RESEARCH ARTICLE OPEN ACCESS

Big Data Analytics in AI-Driven Decision Making: Techniques and Applications

Khaled N S Alsalmi *, Abdullah M. A. Alzafiri**
**(Industrial Institute -Shuwaikh, Public Authority of Applied Education and Training, Kuwait)*
*** (Industrial Institute -Shuwaikh, Public Authority of Applied Education and Training, Kuwait)*

ABSTRACT
As organizations embark on digital transformations, their reliance on data-driven methods for improving performance, operational efficiency, and competitiveness grows. Organizations investigating Artificial Intelligence (AI) are doing so to capitalize on the possibilities made available by the digital transformations. While AI has been groundbreaking to organizations, AI in conjunction with big data analytics has made it transformative. This paper explores how big data analytics using AI is unlocking a methodology through which organizations can capitalize on data driven decision making as one unit, including the techniques, architectures and applications of AI across industry domains, while explicitly providing future research directions and challenges for organizations to remain mindful of when linking big data platforms with AI tools to optimize decision making. The paper as a whole provides a high level overview of the aspects of AI and big data and their interdependence in a decision making environment.

Keywords-Big Data, Artificial Intelligence, Decision Making, Data Analytics, Machine Learning, Predictive Analytics, Data Science, Real-Time Processing.

Date of Submission: 10-04-2025 Date of acceptance: 23-04-2025

I. INTRODUCTION

The convergence of Big Data Analytics (BDA) combined with Artificial Intelligence (AI) is going to change the way decisions are made from healthcare to finance, to transportation to e-commerce. Traditional decision-making models have often relied on human expertise within limited data sets, often taking into account limited and static rule based systems. Today, data from sensors, web applications, social media, and IoT devices are exceeding the capacity of conventional systems. Organizations now have AI driven decision making that are driven by big data platforms [1].

AI systems are moving towards data patterns and predictions to support complex business logic where once in many cases, an organization would have relied on consulting with human expertise in a matter of decision making; advancing both the capacity of human resources and automated systems to act on behalf of the organization. AI, in essence, is a specific type of computer program that has the potential of learning from large datasets, recognizing patterns from data or predicting outcomes based on data, and applying to a business logic situation when required, in real time. Through the use of big data analytics, a systems structure or framework, organizations will be able to develop and learn from available structured and unstructured data. The integration of these technologies allows organizations to create greater agility, enhance customer experiences, improve operations, and better anticipate the future trends [2].

This paper explains the tools and techniques used in big data analytics that encourage AI-influenced decision-making activities. A survey of use cases across different industries, a review of the architectural frameworks involved, and the challenges and future directions of big data analytics will be discussed. The font size for **heading is 11 points bold face** and **subsections with 10 points and not bold.** Do not underline any of the headings, or add dashes, colons, etc.

II. BIG DATA ANALYTICS: AN OVERVIEW

Big Data refers to data that is too large, too fast, or too complex for traditional data-processing applications to deal with. It is often described by the "5 Vs" - Volume, Velocity, Variety, Veracity, and Value [3].

Big Data Analytics (BDA) describes the procedure of using advanced analytic techniques against very large, diverse data sets to uncover hidden patterns, unknown correlations and other useful information [4]. Hadoop, Spark, Flink, as well as NoSQL databases (e.g., Cassandra, MongoDB) provide the platform of BDA technology.

Common types of big data analytics include:

- **Descriptive Analytics:** summarize past data to understand trends.
- **Diagnostic Analytics:** explain why something occurred.
- **Predictive Analytics:** apply a statistical model or machine learning (ML) to predict future outcomes. Prescriptive Analytics: follow suggested action via predictive models.

BDA platforms can be complementary with AI algorithms in the development of decision support systems, particularly with supervised, and unsupervised ML, creating high-impact decisions.

III. AI-DRIVEN DECISION MAKING

AI-driven decision making is the process of using intelligent systems that can interpret data, learn patterns, and make or recommend decisions without the need for direct human engagement. In this regard, delegated machine learning (ML), deep learning (DL), and natural language processing (NLP) are AI techniques typically being considered.

For example, **ML algorithms** (e.g. decision trees, random forests, SVMs) may be used to detect fraud, predict customer action, and evaluate disparity in risk. **Deep learning** using neural networks established major advances in image recognition, speech processing, and autonomous systems. Similarly, **Reinforcement learning** is increasingly applied in dynamic environments—which can be include fields like robotics or financial trading—where the systems learn based on trial and error or develop their optimal action [5].

These methods of AI will depend on the use of enormous, high-quality datasets to train the models, which is provided by the abundance of data supplied by the big data systems. This enables decision making to become more data-driven, automated, and scalable.

IV. MAIN TECHNIQUES IN BIG DATA ANALYTICS FOR AI

The following techniques from big data analytics are worth examining as they blanket the general purpose for the capability of AI models and applications:

A. Data Preprocessing and Feature Engineering

The concept of data pre-processing is a foundational step in big data pipelines, covering cleaning the dataset, normalizing statistics, transforming variables, and extracting features. Given that datasets will be large and heterogeneous (including images, text, tabular), a best practice will be to harness the use of AI to automate feature selection through the use of AutoML or feature importance algorithms (e.g. SHAP, LIME) [6].

B. Stream Processing

Real-time data is necessary for applications like fraud detection, or autonomous vehicles. Stream processing frameworks like Apache Kafka, Apache Storm, and Spark Streaming offer low-latency streaming analytics in order for AI models to operate with the most updated data [7].

C. Distributed Machine Learning

Training machine learning models on big datasets often involves the use of distributed systems. Both TensorFlow for Kubernetes or Pytorch with Horovod allows the training for networks utilizing a GPU cluster. While distributed systems facilitate scale, it also implements speed via either data or model parallelism [8].

D. Integration with Data lakes and Warehouses

Big data platforms such as Amazon Redshift, Google BigQuery, and Azure Synapse Analytics now include built-in ML capabilities. The added functionality enables organizations to develop AI models directly in their data warehouses, and minimize the complexity associated with ETL [9].

V. APPLICATIONS ACROSS INDUSTRIES

A. Healthcare

The use of big data analytics and AI to build predictive models for predicting disease, drug discovery, and patient assessments. Implementation is through systems such as IBM Watson Health which can leverage massive medical data, to provide recommendations for diagnosis, treatment options, and patient safety [10].

B. Finance

In the banking and fintech sectors AI built on big data is deployed to provide credit scoring, fraud detection, algorithmic trading, and personalized financial planning. For example, JPMorgan Chase deploys AI to evaluate market trends and has invested in AI technology to support investing decisions [11].

C. Retail and E-Commerce

Retailers are also using predictive analytics to optimize supply chain efficiency, forecast demand and develop personalized marketing to promote their offerings. AI tools such as recommendation engines (for use by Amazon and Netflix) leverage user behavior data through deep learning [12].

D. Transportation

Organizations such as Uber and Tesla are also using AI with mobility big datasets for route optimization, demand forecasting, and autonomous vehicles. Real-time traffic and GPS data feeds directly into the AI algorithms [13].

E. Government and Smart Cities

Public health monitoring, smart traffic control, and efficient resource allocation uses government AI big data staffs. AI analytics are used by cities such as Singapore, to manage services and predict possible infrastructure failures in urban areas [14].

VI. CHALLENGES AND LIMITATIONS

Although the implications for AI-driven big data analytics can be profound, there are significant challenges:

- **Data Quality and Bias:**

Defective or biased data will produce flawed predictions from AI methods, especially relevant to sensitive issues such as health systems or law enforcement [15].

- **Scalability Issues:**

Real-time at scale is technically complex, involving significant infrastructure and human resources.

- **Data Privacy and Security:**

With enormous personal data sets, concerns about consent, misuse of information and complying with regulations (e.g., GDPR) come to the fore [16].

- **Explainability:**

Many AI systems, and particularly deep learning models, are also "black boxes", and subsequently it can be difficult to explain to users or regulators why a decision is made [17].

VII. FUTURE DIRECTIONS

In addressing current limitations and gaining future opportunities, AI-driven big data analytics may proceed along these future research opportunities:

- **Explainable AI (XAI):**

Producing transparency and trust in how decisions are made.

- **Federated Learning:**

Training models without moving raw data across distributed data sources. **Quantum Computing integration:**
Using quantum algorithms for big data analysis and AI training.

- **Ethical Frameworks for AI:**

Ensuring fairness, accountability and privacy in AI decision system.

- **Edge AI / Real-time processing:**

Getting analytics and intelligence closer to the source of data for better and faster decisions.

A new emerging direction is the adoption of **multimodal AI**, which uses a myriad of different data types (text, images, sensors etc.) for more holistic and accurate decision making. For example, in the context of autonomous vehicles and smart cities, including video streams, location data (localization data) and the weather enhances the predictive capabilities, and thereby decision quality. Also, as **generative AI** tools such as GPT models with enterprise data become ubiquitous, with their summarizing capabilities and decision supporting capability - this will create a more comprehensive element for the AI-BDA ecosystem. Organizations are also increasingly examining the **data-centric AI** practices, where the focus is data quality over the complexity of models, to provide better generalizability and robustness. Each of these trends is evolving for a more intelligent, flexible and ethical space for AI powered decision making.

VIII. CONCLUSION

The emergence of big data analytics and AI is redefining how organizations make decisions. Across a range of popular ideas of transition from health systems, to transport; scalable data platforms combined with intelligent algorithms are facilitating informed, efficient and autonomous decision making in systems. Although technical and ethical considerations remain, the rapid evolution of tools with big data and models with AI augurs well for a future where data will take on an even greater role in informing and even directing decisions in every domain.

Ultimately, pursuing this direction will continue to depend not only on new technology development, but the interplay of that technology and the strategy of AI with data, which is distinct within the goals of businesses, regulatory obligations and societal values. Organizations able to responsibly wield this synergy will ultimately sustain a significant competitive advantage, and significantly contribute to building more intelligent, efficient, and ethical structures. As we move to a world where **data is the new oil, and AI is the engine** to take advantage of that oil, the need for sustainable, explainable and scalable data science frameworks will be paramount.

REFERENCES

[1]. Marr, B. (2018). Data-Driven Decision Making. Wiley.
[2]. Gandomi, A., & Haider, M. (2015). "Beyond the hype: Big data concepts, methods, and analytics." International Journal of Information Management, 35(2), 137–144.
[3]. Laney, D. (2001). "3D Data Management: Controlling Data Volume, Velocity and Variety." Meta Group.
[4]. Chen, M., Mao, S., & Liu, Y. (2014). "Big Data: A survey." Mobile Networks and Applications, 19(2), 171–209.
[5]. Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction. MIT Press.
[6]. Kuhn, M., & Johnson, K. (2013). Applied Predictive Modeling. Springer.
[7]. Zaharia, M., et al. (2013). "Discretized streams: Fault-tolerant streaming computation at scale." SOSP.
[8]. Sergeev, A., & Balso, M. D. (2018). "Horovod: fast and easy distributed deep learning in TensorFlow." arXiv:1802.05799.
[9]. Google Cloud. (2023). BigQuery ML. https://cloud.google.com/bigquery-ml
[10]. Jiang, F. et al. (2017). "Artificial intelligence in healthcare: Past, present and future." Stroke and Vascular Neurology, 2(4), 230–243.
[11]. JPMorgan Chase. (2022). How AI is Transforming Finance. https://www.jpmorganchase.com
[12]. McKinsey & Company. (2021). The state of AI in retail. https://www.mckinsey.com
[13]. Tesla. (2023). AI Day Presentation. https://www.tesla.com/AI
[14]. Smart Nation Singapore. (2023). https://www.smartnation.gov.sg
[15]. Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and Machine Learning. fairmlbook.org
[16]. Voigt, P., & Von dem Bussche, A. (2017). The EU General Data Protection Regulation (GDPR). Springer.
[17]. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Proceedings of the 22nd ACM SIGKDD, 1135–1144.