DATA SCIENCE IN FINANCIAL SERVICES: RISK ANALYSIS, FRAUD DETECTION, AND MARKET INSIGHTS

Ms.Radha.T

Designation: Assistant Professor Department: Commerce Institute: St. Claret College
District: Bangalore City: Bangalore
State: Karnataka

Dr D Appala Naidu

Assistant Professor Department of Economics, Atma Ram Sanatan Dharma College,
University of Delhi
Mail: danaidu@arsd.du.ac.in

NIDHI JINDAL

Designation : Assistant Professor Institute : COER UNIVERSITY
Email : nidhijindal956@gmail.com

Dr.Nikhil Sirohi

Designation: Assistant Professor Department: MBA
Institute: Centre for Management Development, Modinagar District: Ghaziabad
City: Modinagar State: Uttar Pradesh

Dr. Kavita Singh

Designation: Asst. Professor Department: MBA
Institute: IMS Engineering College District: Ghaziabad City: Ghaziabad
State: Uttar Pradesh

Abstract

Introduction: One area of interdisciplinary academic study that is thought to offer deep business insights is data science. The study's objective is to evaluate the value of market, risk, and fraud analysis in data science for enhancing financial services.

Literature review: It is well recognised that using data science to conduct a risk analysis influences commercial organisations' decision-making when it comes to starting new projects. To enhance the banking sector's services, data science's fraud detection division is required. A more complex sort of data science called market analysis can give

Available online at: https://jazindia.com

businesses important insights.

Methodology: A study's methodological choice is crucial since it determines the validity and dependability of the research endeavour. In this instance, it may be said that the study adhered to a theoretical analysis as including theories aids in the investigation of the research problem.

Findings and analysis: In the context of data science, risk prediction models are helpful since they facilitate the start of statistical analysis. In this instance, it should be noted that quick fraud detection is made possible by the efficiency of data processing and encoding. Thus, putting in place such a network can enhance market data analysis by helping to unravel data layers and patterns.

Conclusion: The purpose of the study is to critically examine the effects of fraud detection, risk analysis, and market analysis. This project has focused on leveraging data science to improve financial services. It has been demonstrated that the study's critical evaluation has a significant influence on how data science is applied in the financial industry.

Keywords: Data science, Risk analysis, Market analysis, Fraud detection, Neural Networks.

Introduction
Background

Data science is considered to be a branch of interdisciplinary academic fields that provide extensive insights into a business. As stated by Dicuonzo et al. (2019), the benefits of data science can be obtained by using analytical prospects to assess risks and detect fraudulent cases. Another benefit is market insights that can improve the business operation methods to gain competitive advantages.
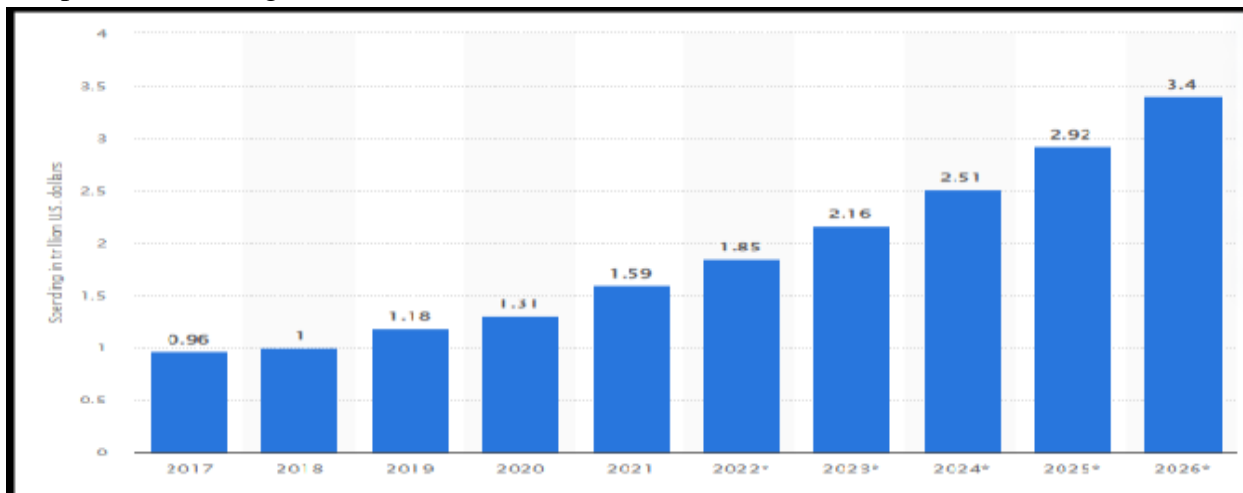

Figure 1: Global expenditure of digital transformation 2017-2026
(Source: Statista, 2023)

Figure 1 shows the worldwide growing expenditure for digital transformation showing that the cost was 1.59 trillion USD in 2021 which is expected to reach 3.4 trillion USD by 2026 (Statista, 2023). Thus, it has demonstrated an increasing financial expenditure that can be challenging for certain businesses.

Aim

The aim of the study is to analyse the importance of risk analysis, fraud detection, and market analysis in data science for improving financial services.

Research objectives

The research objectives of this study have been mentioned below.

RO1: To assess the importance of risk analysis in data science for financial services

RO2: To critically discuss the influence of fraud detection in data science to improve financial services

RO3: To evaluate the impact of market analysis as a part of data science in financial services

RO4: To examine the challenges of using data science for the financial services

Research questions

The research questions of this research project have been outlined below.

RQ1: What is the significance of risk analysis in data science to enhance financial services?

RQ2: What is the impact of fraud detection in data science to encourage financial services?

RQ3: What is the influence of market analysis regarding data science for financial services?

RQ4: What are the issues of data science and its implementation in financial services?

Literature review

Risk analysis is important in data analysis for improving financial services

It has been known that performing a risk analysis through data science is influential for business organizations to initiate decisions for new projects. As per the view of Chakraborty & Ghosh (2020), financial applications are complex at times which can be done more accurately with the help of data analysis. It can detect risk analysis and the actions required to be taken to protect the business interests. Moreover, solutions for the risks can also be identified as a precaution of the risks. Thus, a balance can be maintained between the reduction of risks and the potential risks that are essential for business improvement. However, as argued by Sahani et al. (2019), the structure of the data plays an important role in determining the level of risk associated with the business. Hence, financial decisions and customer services can be improved rapidly projecting a business improvement.

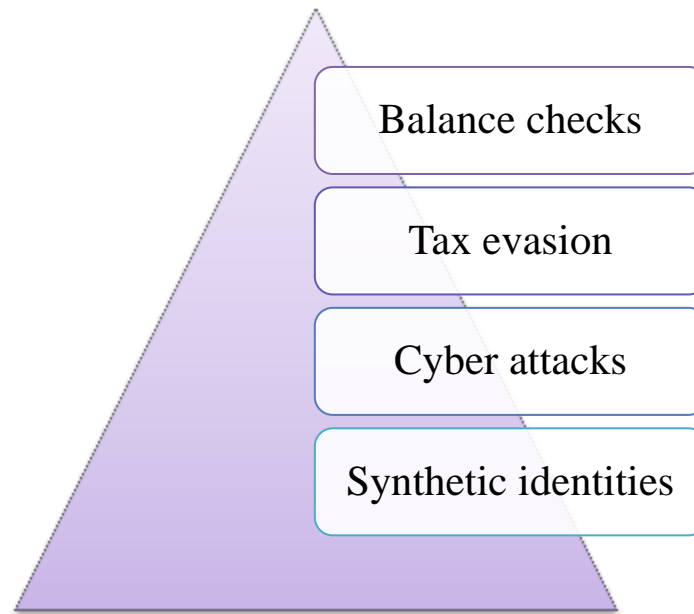**Fraud detection is an essential component of data science regarding its usage in financial services**

Figure 2: Necessity of fraud detection
(Source: Influenced by Khatri et al. 2020)

Fraud detection is necessary in the branch of data science to improve the services of the financial sector. In this case, it can be mentioned that fraudulent credits must be avoided during the transactions (Khatri et al. 2020). In addition to that, balance checks by forging, tax evasion, cyber attacks, and synthetic identities are the major financial threats. Data science is responsible for collecting and storing huge amounts of information and must protect them. In this way, financial services can be significantly improved to detect fraud in this field. On the contrary, as contrasted by Ni et al. (2020), the susceptible characteristics of data science must be used carefully based on the applicability and the right places to be implemented. As a result, the outcome of using data science can be seen through effective fraud detection techniques.

Market analysis is crucial in data sciences to enhance financial services quality

Market analysis is a sophisticated version of data science that can provide valuable information to the business. The understanding of the market and the customers can be known with the help of this specific technique (Raschka et al. 2020). It helps in increasing customer satisfaction which is important to prove a higher efficiency of the business. Moreover, digital marketing has become a significant concept that can be tactically handled by data science. Consequences, the marketing strategies, and their individual performances can be evaluated by applying this method. On the other hand, as opposed to Martínez-Plumed et al. (2019), financial markets can be optimized with data science by knowing the investment history. Thus, the driving factors of customer actions can be known with the appropriate application of this specific component.

Identification of challenges to implementing data science in the financial sector

Communication is the major challenge that is visible among the data scientists that they face while analyzing data and conveying the results. As suggested by Lähnemann et al. (2020), the application of data science tools and techniques must be known by the users which can become problematic without appropriate knowledge and skills. Additionally, the base ideas must be understood with the essential order to be implemented as well. Consequently, business enterprises are continuously trying to mandate skilled employees and organize further training sessions for them (Bhattarai et al. 2019). Data sources are another issue that is present in data science imposing challenging situations in the financial sector. The accuracy of data analysis is a potential issue that exists in data science and for the users as well. Henceforth, recognizing the exact problems is helpful for the data collection and analytical process of the financial analysis.
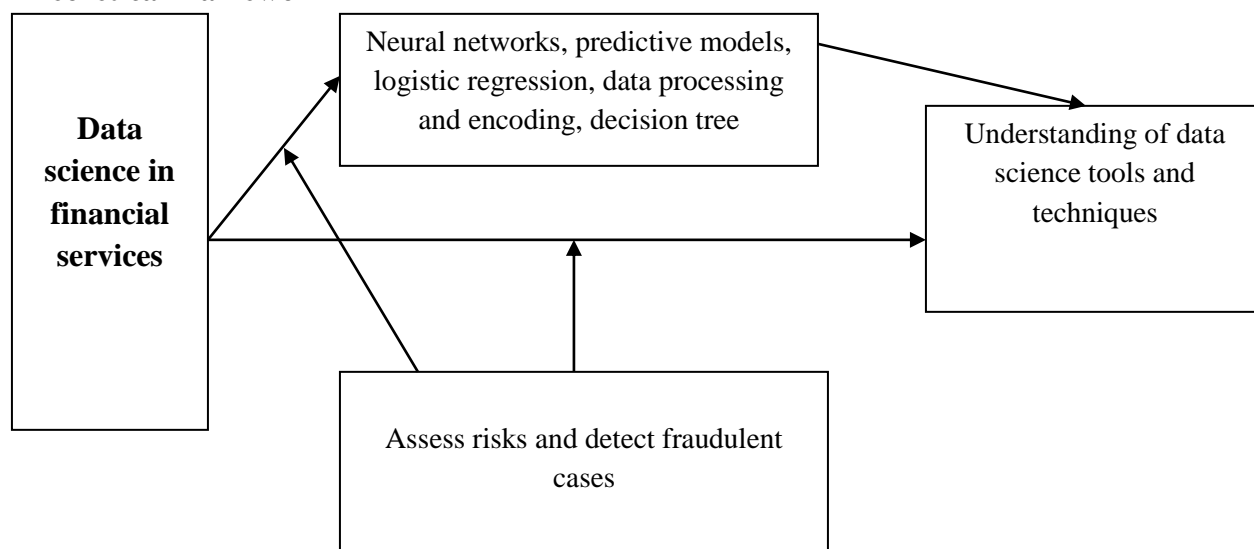
Theoretical framework

Figure 3: Theoretical framework
(Source: Influenced by Dicuonzo et al. 2019)

Figure 3 is the representation of the theoretical framework which has provided a brief and concise idea about the study design. It has shown a connection among the chosen theories and concepts with the major attributions of the study. Hence, it has projected its importance in discussing the influence of data science in the financial sector.

Literature gap

A number of vital gaps in the collected literature have been recognized and will be filled during the current study and its completion. For instance, it can be concluded that the earlier research did not provide statistical knowledge that would have allowed for more accurate proof of the facts (Akter et al. 2022). On the other hand, the assessments from the financial sector and the data science have not addressed the essential influential factors, which have represented a significant gap in these studies.

Methodology

The selection of the methodology for a study is important as it is responsible for the reliability and validity of the research project. In this case, it can be mentioned that this study has

followed a ***theoretical analysis*** as it is helpful in investigating the research problem by incorporating theories (Teo, 2022). The initial data and information in such studies can assist in obtaining the results. Moreover, specification of the research objectives has been done based on the variables such as the factors influencing data science. In this way, themes have been reviewed and explored to discuss the key terms associated with the study.

Findings and analysis

Theme 1: The implementation of prediction models is important to overcome the challenges of risk analysis

Risk prediction models are useful in the context of data science as they help in initiating statistical analysis. As demonstrated by Escher et al. (2019), the analytical techniques and algorithms of Machine Learning can identify the patterns of the dataset. Different business risks can be identified by implementing this specific prediction model to gain further business opportunities. However, as argued by Kou et al. (2019), Predictive models can find future potential risks in a business that can help in finding solutions as precautions. As a result, the business challenges can be overcome by predictive models in business as well as gain opportunities from the market.

The effectiveness of predictive models can be impactful for the financial sector as future financial risks can be identified with this technique. Derived information and analytical power of such models will help the financial sector in estimating the level of risk in the future (Alber et al. 2019). This theoretical model is supportive of identifying the mathematical process to be applied for predicting the list levels of future events. It can also support the analysis of the patterns of the input data, which is another crucial component of predictive analytics (Sujath et al. 2020). Therefore, forecasting such risky activities can analyze their behaviour and potential impact on the business process that can be prevented in the financial business sector. The financial activities of a business can use prediction tools and appropriate techniques based on the list levels with the estimated prognosis.

Theme 2: Data processing and encoding and logistics regression can help in fraud detection while using data science in the financial sector

Fraud detection in the financial sector is crucial as it involves public and money transaction documentation. As illustrated by Yao et al. (2019), discovering fraud in the financial business can be helpful to overcome a significant problem of the failure of specialized business operations. In this case, it can be mentioned that The effectiveness of data processing and encoding can be helpful, in detecting fraud in a short period of time. Data mining as a process of data encoding is supportive of clustering and data segmentation (Ngairangbam et al. 2022). As a result, the association of data can be found that can automatically detect the risks in data analysis and signify the interests.

On another note, the theoretical knowledge of logistic regression is found to be effective for fraud identification. It is widely supervised as a learning technique with the help of regression features to initiate decisions (Zhang et al. 2019). The categorical division of data is supportive of labelling data as fraud or non-fraud in the financial business context. It can be detected during a

transaction process which has been proven to be one of the major benefits of using this model. There are different sets of parameters used in detecting fraud cases such as the data input followed by identity, amount, location, and frequency to gain the output (Tingfei et al. 2020). In this way, the classification of any unusual activities can be found in a fraud transaction.

Theme 3: Neural networks can be implemented to enhance the market analysis process in data science

Neural networks are considered to be effective as they can apply techniques based on their experience and the recognition of patterns. According to Shahid et al. (2019), the prediction of trends in data analysis is constructive due to its implications of enhancing the business process. As a result, the implementation of such a network can improve the data analysis of the market to decode the layers and patterns of data. On the contrary, as contrasted by Manoharan & Sathesh (2020), the nodes and arrangements of neural networks are virtual and need to be handled with expertise. Hence, the utilization and the outcome are entirely dependent on hidden layers of data in this network that can be defined by the input and output data.

The decision-making of the financial business sector can be strengthened with the appropriate application of neural networks. As stated by Tingfei et al. (2020), neural networks use statistical formulas based on the data characteristics that are presented through the networks. Consequently, a trial-and-error process continues to have a general prediction of data. It can also specify the accuracy level of data analysis which is significant in determining the market analysis. Contrarily, as opposed to Zhang et al. (2019), the success of applying neural networks is determined by the applied design. In this way, the input data can be trained to project the maximum output by optimizing information for market analysis.

Theme 4: Using a decision tree can be helpful in overcoming the challenges of data science

Decision trees in data science have been observed to be a constructive method of interpreting data to reduce errors. As per the view of Ngairangbam et al. (2022), both numerical and categorical data can be analyzed by using this technique to outline robust and missing values. The classification issues of data can also be overcome by applying this particular branch of data science. Moreover, objects of data can also be categorized based on the learning features and regression problems. However, as argued by Yao et al. (2019), such data can be overfitting for the analytical process and can be improved with trained data and its proper application. In this way, the generalization of data can be done in a systematic way in order to ensure the swift run of data science.

A continuous predictive data analysis process can be approached with the usage of this theoretical model of data science. Additional learning features can also be included in this system to solve the regression issue (Sujath et al. 2020). Thus, the unforeseen data problems can be avoided by determining the growth of the decision tree. Other attributes of data can be handled by using decision trees by approaching selective measures of the trained data (Alber et al. 2019). As a result, the financial sector can have cost-effective approaches as well as the improvement of computational efficiency.

Discussion

The findings of this study have suggested that data science has become one of the most crucial aspects of today's business world. According to Akter et al. (2022), meaningful insights into the business can only be obtained by using data science which is a multidisciplinary approach. Moreover, the importance of risk analysis, market analysis, and fraud detection has also been elaborated through this study. Moreover, it has been found that this section of technology can help in improving financial services through statistics and mathematics (Pejić Bach et al. 2019). Various prospects and algorithms of data science can enhance the identification of customer transaction data and relevant information from the stock market as well.

The implementation of data science can change the dimension of financial services as it revolutionizes the operations of the business market. As stated by Araz et al. (2020), efficient business operations in the financial sector can be done by having a customer-centric approach and increasing profitability. Furthermore, data science has been proven to be impactful for the financial sector in designing the reporting system and ensuring its accuracy (Hallikainen et al. 2020). As a result, the decision-making of this sector can be made stronger with the specialized skills of data science and its various aspects.

Conclusion

The study has been conducted to critically discuss the impact of market analysis, risk analysis, and fraud detection. The focus of this study has relied on financial service improvement by using data science. The critical assessment of the study has been proven to be impactful in determining the outcome of implementing data science in the financial sector. Theoretical understanding in this study has helped in discussing different components in a critical manner. In this case, it can be mentioned that the theoretical analysis with thematic discussion has played a significant role in improving the outcome of the study. The research questions of the research project have been answered by following this method during the context of the study.

References

Akter, S., Michael, K., Uddin, M. R., McCarthy, G., & Rahman, M. (2022). Transforming business using digital innovations: The application of AI, blockchain, cloud and data analytics. *Annals of Operations Research*, 1-33.https://ro.uow.edu.au/gsbpapers/589

Alber, M., Buganza Tepole, A., Cannon, W. R., De, S., Dura-Bernal, S., Garikipati, K., ... & Kuhl, E. (2019). Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. *NPJ digital medicine*, *2*(1), 115.https://www.nature.com/articles/s41746-019-0193-y

Araz, O. M., Choi, T. M., Olson, D. L., & Salman, F. S. (2020). Data analytics for operational risk management. *Decis. Sci.*, *51*(6), 1316-1319.https://www.weforum.org/reports/the-global-risks-report-2020

2021

Bhattarai, B. P., Paudyal, S., Luo, Y., Mohanpurkar, M., Cheung, K., Tonkoski, R., ... & Zhang, X. (2019). Big data analytics in smart grids: state- of- the- art, challenges, opportunities, and future directions. *IET Smart Grid*, *2*(2), 141-154.https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/iet-stg.2018.0261

Chakraborty, T., & Ghosh, I. (2020). Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: A data-driven analysis. *Chaos, Solitons & Fractals*, *135*, 109850.https://doi.org/10.1016%2Fj.chaos.2020.109850

Dicuonzo, G., Galeone, G., Zappimbulso, E., & Dell'Atti, V. (2019). Risk management 4.0: The role of big data analytics in the bank sector. *International Journal of Economics and Financial Issues*, *9*(6), 40-47.https://ricerca.uniba.it/bitstream/11586/250091/1/2019_Dicuonzo%20et%20al_Risk%20Management%204.0%20The%20Role%20of%20Big%20Data%20Analytics%20in%20the%20Banking%20Sector.pdf

Escher, S. E., Kamp, H., Bennekou, S. H., Bitsch, A., Fisher, C., Graepel, R., ... & Kroese, D. (2019). Towards grouping concepts based on new approach methodologies in chemical hazard assessment: the read-across approach of the EU-ToxRisk project. *Archives of toxicology*, *93*, 3643-3667.https://link.springer.com/article/10.1007/s00204-019-02591-7

Hallikainen, H., Savimäki, E., & Laukkanen, T. (2020). Fostering B2B sales with customer big data analytics. *Industrial Marketing Management*, *86*, 90-98.https://doi.org/10.1016/j.indmarman.2019.12.005

Khatri, S., Arora, A., & Agrawal, A. P. (2020, January). Supervised machine learning algorithms for credit card fraud detection: a comparison. In *2020 10th international conference on cloud computing, data science & engineering (confluence)* (pp. 680-683). IEEE.https://ksra.eu/wp-content/uploads/2020/07/khatri2020.pdf

Kou, G., Chao, X., Peng, Y., Alsaadi, F. E., & Herrera Viedma, E. (2019). Machine learning methods for systemic risk analysis in financial sectors.https://doi.org/10.3846/tede.2019.8740

Lähnemann, D., Köster, J., Szczurek, E., McCarthy, D. J., Hicks, S. C., Robinson, M. D., ... & Schönhuth, A. (2020). Eleven grand challenges in single-cell data science. *Genome biology*, *21*(1), 1-35.https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-1926-6

Manoharan, D. S., & Sathesh, A. (2020). Population based meta heuristics algorithm for performance improvement of feed forward neural network. *Journal of Soft Computing Paradigm*, *2*(1), 36-46.https://doi.org/10.36548/jscp.2020.1.004

2022

Martínez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernández-Orallo, J., Kull, M., Lachiche, N., ... & Flach, P. (2019). CRISP-DM twenty years later: From data mining processes to data science trajectories. IEEE Transactions on Knowledge and Data Engineering, 33(8), 3048-3061.http://hdl.handle.net/10251/184751

Ngairangbam, V. S., Spannowsky, M., & Takeuchi, M. (2022). Anomaly detection in high-energy physics using a quantum autoencoder. *Physical Review D*, *105*(9), 095004.https://journals.aps.org/prd/pdf/10.1103/PhysRevD.105.095004

Ni, F., Zhang, J., & Noori, M. N. (2020). Deep learning for data anomaly detection and data compression of a long- span suspension bridge. *Computer‑ Aided Civil and Infrastructure Engineering*, *35*(7), 685-700.https://www.researchgate.net/profile/Futao-Ni/publication/338211240_Deep_learning_for_data_anomaly_detection_and_data_comp ression_of_a_long-span_suspension_bridge/links/5e0d1adca6fdcc28374fe704/Deep-learning-for-data-anomaly-detection-and-data-compression-of-a-long-span-suspension-bridge.pdf

Pejić Bach, M., Krstić, Ž., Seljan, S., & Turulja, L. (2019). Text mining for big data analysis in financial sector: A literature review. *Sustainability*, *11*(5), 1277.https://www.mdpi.com/2071-1050/11/5/1277/pdf

Raschka, S., Patterson, J., & Nolet, C. (2020). Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence. *Information*, *11*(4), 193.https://www.mdpi.com/2078-2489/11/4/193/pdf

Sahani, J., Kumar, P., Debele, S., Spyrou, C., Loupis, M., Aragão, L., ... & Di Sabatino, S. (2019). Hydro-meteorological risk assessment methods and management by nature-based solutions. *Science of The Total Environment*, *696*, 133936.https://doi.org/10.1016/j.scitotenv.2019.133936

Shahid, N., Rappon, T., & Berta, W. (2019). Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PloS one*, *14*(2), e0212356.https://doi.org/10.1371/journal.pone.0212356

Statista (2023). Spending on digital transformation technologies and services worldwide from 2017 to 2026. Statista. Retrieved on: 16 October 2023 from:https://www.statista.com/statistics/870924/worldwide-digital-transformation-market-size/

Sujath, R. A. A., Chatterjee, J. M., & Hassanien, A. E. (2020). A machine learning forecasting model for COVID-19 pandemic in India. *Stochastic Environmental Research and Risk Assessment*, *34*, 959-972.https://link.springer.com/article/10.1007/s00477-020-01827-8

Teo, T. (2022). What is a white epistemology in psychological science? A critical race-theoretical analysis. *Frontiers in psychology*, *13*, 861584.https://doi.org/10.3389/fpsyg.2022.861584

Tingfei, H., Guangquan, C., & Kuihua, H. (2020). Using variational auto encoding in credit card fraud detection. *IEEE Access*, *8*, 149841-149853.https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9163376

Yao, R., Liu, C., Zhang, L., & Peng, P. (2019, June). Unsupervised anomaly detection using variational auto-encoder based feature extraction. In *2019 IEEE International Conference on Prognostics and Health Management (ICPHM)* (pp. 1-7). IEEE.https://www.researchgate.net/profile/Chongdang-Liu/publication/335495378_Unsupervised_Anomaly_Detection_Using_Variational_Auto-Encoder_based_Feature_Extraction/links/5d6fa8744585151ee49b92f3/Unsupervised-Anomaly-Detection-Using-Variational-Auto-Encoder-based-Feature-Extraction.pdf

Zhang, S., Xiao, K., Carranza, E. J. M., Yang, F., & Zhao, Z. (2019). Integration of auto-encoder network with density-based spatial clustering for geochemical anomaly detection for mineral exploration. *Computers & Geosciences*, *130*, 43-56.https://doi.org/10.1016/j.cageo.2019.05.011