

Intelligent Data Processing: The Impact of AI and Machine Learning on Big Data

Aaditi Manoj Dhumal¹

¹MSc Computer Science, Mamasheeb Mohol College, Paud Road, Pune

Publication Date: 2026/05/22

Abstract: Big Data is a lot of information that comes from different places like social media, detectors and computers. This information is made snappily. Artificial Intelligence and Machine Learning are tools that help us understand and get useful information from Big Data. This paper addresses about how Artificial Intelligence and Machine Learning make it better to reuse, store and dissect Big Data. It does this by using robotization, prophetic analytics and smart decision-timber. We look at how Artificial Intelligence and Machine Learning are used in Big Data operations in numerous fields, like healthcare, finance and education. We also talk about the problems that come with Big Data like making sure the information is good keeping it private and making sure it can be used by a lot of people.

Keywords: Big Data, Artificial Intelligence Machine Learning, Data Analytics, Predictive Modeling.

How to Cite: Aaditi Manoj Dhumal (2026) Intelligent Data Processing: The Impact of AI and Machine Learning on Big Data. *International Journal of Innovative Science and Research Technology*, 11(4), 4991-4994.
<https://doi.org/10.38124/ijisrt/26apr2272>

I. INTRODUCTION

In today's world data is very important for people, companies and societies. Every second a lot of data is being made through media, online buying, mobile apps, sensors and digital devices. This huge amount of data is called Big Data. Big Data has a lot of information comes in fast is very different and is valuable. Old ways of managing and analyzing data have trouble handling such large and varied information. To deal with these problems new technologies like Artificial Intelligence and Machine Learning have become very useful. Artificial Intelligence makes machines think like humans so they can do things like figure out problems and make decisions. Machine Learning is a part of Artificial Intelligence that helps machines learn from data on their own find patterns and get better without being programmed. When Artificial Intelligence and Machine Learning are used with Big Data analysis they make it much easier to understand and use data. The role of Artificial Intelligence and Machine Learning in Big Data is that they can handle amounts of information, automate tasks that use data and give useful ideas in real time. Machine Learning can look at data to make good predictions while Artificial Intelligence models can change quickly when new information comes in. This helps make data analysis more accurate and efficient. It also lets companies make good decisions faster. Artificial Intelligence and Machine Learning are very important, for Big Data because they help companies use Big Data to their advantage. Big Data and Artificial Intelligence and Machine Learning work together to make things better.

II. LITERATURE REVIEW

The combination of Artificial Intelligence and Machine Learning with Big Data has become very popular in the few years. People have been working on using Artificial Intelligence and Machine Learning to get information from large sets of data predict what might happen and make better decisions. Many ways to do this have been found to make Big Data analysis better and more accurate.

➤ Predictive Analytics

Predictive analytics uses Artificial Intelligence and Machine Learning to look at data and guess what will happen in the future. Some common ways to do this are:

- Regression models: These predict numbers like how many things will be sold.
- Classification models: These put things into groups like finding transactions.
- Time-series forecasting: This predicts what will happen next based on what happened.

Some studies, like the one by Chen and others in 2022 have shown that using Machine Learning for models makes forecasts much better in areas like healthcare, money and marketing. Predictive analytics helps companies fix problems and make plans before things happen.

➤ Data Processing and Pattern Recognition

Artificial Intelligence and Machine Learning help process large sets of data finding patterns that are not easy to see. Some common ways to do this are:

- Clustering: This groups things together like customers or unusual events.
- Association analysis: This finds connections between things like what people buy
- Dimensionality reduction: This makes data simpler without losing information.

For example: Kaur and Kaur found in 2021 that using algorithms helps businesses understand what customers do and make better marketing plans.

➤ *Deep Learning*

Deep Learning is a part of Machine Learning that uses layers to handle complicated and unstructured data. It is used for:

- Image recognition: Looking at pictures to diagnose sicknesses.
- Natural Language Processing: Dealing with amounts of text to understand what people think and get information.
- Speech recognition: Turning audio into information.

Research by Li and others in 2023 shows that Deep Learning models are better than Machine Learning models at handling complicated Big Data tasks because they can automatically find important features in raw data.

➤ *Challenges*

Even though Artificial Intelligence and Machine Learning have improved a lot there are still some problems:

- Data quality and preprocessing: Bad or incomplete data can make models less accurate.
- Scalability issues: Handling more Big Data requires Artificial Intelligence and Machine Learning models that can do many things at once.
- Interpretability: Complicated Artificial Intelligence models, Deep Learning can be hard to understand because we do not know how they make decisions.
- Privacy and ethical concerns: Keeping data safe and making sure algorithms are fair is very important.

Overall the literature says that while Artificial Intelligence and Machine Learning make Big Data analysis much better we need to be careful with data choose the models and think about ethics to get useful and reliable information, from Artificial Intelligence and Machine Learning and Big Data.

➤ *Objectives*

The main objectives of this study are to:

- Study how Artificial Intelligence and Machine Learning improve Big Data analysis. We want to see how AI and Machine Learning make Big Data analytics better.
- Look at AI and ML uses in Big Data in areas. This includes checking out how different sectors use these technologies.

- Find out the problems with using AI and Machine Learning in Big Data systems. We need to understand the challenges of implementing these technologies.
- Examine trends and chances, for combining AI Machine Learning and Big Data.

III. METHODOLOGY

This study uses a mix of experiments and comparisons to see how Artificial Intelligence and Machine Learning work with Big Data analysis. The study looks at the role of Artificial Intelligence and Machine Learning, in Big Data analytics to understand it better. The goal is to find out how Artificial Intelligence and Machine Learning are used in Big Data analytics.

➤ *Tools and Frameworks*

- I use the Python programming language.
- My libraries include Scikit-learn and TensorFlow and PyTorch and Pandas and NumPy.
- I work with Apache Spark and Hadoop platforms.
- I build types of models including Supervised and unsupervised Machine Learning models and Deep Learning models like CNN and RNN and LSTM and also Reinforcement Learning models.

➤ *Techniques*

- *Predictive Analytics:*
 - ✓ I train regression and classification models on big data.
 - ✓ Then I compare different Machine Learning algorithms like Random Forest and XGBoost and SVM to find the approach.
- *Data Processing and Pattern Recognition:*
 - ✓ I use clustering techniques like K-Means and DBSCAN to segment datasets.
 - ✓ I also use association rules to discover relationships among variables in the data.
 - ✓ I perform dimensionality reduction using PCA and t-SNE to improve efficiency and make the data easier to visualize.
- *Deep Learning:*
 - ✓ I develop networks for high-dimensional and unstructured data like images and text.
 - ✓ I apply CNNs to image-based data and RNNs and LSTMs to sequential data like time series analysis.
 - ✓ I utilize Natural Language Processing techniques for textual data analysis in document repositories.
- *Optimization:*
 - ✓ I implement models on distributed frameworks like Spark and Hadoop to handle datasets efficiently.

- ✓ I optimize hyper parameters using grid search or Bayesian optimization to improve the performance of the models.

➤ *Evaluation*

- I use Automated Metrics like Accuracy and Precision and Recall and F1-score and RMSE and MAE depending on the task.
- I also evaluate Computational Performance by checking the processing time and memory utilization and scalability analysis.
- Then I do a Comparative Analysis to compare Machine Learning models with Deep Learning models for effectiveness, in handling big data.
- Finally, I do Real-World Validation by testing the models on datasets from finance and healthcare and marketing domains to assess how well they work in life.

IV. EXPECTED RESULTS

➤ *From what we Know far:*

- Predictive Analytics should make it more accurate to forecast trends and make decisions in areas like healthcare, finance and marketing.
- Deep Learning models are likely to work well with complex data like images, text and sensor information.
- Techniques like Clustering and pattern recognition will help find patterns, groups of customers and unusual data points in large datasets.
- Using a mix of approaches like combining Deep Learning and Predictive Analytics might give us the insights and results that we can actually use.
- Distributed frameworks that can handle amounts of data like Spark and Hadoop are expected to process huge datasets quickly without slowing down.

These expected outcomes show that using AI and ML techniques with Big Data will improve accuracy, efficiency and useful insights, for real-world applications.

V. APPLICATIONS

➤ *Healthcare*

- The Healthcare industry uses Artificial Intelligence and Machine Learning models to look at data. This helps doctors make a diagnosis and plan the right treatment.
- We use Predictive Analytics to find out if someone might get a disease before it happens. This also helps us see if a patient is at risk.
- For example: a deep learning model can look at images to find tumors. It does this with a level of accuracy.

➤ *Finance*

- In the Finance Sector Artificial Intelligence helps find transactions and manage financial risks.
- Big Data Analytics assists in predicting what the market will do and planning investments.

- For instance, a Machine Learning model can predict how stock prices will move based on what happened in the past and what people are saying about it.

➤ *Business Intelligence*

- Artificial Intelligence tools look at amounts of data to find patterns and trends.
- They also help us understand our customers better.
- Automated dashboards give us real-time analytics to help us make decisions.
- For example: An Artificial Intelligence-powered engine can look at what people buy and suggest products that they might like.

➤ *Education*

- Adaptive learning platforms change what they teach based on how well students doing.
- Predictive models find students who might be struggling and suggest ways to help them.
- For example: An Artificial Intelligence system can make a study plan, for a student using learning analytics.

➤ *Manufacturing and Industry*

- Artificial Intelligence and Machine Learning make production processes and supply chain management better.
- They also help with maintenance that can predict when something might go wrong.
- Big Data Analytics reduces the time when machines are not working and makes everything run smoothly.
- For instance, we can predict when a machine might break down by looking at sensor data. This helps us schedule maintenance at the time.

VI. CHALLENGES AND FUTURE SCOPE

➤ *Challenges*

- Data Quality and Preprocessing: We have to deal with noisy data that can make Artificial Intelligence and Machine Learning models not work so well.
- Scalability: We need to handle a lot of data that is growing really fast so we need strong and efficient computers to do that.
- Interpretability: Some models, like learning networks are really hard to understand because we do not know how they make predictions.
- Privacy and Security: We have to keep data secret and make sure it is not used in a way, which is a big concern.
- Cross-Domain Limitations: Artificial Intelligence models that are trained on data from one area may not work well in areas unless we make some changes.

➤ *Future Scope*

- Explainable Artificial Intelligence: We want to make models that can explain what they do so we can trust them. Make good decisions.

- Federated Computing: We can process data where it is made and use a team approach to learning, which helps keep data private and makes things faster.
- Integration with Internet of Things and Cloud Computing: We can combine Artificial Intelligence, Machine Learning and Big Data with Internet of Things devices and cloud computers to get real-time information.
- Ethical Artificial Intelligence: We need to make sure Artificial Intelligence is fair and used in a way without any biases.
- Hybrid Approaches: We can use many different Artificial Intelligence and Machine Learning techniques together like deep learning and predictive analytics to get the best results and be more efficient.

This shows that while Artificial Intelligence and Machine Learning are really helpful for Big Data analytics we need to solve these problems and use technologies to make them even better and more useful in the future, for Artificial Intelligence and Machine Learning.

VII. CONCLUSION

The combination of Artificial Intelligence and Machine Learning with Big Data has changed the way companies work with complicated sets of information. Artificial Intelligence and Machine Learning methods, such as predicting what will happen learning from data and finding patterns make it possible to forecast more accurately make decisions more efficiently and automate tasks in many areas like healthcare, finance, education and industry.

Artificial Intelligence and Machine Learning still have problems such as data not being able to handle large amounts of information being hard to understand and ethical issues. However new developments in Artificial Intelligence that can be understood, computing that can be done in places and mixed modeling methods promise to make these technologies more reliable and useful.

Overall, Artificial Intelligence, Machine Learning and Big Data working can greatly help companies get useful information work more efficiently and come up with new ideas. Future studies that focus on Artificial Intelligence, Machine Learning and Big Data being fair able to handle amounts of information and easy to understand will make these technologies have an even bigger impact, on many industries.

REFERENCES

- [1]. Chen, M., Mao, S. And Liu Y. Wrote a survey on Big Data in 2014 for Mobile Networks and Applications.
- [2]. Dean, J. And Ghemawat S. Published a paper on MapReduce in 2008. They talked about how to process amounts of data easily. It was in Communications of the ACM.
- [3]. Davenport T. Wrote an article for Harvard Business Review in 2018. He discussed using Artificial Intelligence in the world.

- [4]. Jordan, M. And Mitchell T. Wrote about Machine Learning in 2015. They shared trends, perspectives and prospects in Science.
- [5]. Manyika, J. And others published a report on Big Data in 2011. They talked about how Big Data can help with innovation, competition and productivity. It was, from McKinsey Global Institute.