UNIVERSITY OF WEST LONDON
The Career University

UWL REPOSITORY

repository.uwl.ac.uk

Big data analytics correlation taxonomy

Husamaldin, Laden and Saeed, Nagham ORCID logoORCID: https://orcid.org/0000-0002-5124-7973 (2019) Big data analytics correlation taxonomy. Information, 11 (1). p. 17.

http://dx.doi.org/10.3390/info11010017

This is the Published Version of the final output.

UWL repository link: https://repository.uwl.ac.uk/id/eprint/6658/

Alternative formats: If you require this document in an alternative format, please contact: open.research@uwl.ac.uk

Copyright: Creative Commons: Attribution 4.0

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy: If you believe that this document breaches copyright, please contact us at open.research@uwl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Rights Retention Statement:

information

MDPI

Concept Paper

Big Data Analytics Correlation Taxonomy

Laden Husamaldin * and Nagham Saeed *

School of Computing and Engineering, University of West London, London W5 5RF, UK

* Correspondence: Laden.husamaldin2@uwl.ac.uk (L.H.); Nagham.saeed@uwl.ac.uk (N.S.)

Received: 14 November 2019; Accepted: 22 December 2019; Published: 25 December 2019

Abstract: Big data analytics (BDA) is an increasingly popular research area for both organisations and academia due to its usefulness in facilitating human understanding and communication. In the literature, researchers have focused on classifying big data according to data type, data security or level of difficulty, and many research papers reveal that there is a lack of information on evidence of a real-world link of big data analytics methods and its associated techniques. Thus, many organisations are still struggling to realise the actual value of big data analytic methods and its associated techniques. Therefore, this paper gives a design research account for formulating and proposing a step ahead to understand the relation between the analytical methods and its associated techniques. Furthermore, this paper is an attempt to clarify this uncertainty and identify the difference between analytics methods and techniques by giving clear definitions for each method and its associated techniques to integrate them later in a new correlation taxonomy based on the research approaches. Thus, the primary outcome of this research is to achieve for the first time a correlation taxonomy combining analytic methods used for big data and its recommended techniques that are compatible for various sectors. This investigation was done through studying various descriptive articles of big data analytics methods and its associated techniques in different industries.

Keywords: big data analytics methods; big data characteristics; big data techniques; big data correlation taxonomy

1. Introduction

Big data analytics (BDA) represents an important part of an information system for development and evolution within organisations. Looking into the literature, a variety of definitions exist for big data analytics, including: "a help to predict future volumes, gain insights, take proactive actions, and give way to better strategic decision-making" [1]. Other definitions provide similar interpretations of the term (for example, see [2–6]. Whereas [7] described BDA as "advanced analytic techniques operating on big data sets to help discover what has changed and how we should react". Furthermore, [8] stated that BDA differ from traditional analytical approaches as instead of tracking care quality and outcomes in a retrospective view by using deductive reasoning, it uses inductive reasoning for prospective analysis of data. The definition in [7] considered data analytics as data analytic techniques, which indicates a vague understanding of the difference between data analytics methods and data analytics techniques.

There are many challenges in big data, the different types of big data challenges were discussed by [9,10]. The broad challenges of big data (BD) were grouped into three main categories, based on the data life cycle: data, process and management challenges. However, suggesting how BDA methods and techniques could address these challenges is out of the paper's scope.

The purpose of this research is to design a correlation taxonomy of the existing methods and its associated techniques. This taxonomy is derived from the limitation in the existing literature, which will be further discussed below after finding the correlation between the methods and the associated

techniques of BDA. The research is constructed according to a design science research (DSR) approach used to design the taxonomy. Furthermore, this paper presents useful budding research in the field of BDA to give a clear understanding of the basic data analytics concepts, to early stage and interested researchers.

The remainder of this paper is structured as follows: Section 2 includes the necessary background with an overview of related literature for big data characteristics definitions and BDA usability, which is followed by Section 3 which includes definitions, different classification of BDA methods based on literature review and the authors' classification preference. Section 4 highlights the need for this research and constitutes of the research methodology, literature review, analysis for the correlation between BDA methods, correlation between BDA methods, the associated techniques, and finally analysis for techniques used in research papers. Section 5 is discussion of the authors' proposed taxonomy and Section 6 highlights the conclusions and outlines the research limitations and potential for further research.

2. Background

The magnitude of data generated and shared by sources and organisations, such as: businesses, public administrations, numerous industries, non-profit sectors, and scientific research has increased immeasurably [11]. To give an example of data magnitude, more than 500 million tweets are posted every single day, 428,571 million share contents of LinkedIn, 7 billion Google searches are launched. Additionally, a comparison done using [12], shown in Figure 1, compares the search interest in big data term over 10 years (the x-axis represents the time and y-axis represents the interest in the topic search). This comparison shows the search of "big data" terms within 2008 (for the whole year) vs. 2018 (for the whole year). The red line shows people's interest in big data in 2008 whereas the blue line shows people's interest in big data 2018. This comparison is made based on interest over time thus the numbers represent search interest relative to the highest point on the graph for the given region and time. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. A score of 0 means that there was not enough data for this term. The graph shows the increase in the interest in big data over the 10 past years. The chart also confirms the graph's results with a big increase over the years.

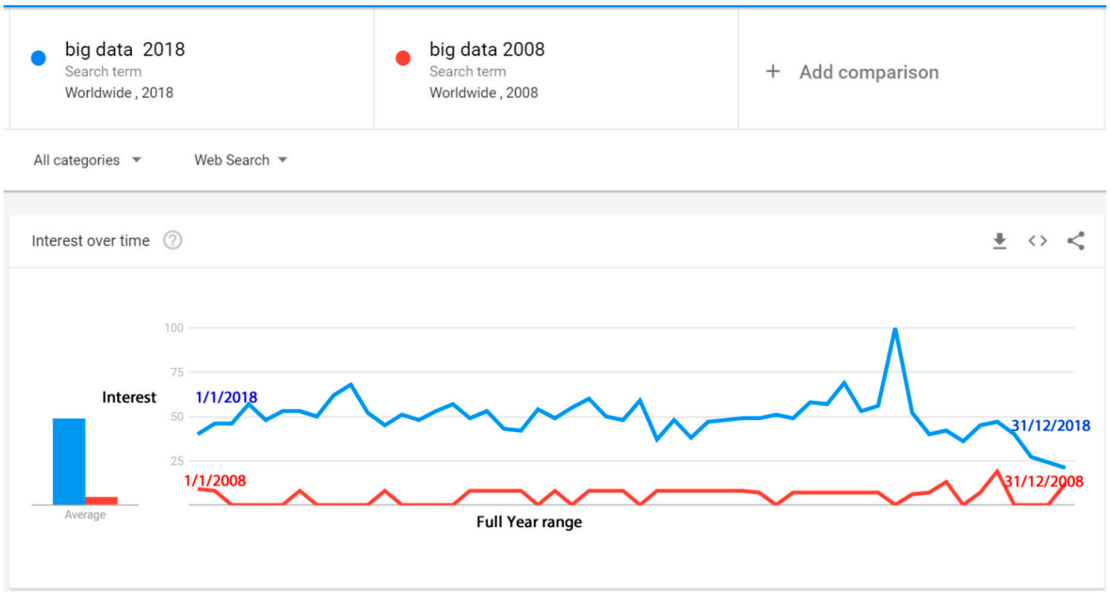

Figure 1. Interest in big data from 2008 to 2018.

Most definitions of big data analytics focus on the size of the data in storage. Size matters, but there are other important attributes of big data analytics as well which has been characterised by Erl et al. (2017) into 5V's: volume, variety, velocity, veracity, and value as presented in Figure 2. The

paper provides a comprehensive definition, and breaks the myth that big data analytics is only about data volume as each of the five V's has its own ramifications for analytics. According to [13], volume refers to "the magnitude of data, which has exponentially increased, posing a challenge to the capacity of existing storage devices" and variety refers to "the fact that data can be generated from heterogeneous sources", for example: sensors, Internet of Things (IoT), mobile devices, online social networks, etc., in structured, semi-structured, and unstructured formats. To give a holistic picture of data classification, structured data is typically stored in databases or spreadsheets. Text, audio, imagery and video refers to unstructured data, which sometimes lack the structural organisation required for analysis. Spanning a continuum between fully structured and unstructured data, the format of semi-structured data does not conform to strict standards. Velocity refers to "the speed of data generation and delivery, which can be processed in batch, real-time, nearly real-time, or stream-lines" [14]. Veracity "stresses the importance of data quality and level of trust due to the concern that many data sources (e.g., social networking sites) inherently contain a certain degree of uncertainty and unreliability" [15]. Finally, value refers to "the process of revealing underexploited values from big data to support decision-making" [15].

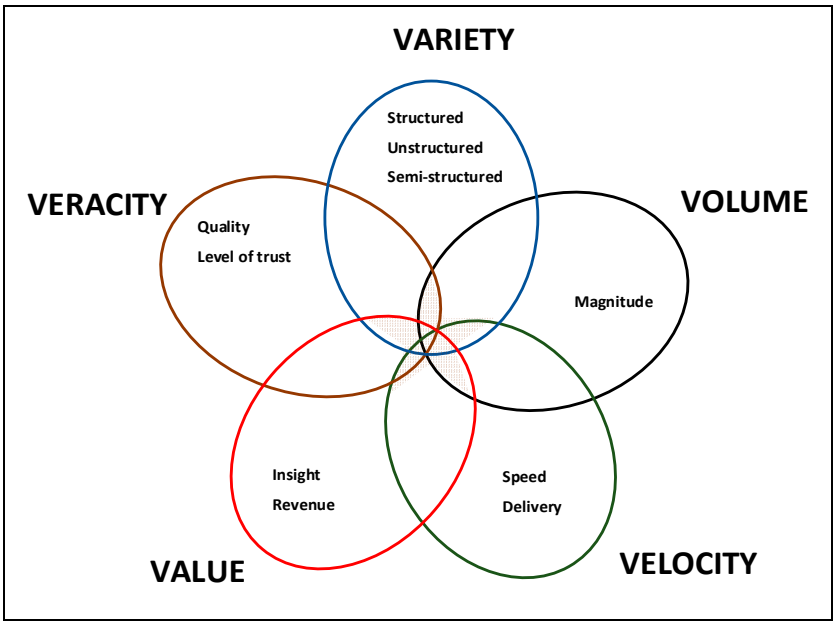

Figure 2. Representation of the five V's of big data.

According to studies by [15–19], more V's and other characteristics have been added to support a better defined big data: such as vision (a purpose), verification (processed data conformed to some specifications), validation (the purpose is fulfilled), variability (data differentiation), venue (different platforms), vocabulary (data terminology) and vagueness (indistinctness of existence in a data), complexity (it is difficult to organise and analyse big data because of evolving data relationships) and immutability (collected and stored big data can be permanent if well managed). It is worth mentioning that the correlation taxonomy proposed in this paper is not affected by a number of the big data analytics characteristics and should be applicable for any number of V's.

The big data analytics concept was used in many sectors, the latest understanding in academia [20–23] specified the potential of big data analytics in five main sectors:

- Healthcare: clinical decision support systems, individual analytics applied for patient profiling, personalised medicine, performance-based pricing for personnel, analysis of disease patterns and improvement of public health.
- Public sector: creating transparency with accessible related data, discovering needs, improving performance, customisation of actions for suitable products and services, decision-making with automated systems to decrease risks, innovating new products and services.

- Retail: in-store behaviour analysis, variety and price optimisation, product placement design, improve performance, labour inputs, optimisation, distribution and logistics optimisation, Web-based markets.
- Manufacturing: improved demand forecasting, supply chain planning, sales support, developing production operations, web-search-based applications.
- Personal location data: smart routing, geo-targeted advertising or emergency response, urban planning, new business models.

The importance of big data analytics has laid the groundwork for investigation of the methods and the techniques involved in big data, which will be explored further in the following section.

3. A Narrative Prospective for Big Data Analytics Methods

As organisations start to adopt analytics as the new science of winning, they need to focus on understanding the method of analytics that support its insights and enables better decisions [24]. One way of appreciating these methods is through classification. Therefore, BDA was classified according to different criteria such as data understanding. In 2009, Thomas Davenport based his idea on looking at all the data first to understand it and then answer the questions: what has happened? Why it has happened? What will happen? And how to make the best of it? These four questions were identified as descriptive, diagnostic, predictive, and prescriptive methods, which will be explained clearly with examples below:

- Descriptive analytics looks at data and analyses past events for insight as to how to approach the future; it asks, "what has happened?" An example is to categorise customers by their product preferences and life stage.
- Diagnostic analytics at this stage, historical data can be measured against other data to answer the question of "why it happened?" providing a possible way to find out dependencies and to identify patterns. For example, a retailer filters the sales down to subcategories. Companies employ diagnostic analytics, as it gives them in-depth insights into a particular problem. At the same time, a company should have detailed information at their disposal; otherwise, data collection may turn out to be individual for every issue and therefore time-consuming.
- Predictive analytics turns data into valuable and actionable information; in other words predictive analytics determines the probable future outcome of an event or a likelihood of a situation occurring; i.e., "what will happen?" For example, for an organisation that offers multiple products, predictive analytics can help analyse customers' spending, usage and other behavior, leading to efficient cross sales, or selling additional products to current customers. This directly leads to higher profitability per customer and stronger customer relationships.
- Prescriptive analytics automatically synthesises big data, mathematical sciences, business rules, and machine learning to make predictions and then suggests decision options to take advantage of the predictions, asking "how to make the best of it?" An example is to determination of the best pricing and advertising strategy to maximise revenue.

In 2019, Davenport added another question "What if we take action? Which was represented by an automated analytics method that follows the prescriptive method [25]. A similar classification was presented in [26] where three analytical stages were presented: descriptive and exploratory, predictive and prescriptive where data preparation methods were considered in an advanced stage. Other BDA classifications are valid too, such as [27] classification for BDA methods according to security data that includes statistical methods, machine learning and knowledge-based methods whereas [28] presented generic architecture for big data analytical approaches, allowing classification according to the data storage layer, the distributed parallel model, and the type of database used.

The Davenport's BDA classification [29] is the preferred and most-popular structure because of the integration, ordering and compatibility between the methods; hence, it will be adopted in this paper. Figure 3 illustrates how big data analytics methods complement each other. The analytics for big data should start with the descriptive method that helps to gain insight from historical data then

grow to the diagnostic analytics for clustering, predictive analytics helps to designate the future outcomes and prescriptive methods involved in making decisions. The following sections will discuss and present the BDA correlations based on the method's integration factor, presented in this section.

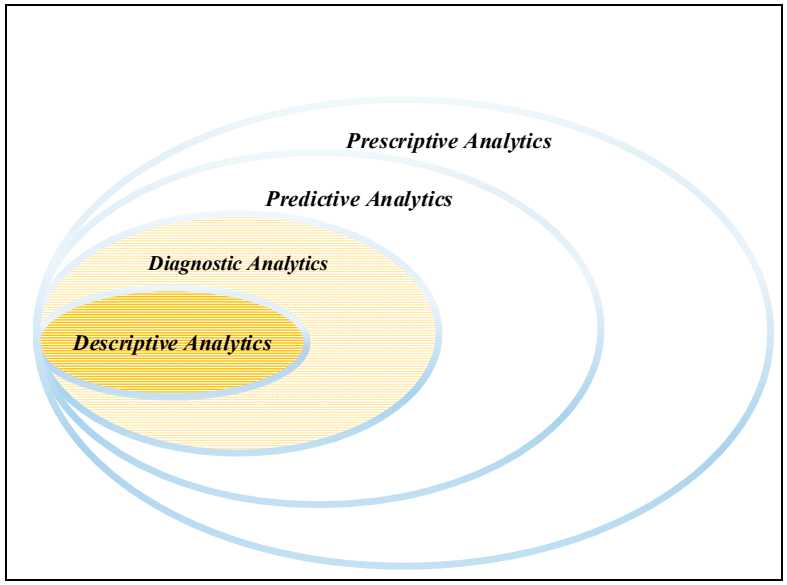

Figure 3. Big data analytics methods integration.

4. Research Methodology

To develop the proposed correlation taxonomy, the authors have applied the problem-centred approach of the design science research methodology (DSRM) presented by [30]. Five DSRM activities were drawn to come to a rigorous and relevant research results [31] as listed and defined in Table 1.

Table 1. The design science research methodology (DSRM) activities application.

DSRM Activity	Research Application
Identify problem and motivate	The existing problem was identified from existing literature and previous research projects. The research problem identified was the lack of realising the actual correction between big data analytic methods and its associated techniques based on a known approach.
Define objectives of a solution	To develop a structural model combining analytic methods used for big data and its recommended techniques based on research approaches that are compatible for various sectors.
Design and development	Several phases were required to fully develop a taxonomy aimed at formatting and structuring big data analytics techniques according to research approaches. The full understanding of each phase evolves and improves the taxonomy concept. The taxonomy was reviewed against big data analytics techniques (and methods) creating finer taxonomy.
Demonstration and evaluation	The demonstration and evaluation formed the basis for the design phase of the following iteration. For the future a final evaluation scenario will be used in order to raise the discussion on the importance and limitation of the presented correlation taxonomy in this research.
Communication	Research published in academic papers.

Thus, to provide a clear structure of the big data analytics correlation taxonomy based on research approaches, a clear understanding for the big data analytical methods correlation is fundamental. Besides, a correlation between the techniques and the methods needed to be investigated. In conclusion, based on existing literature in the big data analytics field, two correlations were explored in this paper: first, correlation between big data analytical methods; and second, correlation between big data analytical methods and techniques. The outcomes are presented below:

4.1. Stages of Big Data Analytical Methods

The relationship between the BDA methods presented in [32] was based on the difficulty criteria, the methods were ranked according to their difficulty and value. For example, the descriptive method is considered the easiest stage with the least value compared with the other three methods, whereas the prescriptive method has the highest value and difficulty. The research in this paper will be based on the concepts in [29,32] to develop the method analytical categories. Figure 4 illustrates the analytical categories based on the relation among the four methods. It shows the outcome from one method which will be the input for the next one. It also represents a development in the analytical level by finishing each stage and providing deeper analysis for the problem, as described below with supported examples adopted from [33]:

Stage One: Descriptive analytics stage where the information is categorised, for example gaining insight from historical data. For example, a healthcare provider will learn how many patients were hospitalised in the previous month.

Stage Two: Diagnostic analytics stage includes the data aggregation, reporting, drilling and clustering of the data into subcategories. For instance, a healthcare provider compares patients' response to a promotional campaign in different regions.

Stage Three: Predictive analytics stage is an important stage because it helps to designate the future outcomes, by using statistical and machine-learning techniques. This stage is where the data will be analysed, classified, simulated and summarised to give a prediction; and it gives a good understanding for the forthcoming events. For example, a health care provider predicts the risk of investing in new medical devices based on registered patients.

Stage Four: The final prescriptive analytics stage is the optimisation stage, where the best framework will be given to secure the best futuristic outcome. It also involves taking decisions using

deeper learning techniques. For example, in the health care industry, you can better manage the patient population by using prescriptive analytics to measure the number of patients who are clinically obese, then add filters for risk factors like diabetes and low-density lipoproteins (LDL) cholesterol levels to determine where to focus treatment.

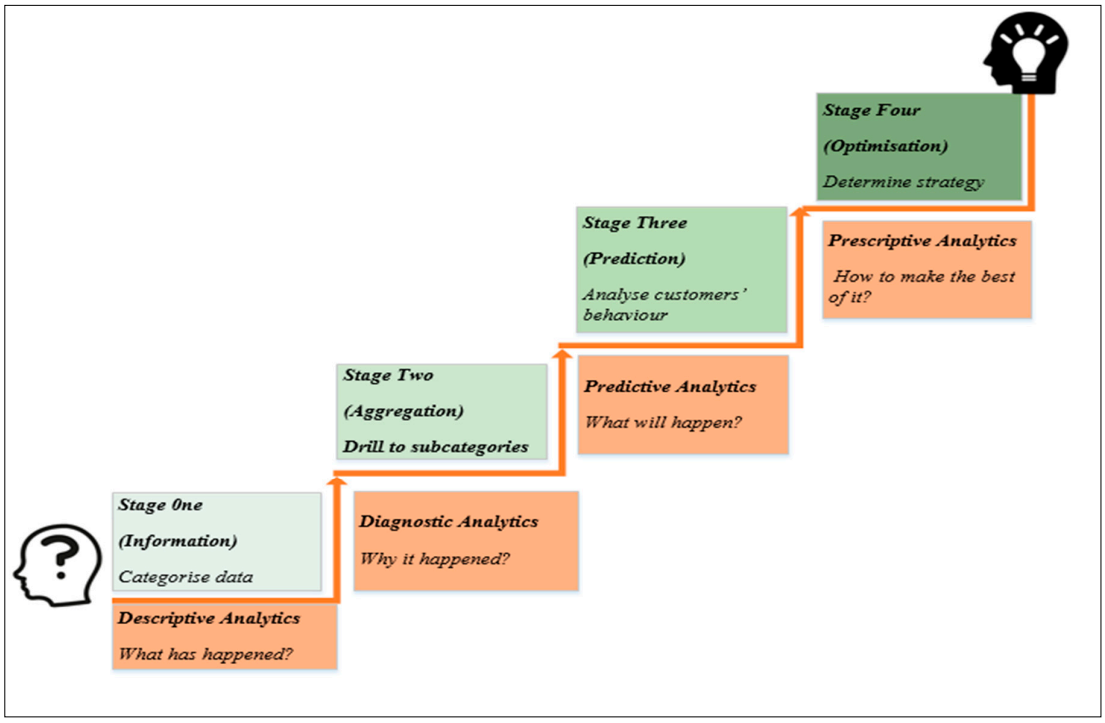

Figure 4. Analytical methods category.

Figure 4 shows a well-built structured correlation chain between the analytical methods based on the four stages. It is clearly shown that these methods cannot function in parallel besides the methods that do not intersect between each other.

4.2. Correlation between Big Data Analytical Methods and Techniques

Not many academic papers are available to provide taxonomy for big data analytics techniques. Only one research paper by [34] was found which offers taxonomy according to big data analytics techniques. The paper developed a categorised indexing techniques taxonomy to provide insight to enable researchers to understand and select a technique as a basis to design an indexing mechanism with reduced time and space consumption for big data and Mobile Cloud Computing. The categories are non-artificial intelligence, artificial intelligence, and collaborative artificial intelligence indexing methods. Furthermore, another study has tried to provide the link between the analytical methods and techniques presented by [26], focused on big data analytics methods and its associated techniques, including those associated with data preparation and descriptive, predictive, and prescriptive analytics as shown in Table 2. Mathematical calculation and visualisation techniques were under the descriptive and exploratory method, whereas machine learning, linear and non-linear regression, classification, data mining, text analytics, Bayesian methods and simulation techniques were considered under the predictive method. Stochastic models of uncertainty, mathematical optimization under uncertainty and optimal solutions techniques were listed under prescriptive method.

Table 2. Big data analytical (BDA) methods and their associate techniques according to [26].

BDA Method	BDA Techniques
Descriptive and Exploratory	Mathematical Calculation Visualisation
Predictive	Machine Learning Linear and Non-Linear Regression Classification Data Mining Text Analytics Bayesian Methods Simulation
Prescriptive	Stochastic Models of Uncertainty Mathematical Optimization Under Uncertainty Optimal Solutions

Table 2 removes the vagueness in understanding the difference between data analytics methods and techniques as it groups the techniques which are suitable for a specific analytical stage. The table groups the techniques according to the three methods proposed in [26]. Therefore, this section will provide examples of known big data analytics technologies that are implemented in research papers based on the findings represented in Figure 4 Section 3 (the four analytical methods proposed: descriptive, diagnostic, predictive and prescriptive methods). Table 3 lists examples of known technologies for each analytic method mentioned before. Yet, given the breadth of the available techniques, an exhaustive list of techniques is beyond the scope of a single paper. As shown in Table 3, the work presented in [35] adopted a descriptive method in the research as it was based on questionnaires and interview techniques to collect the data required, whereas the visualisation technique used was ExcelPro. Similarly, the research work in [36] paper as the 4th generation of supervisory control and data acquisition (SCADA) system was adopted as the descriptive method and implemented Web-application (WebForms class library which was hosted in Microsoft Azure servers) as a visualisation technique. The research in [37] paper used a particle swarm algorithm as a clustering technique for its diagnostic methods and MATLAB software as the visualisation technique. Logistic regression models (statistical models based on traditional mathematical equation) and artificial intelligence (AI) [38] modelling are used as predictive methods alongside genetic algorithms (GA) [39] and neuro-fuzzy [40] techniques used for classification. Excel, and Simulink software were the visualisation techniques. The work in [41] provided a good example for prescriptive methods as a system was developed to provide the optimum routing protocol according to the context. OPNET was the visualisation technique used.

Table 3. Analytical methods sampling techniques.

Big Data Analytic Method	Sample of Techniques	Visualisation Software	Research
Descriptive	Data collection	ExcelPro SCADA	[35,36]
Diagnostic	Clustering techniques	MATLAB	[37]
Predictive	Modelling classification	Excel, and Simulink	[38–40]
Prescriptive	Optimal solutions	OPNET	[41]

It clearly shows from Table 3 that the authors recognised the effect of the visualisation technique and its existence in all big data analytics methods due to its importance in presenting the information to the individuals in a professional and structured manner. Visualisations could be accomplished by using common tools such as bar charts, box plots, and scatter plots. This will have a great impact on the data analysis and reduce the chance of missing important information. From

Table 3. It has been concluded that there is a need for another big data analytics taxonomy which will provide a bigger picture to link between big data analytics methods and techniques.

5. Big Data Analytics Correlation Taxonomy

Analytics methods and techniques have been applied to various phases of the analytical lifecycle, e.g., identification, preparation, aggregation and visualisation, however the application of various analytical techniques to big data has been limited with research having been conducted by only a few over the past 20 years. While limited, such research has been highly effective in proposing different sets and kinds of classification, however a problem that has not been effectively researched is how those techniques are classified and how such an activity can be made systematic via a clear taxonomy that integrates the classification of those analytical methods and its associated techniques. This paper investigates this problem and introduced a novel correlation taxonomy based on the findings presented in the previous sections. The rationale of introducing this taxonomy is offering a better big data analytics topology by examining the existing big data analytics methods, as well as understanding what types of techniques map with the existing methods. The taxonomy also draws a clear relationship between research approach methods and big data analytics methods techniques and represents the correlation between big data analytical methods and techniques based on research approaches. This taxonomy aims to classify big data analytics methods and techniques according to well-known research approach methods. This taxonomy determines the right techniques (and consequently the methods) that will equip new researchers in the field with the right tools. The taxonomy proposed is presented in Figure 5. The new structural model illustrates the relationship among big data analytics techniques and their methods according to quantitative and qualitative approaches. A quantitative approach is a strategy for systematic collection, organisation and interpretation of information, this approach is implemented to generate data to allow researchers to quantify the problem by having data sets. Whereas a qualitative approach aims to gain a degree of understanding for the affecting parameters. Based on these definitions the techniques were classified. In this taxonomy, data gathering techniques were counted as a quantitative research approach as shown in Figure 5, whereas clustering, classification, prediction, summarisation and optimisation techniques were all considered as qualitative research approach.

BD analytics methods could have a different classification based on the findings from the previous sections which suggest data gathering techniques belong to descriptive analytic methods and clustering is a good example of a diagnostic analytics method whereas classification, prediction and summarisation techniques are all listed under predictive analytics methods, whereas, optimisation techniques are considered as prescriptive analytics methods. In the new proposed classification the descriptive analytics method belongs to the quantitative research approach, whereas diagnostic, predictive and prescriptive analytics methods belong to the qualitative research approach.

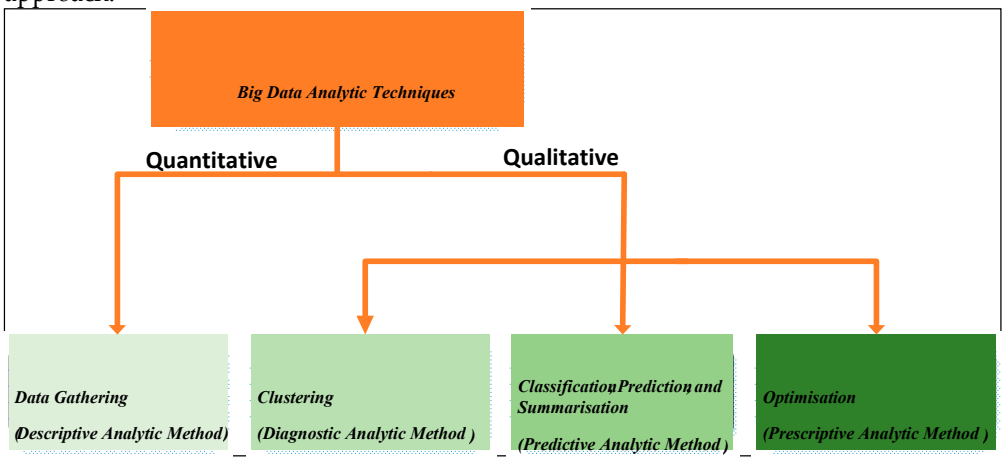

Figure 5. Big data analytic correlation taxonomy.

6. Conclusion

The development of the big data analytics correlation taxonomy for identifying BDA methods and its associated techniques based on research approach methods was achieved through the cross-fertilisation of two stages to understand the correlation between the analytical methods and the correlation of the analytical methods and its associated techniques. These were deemed relevant and helpful in addressing the lack of guidelines for researchers to understand the correlation between big data analytic methods and its associated techniques. Therefore, the main contribution and novelty of this paper is our proposal of a big data analytics taxonomy techniques according to a research approach and big data analytics methods.

This paper has reviewed and reflected big data analytics and presented the current interest in this field. It has defined what was meant by big data, then highlighted the fact that big data analytics has many characteristics and it is worth mentioning that the correlation taxonomy proposed in this paper is not affected by a number of the big data analytics characteristics and should be applicable for any number of Vs'. Furthermore, the paper has investigated the most popular sectors where big data analytics made an impact. After defining the existing analytical methods, this paper made the case for a new analytical methods category and how they complement each other. Design science research has been adopted and five DSRM elements were chosen to accomplish the project, which helped in developing the concept successfully.

Finally, the work presented here is ongoing. The following phases of the research will be to: (1) continue the investigation on the validity of big data analytics characteristics used in industry; (2) continually test and assess the proposed taxonomy in a specific sector; (3) identify and highlight the positive and negative implication examples; and (4) specify the limitations and the difficulties raised; and (5) update the taxonomy based on the outcomes from the previous points.

Author Contributions: All authors have read and agree to the published version of the manuscript. Conceptualization, L.H. and N.S.; methodology, L.H.; validation, L.H. and N.S.; formal analysis, N.S.; investigation, L.H.; resources, L.H. and N.S.; writing—original draft preparation, L.H. and N.S.; writing—review and editing, L.H. and N.S.; visualization, N.S.; supervision, L.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

1. Zaker, J.; Seymour, T.; Berg, K. Big data analytics. *Inf. Syst. J.* **2015**, *16*, 81–90.
2. Hardoon, D.R.; Shmueli, G. *Getting Started with Business Analytics: Insightful Decision-Making*, 1st ed.; Taylor & Francis Group: London, UK, 2014.
3. Evans, J.R. *Business Analytics: Methods, Models, and Decisions*, 2nd ed.; Pearson: Upper Saddle River, NJ, USA, 2015.
4. Orozco, J.; Tarhini, A.; Masa'deh, R. A framework of IS/business alignment management practices to improve the design of IT Governance architectures. *Int. J. Bus. Manag.* **2015**, *10*, 1–12.
5. Sharma, R.; Reynolds, P.; Scheepers, R.; Seddon, P.; Shanks, G. Business Analytics and Competitive Advantage: A Review and Research Agenda, In Proceedings of the 5th International Federation of Information Processing WG 8.3 International Conference on Decision Support Systems, Lisbon, Portugal, 7–10 July 2010.
6. Chen, H.; Chiang, R.; Storey, V. Business intelligence and analytics: From big data to Big impact. *MIS Q. Manag. Inf. Syst.* **2012**, *36*, 1165–1188.
7. Russom, P. Big Data Analytics. Available online: http://download.101com.com/pub/tdwi/Files/TDWI_BPReport_Q411_Big_Data_Analytics_Web.pdf (accessed on 24 December 2019).
8. Lee, C.; Yoon, H. Medical big data: Promise and challenges. *Kidney Res. Clin. Pract.* **2017**, *36*, 3–11.
9. Akerkar, R. (Ed.) *Big Data Computing*; CRC Press, Taylor & Francis Group*: Boca Raton, FL, USA*, 2014; pp. 103–128.

10. Sivarajah, U.; Kamal, M.M.; Irani, Z.; Weerakkody, V. Critical analysis of big data challenges and analytical methods. *J. Bus. Res.* **2017**, *70*, 263–286.
11. Agarwal, R.; Dhar, V. Editorial—Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research. *Inf. Syst. Res.* **2014**, *25*, 443–448.
12. Google Trends. 2019. Available online: https://trends.google.co.uk/trends/trendingsearches (accessed on 20 October 2019).
13. Chen, C.L.P.; Zhang, C.-Y. Data-intensive applications challenges techniques and technologies: A survey on Big Data. *Inf. Sci.* **2014**, *275*, 314–347.
14. Assunção, M.D.; Calheiros, R.N.; Bianchi, S.; Netto, M.A.S.; Buyya, R. big data computing and clouds: Trends and future directions. *J. Parallel Distrib. Comput.* **2015**, *79*, 3–15.
15. Gandomi, A.; Haider, M. Beyond the hype: Big data concepts, methods, and analytics. *Int. J. Inf. Manag.* **2015**, *35*, 137–144.
16. Bagga, S.; Sharma, A. Big Data and Its Challenges: A Review. In Proceedings of the IEEE 2018 4th International Conference on Computing Sciences (ICCS), Phagwara, India, 30–31 August 2018; pp. 183–187.
17. Emani, C.K.; Cullot, N.; Nicolle, C. Understandable Big Data: A survey. *Comput. Sci. Rev.* **2015**, *17*, 70–81.
18. Panimalar, A.; Shree, V.; Kathrine, V. The 17 V's of Big Data. *Int. Res. J. Eng. Technol.* **2017**, *4*, 329–333.
19. Shafer, T. The 42 V's of Big Data and Data Science. 2019. Available online: https://www.elderresearch.com/blog/42-v-of-Big-data (accessed on 10 December 2019).
20. Manyika, J.; Chui, M.; Brown, B.; Bughin, J.; Dobbs, R.; Roxburgh, C.; Byers, A.H. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*; McKinsey Global Institute: Washington, DC, USA, 2011.
21. Duan, L.; Xiong, Y. Big data analytics and business analytics. *J. Manag. Anal.* **2015**, *2*, 1–21.
22. Nguyen, T.; Zhou, L.; Spiegler, V.; Ieromonachou, P.; Lin, Y. Big data analytics in supply chain management: A state-of-the-art literature review. *Comput. Oper. Res.* **2018**, *98*, 254–264.
23. Raguseo, E. Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *Int. J. Inf. Manag.* **2018**, *38*, 187–195.
24. Erl, T.; Khattak, W.; Buhler, P. *Big Data Fundamentals: Concepts, Drivers & Techniques*; Prentice Hall Press: Upper Saddle River, NJ, USA, 2016.
25. MassTLC. Tom Davenport, Automation vs. Augmentation; Big data Summit 2015. Available online: https://www.slideshare.net/MassTLC/tom-davenport-slides-automation-augmentation (accessed on 8 December 2019).
26. Academies, N.I. *Strengthening Data Science Methods for Department of Defence Personnel and Readiness Missions*; The National Academies of Sciences, Engineering and Medicine, The National Academies Press: Washington, DC, USA, 2017.
27. Jing, X.; Yan, Z.; Pedrycz, W. Security Data Collection and Data Analytics in the Internet: A Survey. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 586–618.
28. Cardinale, Y.; Guehis, S.; Rukoz, M. Big Data Analytic Approaches Classification. In *Proceedings of the 12th International Conference on Software Technologies, Madrid, Spain, 24–26 July 2017*; ICSOFT: Madrid, Spain, 2017; Volume 1, pp. 151–162, ISBN 978-989-758-262-2, doi:10.5220/0006437801510162.
29. Davenport, T. Make better decisions. *Harv. Bus. Rev.* **2009**, *87*, 117–123.
30. Peffers, K.; Tuunanen, T.; Rothenberger, M.; Chatterjee, S. A Design Science Research Methodology for Information Systems Research. *J. Manag. Inf. Syst.* **2007**, *24*, 45–77.
31. Aldin, L.; Cesare, S.D. A Literature Review on Business Process Modelling: New Frontiers of Reusability, *Enterp. Inf. Syst. J.* **2011**, *5*, 359–383.
32. Analytics, G. Data Classification and Analysis-Governance Analytics Knowledge Base. 2019. Available online: https://www.governanceanalytics.org/knowledge-base/Main_Tools/Data_classification_and_analysis (accessed on 8 December 2019).
33. Greg, M. Business Analytics and Intelligent: An Introduction and Consideration for Getting Started. Information Services and Technology. Available online: http://dcotech.org/wp-content/uploads/2018/03/BIWhitePaperV5.pdf (accessed on 24 December 2019).
34. Gani, A.; Siddiqa, A.; Shamshirband, S.; Hanum, F. A survey on indexing techniques for big data: Taxonomy and performance evaluation. *Knowl. Inf. Syst.* **2015**, *46*, 241–284.

35. Mohamedali, F.; Saeed, N. *Use of Feed-Forward to Improve Students' Engagement and Achievement*; University of West London: London, UK, 2019.
36. Urbonaviciusa, A.; Saeed, N. IoT leak detection system for building hydronic pipes. *Int. J. Eng. Manuf.* **2019**, *9*, 1–21. ISSN 2305-3631.
37. Raoof, O.; Saeed, N.H.; Abbod, M.F.; Al-Raweshidy, H.S. Intelligent MANET Routing Optimizer. In Proceedings of the 2008 fifth International ICST Conference on Broadband Communications, Networks and Systems, Okinawa, Japan, 25–28 March 2008; pp. 359–365.
38. Abbod, M.F.; Saeed, N.H.; Al-Raweshidy, H.S. Modelling MANET Utilizing Artificial Intelligent. In Proceedings of the IEEE 2008 Second UKSIM European Symposium on Computer Modelling and Simulation, Liverpool, UK, 8–10 September 2008; pp. 117–122.
39. Saeed, N.H.; Abbod, M.F.; Sulamin, T.H.; Al-Raweshidy, H.S.; Kurid H. Intelligent MANET routing protocol selector. In Proceedings of the IEEE 2008 Second International Conference on Next Generation Mobile Applications, Services, and Technologies, Cardiff, UK, 16–19 September 2008; pp. 389–394.
40. Saeed, N.; Abbod, M. Modelling Oil Pipelines Grid: Neuro-fuzzy Supervision System. *Int. J. Intell. Syst. Appl.* **2017**, *9*, 1–11.
41. Saeed, N.H.; Abbos, M.F.; Al-Raweshidy, H.S. IMAN: An Intelligent MANET routing system. In Proceedings of the IEEE 2010 17th International Conference on Telecommunications (ICT), Doha, Qatar, 4–7 April 2010; pp. 401–404.

© 2019 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 (http://creativecommons.org/licenses/by/4.0/).