

Data Science: Trends and Its Role in Various Fields

Dedi Iskanto

Management Business Telecommunications and Informatics, Telkom University, Indonesia

deditaba@telkomuniversity.ac.id

ARTICLE INFO

Research Paper

Article history:

Received: 12 May 2023

Revised: 15 June 2023

Accepted: 12 July 2023



<https://doi.org/10.54099/aijms.v2i2.606>

ABSTRACT

Data Science is a field that has developed rapidly in recent years, which utilizes technology and data analysis methods to produce useful information for various fields. In this paper, we will explain the latest trends in the field of Data Science and their role in various fields, including the fields of business, health, finance, and the public sector and discuss the technologies and algorithms used in Data Science, such as Machine Learning, Natural language Processing., data visualization, Big Data Analytics, and provides examples of how this technology can be used to solve problems in various fields. This paper aims to explore that Data Science has a very important role in various fields, and will continue to develop along with technological developments. The literature study methodology was carried out to collect data and information related to this topic, which aims to ensure that the data analysis carried out is accurate, and can be trusted. The results of the study found that Data Science science and technology has helped a number of parties with its application to certain aspects related to their fields. Thus, this paper can be a useful source of information for professionals and researchers who are interested in Data Science and its role in various fields.

Keywords : Data Science, Natural Data Processing, Data Science Technology, Big Data Analytics, Machine Learning, Artificial Intelligence, Data Visualization.

This work is licensed under a Creative Commons Attribution-Noncommercial 4.0 International License.

INTRODUCTION

Data science and analytics is a rapidly growing field in recent years, which utilizes technology and data analysis methods to make decisions and solve problems (Provost & Fawcett, 2013b). In the current digital era the amount of data generated and collected continues to increase, and this raises the need to analyze data effectively and efficiently, therefore Data Science is the "Sexiest" and most needed job today (Davenport H Thomas & Patil DJ , 2012). Technologies and methods in data science and analytics are constantly evolving and enabling organizations to collect and analyze data more accurately and quickly. Technologies and methods in data science and analytics continue to develop rapidly. Along with the development of computing technology, internet speed, and increasing data storage capacity, organizations can collect and analyze more massive and faster amounts of data. This combination of technologies enables organizations to collect and analyze massive amounts of data quickly. The aim of this journal is to provide insight into the latest trends in data science and analytics, and their applications and benefits in various fields. In this journal, we will explain the technologies and methods in data science and analytics such as machine learning, big data analytics, data visualization, and natural language processing. Apart from that, it will also discuss the

applications and benefits of data science and analytics in fields such as business, health, finance, and the public sector. By reading this journal,

Methodology

The methodology used in this journal is the literature method used in the process of searching, collecting, evaluating, and analyzing information from written sources such as books, journals, articles, and other documents. The literature method is one of the important aspects of data science because it allows researchers to understand the field they are researching thoroughly and build new knowledge based on previous findings. Trends in the use of literature methods in Data Science continue to change with the development of technology and needs. Data and information from these literature sources will then be analyzed and presented in this journal.

FINDINGS AND DISCUSSION

Definition of Data Science

According to Sarker (2021) Data Science is like a combination of statistics, data analysis, and related methods to analyze certain phenomena using data. A Data Science Practitioner is called a Data Scientist, namely an analytical model practitioner who solves business problems by combining advanced approaches and data visualization tools to find patterns in data, generate Insights (insights/information) from data, and communicate these Big Data-based insights to decision makers business decisions, such as bosses etc., in a way that allows them to understand insights derived from Big Data (Chatfield et al., 2014, p. 9). Data Science includes techniques such as machine learning, AI, Deep Learning, and Data Mining. The purpose of Data Science is to solve business problems,

Definition of Analytics

Analytics is a method or data processing technique that produces useful information for decision making. According to Sarker (2021), Analytics uses conventional theory (such as classical, empirical, or logical statistics), technology, and tools to extract information for practical purposes. The goal of analytics is to get better insights about data, use those insights for better decision making, and optimize overall business performance. Data analysis and statistics are key in this process. "Data mining" is another term popularized over the last decade, which has a similar meaning to the term data analysis.

Difference between Data Science and Analytics

Data Science and Analytics are often used interchangeably, even though the two differ in their approach and purposes for which they are used. Expert opinion on the definition of Data Science and Analytics also varies. According to Davenport and Patil (2012), Data Science focuses more on collecting and processing big data, developing models and algorithms that can predict patterns and behavior based on historical data, and making fast decisions, while Analytics focuses more on data processing, using mathematical techniques. and statistics to obtain useful information for business decision making. The two are interrelated and are often used together to produce more comprehensive and in-depth information from data (Bollen, nd). According to Provost and Fawcett (2013),

In an article written by Will Hillier, a Public Relations (PR) and seasoned article writer, Data Science is a scientific discipline that explores aspects of all kinds of unstructured data and how it relates to the world. Data Analytics is a key process in the field of data science, which is used to create meaningful insights based on structured data sets (Hillier, 2022).

The Importance of Data Science and Analytics

a. In Decision Making

Data Science and Analytics play an important role in decision making, especially in the rapidly growing digital era. According to Provost and Fawcett (2013), Data Science and Analytics enable

organizations to make more accurate and fact-based decisions, by analyzing data collected from various sources. In a business context, according to Chen, Chiang, and Storey (2012) decision-making based on data science and analytics can help organizations improve operational efficiency and effectiveness, find new market opportunities, optimize marketing strategies, and improve customer experience. For example, according to Chen, Chiang, and Storey (2012), data science and analytics can help organizations to identify patterns of consumer behavior and customer preferences, so they can develop more effective marketing strategies and increase customer retention. In addition, data science and analytics can also assist organizations in making better strategic decisions, for example in determining more effective investments and optimizing resource allocation. (Chen et al., 2012)

b. Within the Organization

Apart from enabling more accurate decision making, Data Science and Analytics also provide great benefits to organizations. James, Witten, Hastie, and Tibshirani (2013) stated that the main benefit of data science and analytics is its ability to transform complex and unstructured data into useful insights for organizations (James et al., 2013)

For example, data science and analytics can help organizations optimize business and operational processes, thereby increasing efficiency and productivity. In one of its journals, McKinsey Global Institute (2018, p. 22) reports that they are working with clients to build analytics-driven organizations, helping them develop strategies, operations, and capabilities for fast and sustainable impact from analytics. McKinsey Global Institute says that deep learning technologies in data science and analytics can help organizations automate business processes, such as image and voice recognition, and document processing. This can save costs and improve service quality. Besides that,

Data Science and Analytics Technologies

a. Machine Learning Technology

Machine Learning technology is a technology that enables machines to learn from data, improve themselves, and provide predictions or results based on patterns found in data (Dicoding Intern, 2020). According to provost and Fawcett (2013a), Machine Learning technology is used to predict customer behavior, optimize business decisions, and carry out supervision or monitoring in real-time. Machine learning technology can be used to solve various problems, such as regression, which is an analytical technique to determine the relationship or relationship between two or more variables to reduce errors in forecasting value guesses, classification, which is a technique for classifying or classifying many unlabeled items into a set of different classes. , and grouping i.e. grouping is done if we want to find the specified phrase data. Clusters are collections of data or objects that are similar in one group or group and different from objects in other groups (Afifah, 2020). In addition, Machine learning is able to detect hidden structures in data and use them to make intelligent decisions (Meylani & Negara, 2022, p. 209).

b. Big Data Analytics Technology

According to Chen, Chiang, and Storey (2012, p. 1166) Big Data Analytics technology is a technology used to process, analyze, and extract insight from very large and complex data. Big Data Analytics can take advantage of the opportunities offered by large amounts of data and specific analysis needed in many fields (Chen et al., 2012, p. 1168). Big Data Analytics aims to generate valuable information that can be utilized in various fields to improve business strategy and decision making. The big data analytics process includes discovering, managing, and analyzing patterns and inferences from useful big data. (Ahmad et al., 2018, p. 2).

c. Data Visualization Technology

Data Visualization Technology is a technology used to visualize data as a representation of the original data so that it is easy to understand and present (Ilyas & Pudjiantoro, nd). Data visualization technology can show patterns, trends, and connections or relationships between various data elements (Shadare et al., 2016, pp. 11–12). James, Witten, Hastie, and Tibshirani (2013) said that Data Visualization Technology can assist decision making by facilitating understanding of complex data and providing valuable insights. Data Visualization Technology can be used to visualize data in the form of scatter plots, histograms, pie charts, and so on. For example, Data Visualization technology is used in a data analytics project in sales to identify sales patterns and define marketing strategies.

d. Natural Language Processing Technologies

Natural language processing technology (Natural Language Processing) is a branch of science that enables computers to understand, process, and produce human language naturally (Verspoor & Cohen, 2013). This technology is very important in data analysis because a lot of data is generated by humans in the form of text or voice. In natural language processing, there are several techniques such as tokenization, lemmatization, entity recognition, and language modeling (Jurafsky & Martin, 2019). This technology has many applications, including in chatbots, sentiment analysis, and document classification.

One of the studies on natural language processing was conducted by Ji et al. (2019) who developed a new algorithm for document classification. This algorithm uses the transfer learning method and aims to improve the accuracy of document classification in various languages. The results of the research show that the developed algorithm is capable of achieving higher accuracy than the existing document classification algorithm.

Apart from that, natural language processing technology is also used in chatbot development. One example of research related to chatbots is that conducted by Toba et al. (201) who developed a chatbot to facilitate communication between patients and doctors. This chatbot can perform sentiment analysis of patient complaints and provide advice on appropriate diagnosis and treatment. The results showed that the developed chatbot was able to increase patient satisfaction and reduce the workload for doctors.

Application of Data Science and Analytics in Various Fields

a. Application of Data Science and Analytics in the Business Sector

The application of Data Science and Analytics in business can help companies make decisions based on data that has been processed and analyzed. Data Science can help the company's operational efficiency. An example of the application of Data Science and Analytics in the business sector is in marketing, where companies can utilize Machine Learning technology to segment markets and predict consumer needs. In addition, companies can also use Big Data Analytics technology to process transaction data and optimize their business processes (Provost & Fawcett, 2013b, p. 5).

According to Provost and Fawcett (2013), the application of Data Science and Analytics in business can also help companies to manage risk. Companies can utilize Machine Learning technology to detect fraud and predict credit risk. This can help companies make more accurate and efficient business decisions.

Among the several methods that can be applied in decision making or Decision Making is Regression Analysis. This method is useful for forming a model that can represent the relationship between the dependent and independent variables. In the context of product marketing, the dependent variable can be product sales, while the independent variables can be factors such as product prices, promotions, and demographic factors such as age and income. . The use of this method is not limited to just one thing, because this method can also be used in analyzing Demand Statistics. Demand Statistics itself is a set of statistical procedures that are intended to reveal the important factors that affect sales and their relative influence. Typically, this method examines factors such as price, income, population numbers, and promotional efforts in Demand Statistics analysis. (Kotler & Wong, n.d., p. 371).

b. Application of Data Science and Analytics in the Health Sector

The application of Data Science and Analytics in the health sector can assist doctors and researchers in analyzing patient health data and making more accurate decisions. In his journal "The role of data science in healthcare advancements: applications, benefits, and future prospects", Subrahmanya et al. (2022, p. 1467) explains There are six areas of application of analytics in healthcare including disease surveillance, healthcare management and administration, privacy protection and fraud detection, mental health, public health, and Pharmacovigilance which is a term that refers to a set of activities related to monitoring and management of side effects or other problems related to drug use.

Still from the same journal, Subrahmanya et al. (2022, p. 1467) explained that image processing or image processing in healthcare data offers valuable knowledge about the anatomy and function of organs as well as identifying diseases and health conditions of patients. Current techniques have been used for imaging of organs, identification of lung tumors, diagnosis of spinal deformities, detection of arterial stenosis, detection of aneurysms, etc.

c. Application of Data Science and Analytics in Finance

The application of data science and analytics has also proven to be very beneficial in finance. One example is in the management of financial risk. Data science can assist companies in modeling and analyzing financial risk data so that they can determine a more effective risk management strategy. In addition, data science can also assist in the investment decision-making process by analyzing market data and predicting future market behavior (Pompiliu et al., 2022).

In a journal entitled "Using Data Mining In The Sentiment Analysis Process On The Financial Market" by M. Pompiliu et al. (2022) stated that data science can assist in making investment decisions in the stock market. In their journal, they use a data mining approach and sentiment analysis to predict stock price movements. To get a sentiment score, they retrieve text data from the Finviz platform where the polarity of opinion can be extracted. They also use the Valence Aware Dictionary for Sentiment Reasoning (VADER), by running Python scripts using the BeautifulSoup library. After that, they use Pandas (Python Data Analysis Library) to analyze and obtain a sentiment score on the article title. The results show that the script is able to generate sentiment scores for a range of selected stocks, while also showing graphical charts for past and future trends of the stocks, in terms of overall opinion on the stock's performance. The results of their research indicate that this approach can direct investors to choose the right investment decision.

d. Application of Data Science and Analytics in the Public Sector

Apart from business and finance, data science and analytics can also be applied in the public sector, such as government and public services. One example is in public policy decision-making based on data. By using data science and analytics, the government can obtain more accurate and detailed information about the situation and needs of the community, so that it can determine more effective public policies (Kwan, 2021).

Data science can also help improve government performance by processing and analyzing existing data. By using data science, governments can collect, store, process and analyze data more efficiently and accurately. Governments can create dashboards and data visualizations to assist in data-driven decision making. In addition, data analysis can also be used to optimize the distribution of funds between departments with more accurate predictions and forecasts. By increasing the efficiency of government performance, the positive impact that can be felt by the people is better and more efficient public services. It can also help increase public trust in the government and improve the quality of life of society as a whole. Therefore,

Next is the Smart City Concept. Smart City is a concept that is the future of all countries with the aim of increasing overall efficiency. Many countries have implemented the Smart City concept, including America and China, which have implemented this concept by using data science in government. An example is the Traffic Management system in Los Angeles, America, which uses a Machine Learning model to recognize patterns in traffic jams in real-time. With this concept, the red light can automatically change depending on the level of congestion, so it can help reduce congestion and noise pollution on the streets (Kwan, 2021).

e. Application of Data Science and Analytics in Education

In the journal "Learning Analytics: Measurement Innovations to Support Student Success" by John Fritz and Brandy A. Randall, it is discussed about the application of data science to measure student success. This journal discusses the use of data science to analyze learning data, including data on student behavior, interactions in the classroom, and learning activities outside the classroom.

In this journal, the authors discuss how data science can help make better decisions in education. Data science can help identify students who are at high risk of not being successful, provide feedback to students on their progress, and help develop more effective learning programs.

In addition, this journal also discusses various data science techniques that can be used in learning analysis, such as longitudinal data analysis, classification analysis, and cluster analysis. In this paper, the authors conclude that data science is an important tool for increasing educational effectiveness, and that its use should be integrated into overall teaching and learning strategies.

In conclusion, the journal "Learning Analytics: Measurement Innovations to Support Student Success" discusses the use of data science in education. The authors explain that data science can help identify students who are at high risk of not succeeding, provide feedback to students on their progress, and assist the development of more effective learning programs. This journal also discusses various data science techniques that can be used in learning analysis and emphasizes that their use must be integrated into overall teaching and learning strategies.

CONCLUSION

Data Science and Analytics is an important and rapidly growing field in today's digital era. Good data analysis skills will help in making better, more effective and accurate decisions. The difference between the two is that Data Science focuses more on collecting, processing and analyzing big data, developing models and algorithms, and making fast decisions. Meanwhile, Analytics focuses more on processing data with mathematical and statistical techniques to obtain information that is useful for making business decisions. Both involve techniques and algorithms such as machine learning, statistical analysis, and data mining. In addition, the application of Data Science and Analytics is useful in improving quality in various sectors such as business, health, finance and the public sector.

Within the business itself, Data Science and Analytics can help companies improve the quality and efficiency of business operations, such as market segmentation and risk management. Then, in the field of health Data Science and Analytics can assist doctors and researchers in analyzing patient health data and predicting disease. And in the field of finance, Data Science and Analytics can assist companies in modeling and analyzing financial risk data and in the process of making investment decisions. Meanwhile in the public sector, Data Science and Analytics can assist the government in optimizing the use of resources and increasing the efficiency of public services, as well as in making public policy decisions based on data.

Reference

Afifah, L. (2020). What is Regression, Classification and Clustering? - Data SciencePy. <https://ilmudatapy.com/apa-itu-regresi-klasifikasi-dan-clustering-klasterisasi/>

- Ahmad, S., Singh, P., & Sagar, AK (2018). A Survey on Big Data Analytics. Proceedings - IEEE 2018 International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2018, 4, 256–260. <https://doi.org/10.1109/ICACCCN.2018.8748774>
- AltexSoft. (2021). Data Science vs Machine Learning vs AI vs Deep Learning vs Data Mining Know the Differences. <https://www.altexsoft.com/blog/data-science-artificial-intelligence-machine-learning-deep-learning-data-mining/>
- Bollen, R. (n.d.). Competencies of the Sexiest Job of the 21st Century: Developing and Testing a Measurement Instrument for Data Scientists. 66. Retrieved March 11, 2023, from <http://arno.uvt.nl/show.cgi?fid=144899>
- Chatfield, AT, Shlemoon, VN, Redublado, W., & Rahman, F. (2014). Data scientists as game changers in big data environments. <https://ro.uow.edu.au/eispapers>
- Chen, H., Chiang, RHL, & Storey, VC (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly: Management Information Systems*, 36(4), 1165–1188. <https://doi.org/10.2307/41703503>
- Chui, M., Francisco, S., & Manyika, J. (2018). NOTES FROM THE AI FRONTIER INSIGHTS FROM HUNDREDS OF USE CASES. <https://www.mckinsey.com/~media/mckinsey/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning/notes-from-the-ai-frontier-insights-from-hundreds-of-use-cases-discussion-paper.ashx>
- Davenport H Thomas, & Patil DJ (2012). Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review*, Oct. 2012, 72.
- Coded Internal. (2020). What is Machine Learning? Along with Understanding and How it Works - Dicoding Blog. <https://www.dicoding.com/blog/machine-learning-is/>
- Hillier, W. (2022). Whats the Difference Between Data Science, Data Analytics, and Machine Learning. <https://careerfoundry.com/en/blog/data-analytics/data-science-vs-data-analytics-vs-machine-learning/#:~:text=Data science is a scientific, on sets of structured data.>
- Ilyas, R., & Pudjiantoro, TH (nd). Data Visualization on Complaint Management System and Survey Engine. 217–222.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning. <http://www.springer.com/series/417>
- Jurafsky, D., & Martin, JH (2019). *Speech and Language Processing An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition Third Edition draft Summary of Contents.*
- Kotler, P., & Wong, V. (nd). Principles of Marketing. <http://library.wbi.ac.id/repository/212.pdf>
- Kwan, W. (2021). 6 IMPORTANT ROLE OF DATA SCIENCE IN GOVERNMENT.pdf. <https://algorit.ma/blog/6-peran-data-science-government/>
- Meylani, A., & Country, ES (2022). Health Prediction Application Using Machine Learning. *Journal of Jupiter*, 14(2), 208–215. <https://jurnal.polsri.ac.id/index.php/jupiter/article/view/5055>
- Pompiliu, M., Andreea, R., & Alexandru, D. (2022). USING DATA MINING IN THE SENTIMENT ANALYSIS PROCESS ON THE FINANCIAL MARKET. 11(1), 36–58.

- Provost, F., & Fawcett, T. (2013a). *Data Science for Business*. <https://www.oreilly.com/library/view/data-science-for/9781449374273/>
- Provost, F., & Fawcett, T. (2013b). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, 1(1), 51–59. <https://doi.org/10.1089/big.2013.1508>
- Sarker, IH (2021). Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-Making and Applications Perspective. *SN Computer Science*, 2(5), 1–22. <https://doi.org/10.1007/S42979-021-00765-8/FIGURES/6>
- Shadare, AE, Musa, SM, & Akujuobi, C. (2016). Data visualization. December.
- Subrahmanya, SVG, Shetty, DK, Patil, V., Hameed, BMZ, Paul, R., Smriti, K., Naik, N., & Somani, BK (2022). The role of data science in healthcare advancements: applications, benefits, and future prospects. *Irish Journal of Medical Science*, 191(4), 1473–1483. <https://doi.org/10.1007/S11845-021-02730-Z/FIGURES/5>
- Verspoor, KM, & Cohen, KB (2013). *Encyclopedia of Systems Biology*. *Encyclopedia of Systems Biology*, June 2018. <https://doi.org/10.1007/978-1-4419-9863-7>