BI: Extending Excel to Manipulate, Analyze, and Visualize Diverse Data. *Ser. Rev.* **2019**, *45*, 184–188. [\[CrossRef\]](#)
- Park, J.S.; Chen, M.-S.; Yu, P.S. An Effective Hash-Based Algorithm for Mining Association Rules. *ACM Sigmod Rec.* **1995**, *24*, 175–186. [\[CrossRef\]](#)
- Brin, S.; Motwani, R.; Ullman, J.D.; Tsur, S. Dynamic Itemset Counting and Implication Rules for Market Basket Data. In Proceedings of the 1997 ACM SIGMOD International Conference on Management of Data, Tucson, AZ, USA, 13–15 May 1997; pp. 255–264.
- Yang, X.Y.; Liu, Z.; Fu, Y. MapReduce as a Programming Model for Association Rules Algorithm on Hadoop. In Proceedings of the 3rd International Conference on Information Sciences and Interaction Sciences, Chengdu, China, 23–25 June 2010; pp. 99–102.
- Li, N.; Zeng, L.; He, Q.; Shi, Z. Parallel Implementation of Apriori Algorithm Based on Mapreduce. In Proceedings of the 2012 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, Kyoto, Japan, 8–10 August 2012; pp. 236–241.
- Sornalakshmi, M.; Balamurali, S.; Venkatesulu, M.; Krishnan, M.N.; Ramasamy, L.K.; Kadry, S.; Lim, S. An Efficient Apriori Algorithm for Frequent Pattern Mining Using Mapreduce in Healthcare Data. *Bull. Electr. Eng. Inform.* **2021**, *10*, 390–403. [\[CrossRef\]](#)

23. Qiu, H.; Gu, R.; Yuan, C.; Huang, Y. Yafim: A Parallel Frequent Itemset Mining Algorithm with Spark. In Proceedings of the 2014 IEEE International Parallel & Distributed Processing Symposium Workshops, Phoenix, AZ, USA, 19–23 May 2014; pp. 1664–1671.
24. Rathee, S.; Kaul, M.; Kashyap, A. R-Apriori: An Efficient Apriori Based Algorithm on Spark. In *PIKM '15 Proceedings of the 8th Workshop on Ph.D. Workshop in Information and Knowledge Management, Melbourne, Australia, 19 October 2015*; ACM: New York, NY, USA, 2015; pp. 27–34.
25. Akil, B.; Zhou, Y.; Röhm, U. On the Usability of Hadoop MapReduce, Apache Spark & Apache Flink for Data Science. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; pp. 303–310.
26. Patil, S.D.; Deshmukh, R.R.; Kirange, D.K. Adaptive Apriori Algorithm for Frequent Itemset Mining. In Proceedings of the 2016 International Conference System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 25–27 November 2016; pp. 7–13.
27. Qiu, Y.; Lan, Y.-J.; Xie, Q.-S. An Improved Algorithm of Mining from FP-Tree. In Proceedings of the 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.04EX826), Shanghai, China, 26–29 August 2004; Volume 4, pp. 1665–1670.
28. Lin, C.-W.; Hong, T.-P.; Lu, W.-H. Linguistic Data Mining with Fuzzy FP-Trees. *Expert Syst. Appl.* **2010**, *37*, 4560–4567. [[CrossRef](#)]
29. Li, H.; Wang, Y.; Zhang, D.; Zhang, M.; Chang, E.Y. Pfp: Parallel Fp-Growth for Query Recommendation. In *RecSys '08 Proceedings of the 2008 ACM conference on Recommender systems, Lausanne, Switzerland, 23–25 October 2008*; ACM Press: Lausanne, Switzerland, 2008; p. 107.
30. Zhou, L.; Zhong, Z.; Chang, J.; Li, J.; Huang, J.Z.; Feng, S. Balanced Parallel FP-Growth with MapReduce. In Proceedings of the 2010 IEEE Youth Conference on Information, Computing and Telecommunications, Beijing, China, 28–30 November 2010; pp. 243–246.
31. Chen, D.; Lai, C.; Hu, W.; Chen, W.; Zhang, Y.; Zheng, W. Tree Partition Based Parallel Frequent Pattern Mining on Shared Memory Systems. In Proceedings of the 20th IEEE International Parallel & Distributed Processing Symposium, Rhodes Island, Greece, 25–29 April 2006; p. 8.
32. MacQueen, J. Classification and Analysis of Multivariate Observations. In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 21 June–18 July 1967; pp. 281–297.
33. Bezdek, J.C.; Ehrlich, R.; Full, W. FCM: The Fuzzy c-Means Clustering Algorithm. *Comput. Geosci.* **1984**, *10*, 191–203. [[CrossRef](#)]
34. Kaufman, L.; Rousseeuw, P.J. (Eds.) *Finding Groups in Data*; Wiley Series in Probability and Statistics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1990; ISBN 978-0-470-31680-1.
35. Huang, Z.; Ng, M.K. A Fuzzy K-Modes Algorithm for Clustering Categorical Data. *IEEE Trans. Fuzzy Syst.* **1999**, *7*, 446–452. [[CrossRef](#)]
36. Ng, R.T.; Han, J. Efficient and Effective Clustering Methods for Spatial Data Mining. In *VLDB'94 Proceedings of the 20th International Conference on Very Large Data Bases, Santiago de Chile, Chile, 12–15 September 1994*; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1994; pp. 144–155.
37. Zhang, T.; Ramakrishnan, R.; Livny, M. BIRCH: An Efficient Data Clustering Method for Very Large Databases. *ACM Sigmod Rec.* **1996**, *25*, 103–114. [[CrossRef](#)]
38. Guha, S.; Rastogi, R.; Shim, K. CURE: An Efficient Clustering Algorithm for Large Databases. *ACM Sigmod Rec.* **1998**, *27*, 73–84. [[CrossRef](#)]
39. Guha, S.; Rastogi, R.; Shim, K. ROCK: A Robust Clustering Algorithm for Categorical Attributes. *Inf. Syst.* **2000**, *25*, 345–366. [[CrossRef](#)]
40. Karypis, G.; Han, E.-H.; Kumar, V. Chameleon: Hierarchical Clustering Using Dynamic Modeling. *Computer* **1999**, *32*, 68–75. [[CrossRef](#)]
41. Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. Density-Based Spatial Clustering of Applications with Noise. In Proceedings of the Second International Conference Knowledge Discovery and Data Mining, Portland, OR, USA, 2–4 August 1996; Volume 240.
42. Hinneburg, A.; Keim, D.A. *An Efficient Approach to Clustering in Large Multimedia Databases with Noise*; Bibliothek der Universität Konstanz: Konstanz, Germany, 1998; Volume 98.
43. Ankerst, M.; Breunig, M.M.; Kriegel, H.-P.; Sander, J. OPTICS: Ordering Points to Identify the Clustering Structure. *ACM Sigmod Rec.* **1999**, *28*, 49–60. [[CrossRef](#)]
44. Wang, W.; Yang, J.; Muntz, R. STING: A Statistical Information Grid Approach to Spatial Data Mining. *Vldb* **1997**, *97*, 186–195.
45. Agrawal, R.; Gehrke, J.; Gunopulos, D.; Raghavan, P. Automatic Subspace Clustering of High Dimensional Data for Data Mining Applications. In *SIGMOD '98 Proceedings of the 1998 ACM SIGMOD international conference on Management of data, Seattle, WA, USA, 1–4 June 1998*; ACM: New York, NY, USA, 1998; pp. 94–105.
46. Sheikholeslami, G.; Chatterjee, S.; Zhang, A. WaveCluster: A Wavelet-Based Clustering Approach for Spatial Data in Very Large Databases. *VLDB J.* **2000**, *8*, 289–304. [[CrossRef](#)]
47. Arifovic, J. Genetic Algorithm Learning and the Cobweb Model. *J. Econ. Dyn. Control.* **1994**, *18*, 3–28. [[CrossRef](#)]
48. Reynolds, D.A. Gaussian Mixture Models. *Encycl. Biom.* **2009**, *741*, 659–663.
49. Kohonen, T. *Self-Organizing Maps*; Springer Science & Business Media: Berlin, Germany, 2012; Volume 30, ISBN 3-642-56927-7.
50. Kutner, M.H.; Nachtsheim, C.J.; Neter, J.; Wasserman, W. *Applied Linear Regression Models*; McGraw-Hill/Irwin: New York, NY, USA, 2004; Volume 4.

51. Hosmer, D.W., Jr.; Lemeshow, S.; Sturdivant, R.X. *Applied Logistic Regression*; John Wiley & Sons: Hoboken, NJ, USA, 2013; Volume 398, ISBN 0-470-58247-2.
52. Jain, G.; Mallick, B. A Study of Time Series Models ARIMA and ETS. *SSRN J.* **2017**. [[CrossRef](#)]
53. Gardner, E.S. Exponential Smoothing: The State of the Art. *J. Forecast.* **1985**, *4*, 1–28. [[CrossRef](#)]
54. Hyndman, R.J.; Koehler, A.B.; Snyder, R.D.; Grose, S. A State Space Framework for Automatic Forecasting Using Exponential Smoothing Methods. *Int. J. Forecast.* **2002**, *18*, 439–454. [[CrossRef](#)]
55. Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C. *Time Series Analysis: Forecasting and Control*, 3rd ed.; Prentice Hall: Englewood Cliffs, NJ, USA, 1994; ISBN 978-0-13-060774-4.
56. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
57. Cover, T.; Hart, P. Nearest Neighbor Pattern Classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [[CrossRef](#)]
58. Uddin, S.; Haque, I.; Lu, H.; Moni, M.A.; Gide, E. Comparative Performance Analysis of K-Nearest Neighbour (KNN) Algorithm and Its Different Variants for Disease Prediction. *Sci. Rep.* **2022**, *12*, 6256. [[CrossRef](#)] [[PubMed](#)]
59. Quinlan, J.R. Discovering Rules by Induction from Large Collections of Examples. *Expert Syst. Micro Electron. Age* **1979**.
60. Quinlan, J.R. *C4. 5: Programs for Machine Learning*; Elsevier: Amsterdam, The Netherlands, 2014; ISBN 0-08-050058-7.
61. Lewis, R.J. An Introduction to Classification and Regression Tree (CART) Analysis. In Proceedings of the Annual meeting of the society for academic emergency medicine, San Francisco, CA, USA, 22–25 May 2000; Volume 14.
62. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
63. Freund, Y.; Schapire, R.; Abe, N. A Short Introduction to Boosting. *J.-Jpn. Soc. Artif. Intell.* **1999**, *14*, 1612.
64. Freund, Y.; Schapire, R.E. A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [[CrossRef](#)]
65. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
66. Chen, T.; Guestrin, C. Xgboost: A Scalable Tree Boosting System. In *KDD '16 Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, San Francisco, CA, USA, 13–17 August 2016*; ACM: New York, NY, USA, 2016; pp. 785–794.
67. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4 December 2017; Volume 30, pp. 3149–3157.
68. Dorogush, A.V.; Ershov, V.; Gulin, A. CatBoost: Gradient Boosting with Categorical Features Support. *arXiv* **2018**, arXiv:1810.11363.
69. Wolpert, D.H. Stacked Generalization. *Neural Netw.* **1992**, *5*, 241–259. [[CrossRef](#)]
70. Jain, A.K.; Mao, J.; Mohiuddin, K.M. Artificial Neural Networks: A Tutorial. *Computer* **1996**, *29*, 31–44. [[CrossRef](#)]
71. Sanger, T.D. Optimal Unsupervised Learning in a Single-Layer Linear Feedforward Neural Network. *Neural Netw.* **1989**, *2*, 459–473. [[CrossRef](#)]
72. Murat, H.S. A brief review of feed-forward neural networks. *Commun. Fac. Sci. Univ. Ank.* **2006**, *50*, 11–17. [[CrossRef](#)]
73. Sherstinsky, A. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network. *Phys. D Nonlinear Phenom.* **2020**, *404*, 132306. [[CrossRef](#)]
74. Gopalsamy, K.; He, X. Stability in Asymmetric Hopfield Nets with Transmission Delays. *Phys. D Nonlinear Phenom.* **1994**, *76*, 344–358. [[CrossRef](#)]
75. Ackley, D.H.; Hinton, G.E.; Sejnowski, T.J. A Learning Algorithm for Boltzmann Machines. *Cogn. Sci.* **1985**, *9*, 147–169. [[CrossRef](#)]
76. Gers, F.A.; Schmidhuber, J.; Cummins, F. Learning to Forget: Continual Prediction with LSTM. *Neural Comput.* **2000**, *12*, 2451–2471. [[CrossRef](#)] [[PubMed](#)]
77. Salinas, D.; Flunkert, V.; Gasthaus, J. DeepAR: Probabilistic Forecasting with Autoregressive Recurrent Networks. *Int. J. Forecast.* **2020**, *36*, 1181–1191. [[CrossRef](#)]
78. Rangapuram, S.S.; Seeger, M.W.; Gasthaus, J.; Stella, L.; Wang, Y.; Januschowski, T. Deep State Space Models for Time Series Forecasting. *Adv. Neural Inf. Process. Syst.* **2018**, *31*, 7796–7805.
79. Wen, R.; Torkkola, K.; Narayanaswamy, B.; Madeka, D. A Multi-Horizon Quantile Recurrent Forecaster. *arXiv* **2017**, arXiv:1711.11053.
80. Lu, W.; Li, J.; Li, Y.; Sun, A.; Wang, J. A CNN-LSTM-Based Model to Forecast Stock Prices. *Complexity* **2020**, *2020*, 1–10. [[CrossRef](#)]
81. Huang, C.-J.; Kuo, P.-H. A Deep CNN-LSTM Model for Particulate Matter (PM_{2.5}) Forecasting in Smart Cities. *Sensors* **2018**, *18*, 2220. [[CrossRef](#)] [[PubMed](#)]
82. Kim, T.-Y.; Cho, S.-B. Predicting Residential Energy Consumption Using CNN-LSTM Neural Networks. *Energy* **2019**, *182*, 72–81. [[CrossRef](#)]
83. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*. [[CrossRef](#)]
84. Li, S.; Jin, X.; Xuan, Y.; Zhou, X.; Chen, W.; Wang, Y.-X.; Yan, X. Enhancing the Locality and Breaking the Memory Bottleneck of Transformer on Time Series Forecasting. *Adv. Neural Inf. Process. Syst.* **2019**, *32*. [[CrossRef](#)]
85. Zerveas, G.; Jayaraman, S.; Patel, D.; Bhamidipaty, A.; Eickhoff, C. A Transformer-Based Framework for Multivariate Time Series Representation Learning. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, Singapore, 14–18 August 2021; pp. 2114–2124.

86. Zhou, H.; Zhang, S.; Peng, J.; Zhang, S.; Li, J.; Xiong, H.; Zhang, W. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. *AAAI Conf. Artif. Intell.* **2021**, *35*, 11106–11115. [[CrossRef](#)]
87. Lim, B.; Arık, S.Ö.; Loeff, N.; Pfister, T. Temporal Fusion Transformers for Interpretable Multi-Horizon Time Series Forecasting. *Int. J. Forecast.* **2021**, *37*, 1748–1764. [[CrossRef](#)]
88. Oreshkin, B.N.; Carpo, D.; Chapados, N.; Bengio, Y. N-BEATS: Neural Basis Expansion Analysis for Interpretable Time Series Forecasting. *arXiv* **2020**, arXiv:1905.10437.
89. Williams, H.P. *Model Building in Mathematical Programming*; John Wiley & Sons: Hoboken, NJ, USA, 2013; ISBN 1-118-50618-9.
90. Klee, V.; Minty, G.J. How Good Is the Simplex Algorithm. *Inequalities* **1972**, *3*, 159–175.
91. Ruder, S. An Overview of Gradient Descent Optimization Algorithms. *arXiv* **2016**, arXiv:1609.04747.
92. Dennis, J.; Moré, J.J. Quasi-Newton Methods, Motivation and Theory. *SIAM Rev.* **1977**, *19*, 46–89. [[CrossRef](#)]
93. Shewchuk, J.R. *An Introduction to the Conjugate Gradient Method without the Agonizing Pain*; Carnegie-Mellon University, Department of Computer Science Pittsburgh: Pittsburgh, PA, USA, 1994.
94. DeVore, R.A.; Temlyakov, V.N. Some Remarks on Greedy Algorithms. *Adv. Comput. Math.* **1996**, *5*, 173–187. [[CrossRef](#)]
95. Johnson, D.S.; Papadimitriou, C.H.; Yannakakis, M. How Easy Is Local Search? *J. Comput. Syst. Sci.* **1988**, *37*, 79–100. [[CrossRef](#)]
96. Tsamardinos, I.; Brown, L.E.; Aliferis, C.F. The Max-Min Hill-Climbing Bayesian Network Structure Learning Algorithm. *Mach. Learn.* **2006**, *65*, 31–78. [[CrossRef](#)]
97. Kennedy, J.; Eberhart, R. Particle Swarm Optimization. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.
98. Dorigo, M.; Birattari, M.; Stutzle, T. Ant Colony Optimization. *IEEE Comput. Intell. Mag.* **2006**, *1*, 28–39. [[CrossRef](#)]
99. Kirkpatrick, S.; Gelatt, C.D., Jr.; Vecchi, M.P. Optimization by Simulated Annealing. *Science* **1983**, *220*, 671–680. [[CrossRef](#)] [[PubMed](#)]
100. Glover, F.; Laguna, M. Tabu Search. In *Handbook of Combinatorial Optimization*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 2093–2229.
101. Nair, V.; Bartunov, S.; Gimeno, F.; von Glehn, I.; Lichocki, P.; Lobov, I.; O'Donoghue, B.; Sonnerat, N.; Tjandraatmadja, C.; Wang, P.; et al. Solving Mixed Integer Programs Using Neural Networks. *arXiv* **2021**, arXiv:2012.13349.
102. Vinyals, O.; Fortunato, M.; Jaitly, N. Pointer Networks. In Proceedings of the 28th International Conference on Neural Information Processing Systems, Cambridge, MA, USA, 7 December 2015; Volume 2, pp. 2692–2700.
103. Schuetz, M.J.A.; Brubaker, J.K.; Katzgraber, H.G. Combinatorial Optimization with Physics-Inspired Graph Neural Networks. *Nat. Mach. Intell.* **2022**, *4*, 367–377. [[CrossRef](#)]
104. Lillicrap, T.P.; Hunt, J.J.; Pritzel, A.; Heess, N.; Erez, T.; Tassa, Y.; Silver, D.; Wierstra, D. Continuous Control with Deep Reinforcement Learning. *arXiv* **2015**, arXiv:1509.02971.
105. Boute, R.N.; Gijsbrechts, J.; van Jaarsveld, W.; Vanvuchelen, N. Deep Reinforcement Learning for Inventory Control: A Roadmap. *Eur. J. Oper. Res.* **2021**, *298*, 401–412. [[CrossRef](#)]
106. de Oliveira, M.P.V.; McCormack, K.; Trkman, P. Business Analytics in Supply Chains—The Contingent Effect of Business Process Maturity. *Expert Syst. Appl.* **2012**, *39*, 5488–5498. [[CrossRef](#)]
107. Wu, P.-J.; Huang, P.-C. Business Analytics for Systematically Investigating Sustainable Food Supply Chains. *J. Clean. Prod.* **2018**, *203*, 968–976. [[CrossRef](#)]
108. Trkman, P.; McCormack, K.; de Oliveira, M.P.V.; Ladeira, M.B. The Impact of Business Analytics on Supply Chain Performance. *Decis. Support Syst.* **2010**, *49*, 318–327. [[CrossRef](#)]
109. Nam, D.; Lee, J.; Lee, H. Business Analytics Use in CRM: A Nomological Net from IT Competence to CRM Performance. *Int. J. Inf. Manag.* **2019**, *45*, 233–245. [[CrossRef](#)]
110. Acito, F.; Khatri, V. Business Analytics: Why Now and What Next? *Bus. Horiz.* **2014**, *57*, 565–570. [[CrossRef](#)]
111. Zhang, Z.; Xiao, Y.; Fu, Z.; Zhong, K.; Niu, H. A Study on Early Warnings of Financial Crisis of Chinese Listed Companies Based on DEA-SVM Model. *Mathematics* **2022**, *10*, 2142. [[CrossRef](#)]
112. Zhou, W.; Chen, M.; Yang, Z.; Song, X. Real Estate Risk Measurement and Early Warning Based on PSO-SVM. *Socio-Econ. Plan. Sci.* **2021**, *77*, 101001. [[CrossRef](#)]
113. Jianying, F.; Bianyu, Y.; Xin, L.; Dong, T.; Weisong, M. Evaluation on Risks of Sustainable Supply Chain Based on Optimized BP Neural Networks in Fresh Grape Industry. *Comput. Electron. Agric.* **2021**, *183*, 105988. [[CrossRef](#)]
114. Jiang, H.; Ching, W.-K.; Yiu, K.F.C.; Qiu, Y. Stationary Mahalanobis Kernel SVM for Credit Risk Evaluation. *Appl. Soft Comput.* **2018**, *71*, 407–417. [[CrossRef](#)]
115. Gerrard, M.; Gibbons, F.X.; Houlihan, A.E.; Stock, M.L.; Pomery, E.A. A Dual-Process Approach to Health Risk Decision Making: The Prototype Willingness Model. *Dev. Rev.* **2008**, *28*, 29–61. [[CrossRef](#)]
116. Pröllochs, N.; Feuerriegel, S. Business Analytics for Strategic Management: Identifying and Assessing Corporate Challenges via Topic Modeling. *Inf. Manag.* **2020**, *57*, 103070. [[CrossRef](#)]
117. Appelbaum, D.; Kogan, A.; Vasarhelyi, M.; Yan, Z. Impact of Business Analytics and Enterprise Systems on Managerial Accounting. *Int. J. Account. Inf. Syst.* **2017**, *25*, 29–44. [[CrossRef](#)]
118. Nielsen, S. The Impact of Business Analytics on Management Accounting. *SSRN J.* **2015**. [[CrossRef](#)]
119. Rikhardsson, P.; Yigitbasioglu, O. Business Intelligence & Analytics in Management Accounting Research: Status and Future Focus. *Int. J. Account. Inf. Syst.* **2018**, *29*, 37–58. [[CrossRef](#)]

120. van der Togt, J.; Rasmussen, T.H. Toward Evidence-Based HR. *JOEPP* **2017**, *4*, 127–132. [[CrossRef](#)]
121. Margherita, A. Human Resources Analytics: A Systematization of Research Topics and Directions for Future Research. *Hum. Resour. Manag. Rev.* **2022**, *32*, 100795. [[CrossRef](#)]
122. Pape, T. Prioritising Data Items for Business Analytics: Framework and Application to Human Resources. *Eur. J. Oper. Res.* **2016**, *252*, 687–698. [[CrossRef](#)]
123. Stadler, J.G.; Donlon, K.; Siewert, J.D.; Franken, T.; Lewis, N.E. Improving the Efficiency and Ease of Healthcare Analysis Through Use of Data Visualization Dashboards. *Big Data* **2016**, *4*, 129–135. [[CrossRef](#)] [[PubMed](#)]
124. Stelzer, G.; Rosen, N.; Plaschkes, I.; Zimmerman, S.; Twik, M.; Fishilevich, S.; Stein, T.I.; Nudel, R.; Lieder, I.; Mazor, Y.; et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinform.* **2016**, *54*, 1–30. [[CrossRef](#)] [[PubMed](#)]
125. Fanelli, D.; Piazza, F. Analysis and Forecast of COVID-19 Spreading in China, Italy and France. *Chaos Solitons Fractals* **2020**, *134*, 109761. [[CrossRef](#)] [[PubMed](#)]
126. Ward, M.J.; Marsolo, K.A.; Froehle, C.M. Applications of Business Analytics in Healthcare. *Bus. Horiz.* **2014**, *57*, 571–582. [[CrossRef](#)] [[PubMed](#)]
127. Commission, E. *A New Circular Economy Action Plan*; Office of the European Union Brussels: Brussels, Belgium, 2020; pp. 1–19.
128. Kristoffersen, E.; Mikalef, P.; Blomsma, F.; Li, J. The Effects of Business Analytics Capability on Circular Economy Implementation, Resource Orchestration Capability, and Firm Performance. *Int. J. Prod. Econ.* **2021**, *239*, 108205. [[CrossRef](#)]
129. Kristoffersen, E.; Mikalef, P.; Blomsma, F.; Li, J. Towards a Business Analytics Capability for the Circular Economy. *Technol. Forecast. Soc. Change* **2021**, *171*, 120957. [[CrossRef](#)]
130. Zhao, R.; Liu, Y.; Zhang, N.; Huang, T. An Optimization Model for Green Supply Chain Management by Using a Big Data Analytic Approach. *J. Clean. Prod.* **2017**, *142*, 1085–1097. [[CrossRef](#)]
131. Kaur, M.; Kang, S. Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining. *Procedia Comput. Sci.* **2016**, *85*, 78–85. [[CrossRef](#)]
132. Videla-Cavieres, I.F.; Ríos, S.A. Extending Market Basket Analysis with Graph Mining Techniques: A Real Case. *Expert Syst. Appl.* **2014**, *41*, 1928–1936. [[CrossRef](#)]
133. Griva, A.; Bardaki, C.; Pramataris, K.; Papakiriakopoulos, D. Retail Business Analytics: Customer Visit Segmentation Using Market Basket Data. *Expert Syst. Appl.* **2018**, *100*, 1–16. [[CrossRef](#)]
134. Hwangbo, H.; Kim, Y.S.; Cha, K.J. Recommendation System Development for Fashion Retail E-Commerce. *Electron. Commer. Res. Appl.* **2018**, *28*, 94–101. [[CrossRef](#)]
135. Isinkaye, F.O.; Folajimi, Y.O.; Ojokoh, B.A. Recommendation Systems: Principles, Methods and Evaluation. *Egypt. Inform. J.* **2015**, *16*, 261–273. [[CrossRef](#)]
136. Kim, K.; Lee, K.; Ahn, H. Predicting Corporate Financial Sustainability Using Novel Business Analytics. *Sustainability* **2018**, *11*, 64. [[CrossRef](#)]
137. Troilo, M.; Bouchet, A.; Urban, T.L.; Sutton, W.A. Perception, Reality, and the Adoption of Business Analytics: Evidence from North American Professional Sport Organizations. *Omega* **2016**, *59*, 72–83. [[CrossRef](#)]
138. Wang, L. Heterogeneous Data and Big Data Analytics. *ACIS* **2017**, *3*, 8–15. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.