

The Role of Artificial Intelligence in Analyzing Large Datasets and Its Utilization in Commercial Decision-Making

Information Technology
Department of Information Technology, Faculty of Technology
Bachelor's Thesis

Author:
Riku Lauttia

January 2025

According to the quality system of the University of Turku, the originality of this publication has been checked using the Turnitin OriginalityCheck system.

UNIVERSITY OF TURKU

Department of Information Technology, Faculty of Technology

RIKU LAUTTIA: The Role of Artificial Intelligence in Analyzing Large Datasets and Its Utilization in Commercial Decision-Making

26 pages

January 2025

Artificial intelligence (AI) has become a central tool in the analysis of large datasets, enabling large enterprises to refine vast amounts of information into valuable business insights.

Efficient processing and analysis of large data volumes have become essential for maintaining a competitive edge, and AI facilitates the understanding and utilization of even complex data structures. This helps businesses identify hidden trends, predict market developments, and make data-driven decisions more quickly and accurately.

This thesis is based on a literature review analyzing the application of AI in large-scale data analysis and its impact on commercial decision-making in large enterprises. The analysis focuses particularly on recent studies that explore AI's role in data preprocessing, validation, and optimization. Key findings indicate that AI can significantly enhance decision-making in large enterprises by streamlining data analysis processes and providing deeper insights into business environments. As a result, AI also helps reduce human errors and improves the efficiency of resource utilization.

Keywords: artificial intelligence, machine learning, big data analysis, commercial decision-making, large enterprises, data analytics, technology applications

Table of Contents

1 INTRODUCTION	1
2 ARTIFICIAL INTELLIGENCE IN LARGE-SCALE DATA ANALYSIS	3
2.1 Artificial Intelligence and Data Processing	4
2.1.1 Big Data Challenges and the Role of Artificial Intelligence	4
2.1.2 Stages and Importance of Data Preprocessing	6
2.1.3 Cloud Services and Scalable Data Processing	6
2.2 Utilizing Machine Learning in Data Analysis	7
2.2.1 Fundamentals of Machine Learning and Learning Methods	7
2.2.2 Algorithms and Modeling for Large Datasets	8
2.2.3 The Role of Deep Learning in Analysis	9
2.2.4 Machine Learning Tools	9
2.2.5 Applications of Machine Learning	10
2.2.6 Practical Examples	11
2.3 Summary	12
3.1 Practical Applications in Decision-Making	13
3.1.1 Market Trend Analysis and Forecasting	13
3.1.2 Customer Behavior Analysis and Personalization	14
3.1.3 Resource Optimization and Management	15
3.1.4 Risk Management and Fraud Detection	16
3.1.5 Supporting Strategic Decision-Making	17
3.2 Commercial Benefits and Challenges	19
3.2.1 Efficiency Improvements Brought by AI	19
3.2.2 Gaining a Competitive Advantage with AI	19
3.2.3 Technological Challenges and Legacy Systems	20
3.2.4 Infrastructure Requirements and Environmental Impact	21
3.2.5 Challenges and Strategies for AI Adoption	22
3.2.6 Ethics and Data Privacy	23
3.3 Summary	23
4 SUMMARY	25

1 Introduction

The rapid advancement of artificial intelligence (AI) has revolutionized the way large enterprises process and analyze vast amounts of data. Efficient analysis of large datasets has become essential for maintaining a competitive advantage in the global business environment (Kirola et al., 2023; Dandekar et al., 2024). Traditional analytical methods often fail to manage the increasing volume, variety, and velocity of data, emphasizing the significance of AI and machine learning in modern data analytics (Dandekar et al., 2024).

AI applications, such as predictive analytics, natural language processing, and neural networks, enable a deeper understanding of data and the identification of hidden patterns (Jamalpur et al., 2024). For instance, machine learning models allow businesses to analyze market trends, predict customer behavior, and optimize resource utilization, leading to more effective strategic decision-making (Johri et al., 2023; Sun et al., 2022). Efficient use of these technologies can significantly enhance operational efficiency and improve the quality of decision-making within organizations.

The objective of this study is to explore how various AI technologies can be effectively applied to analyze large datasets and examine their impact on commercial decision-making in large enterprises. The study focuses particularly on practical applications of AI and machine learning in data preprocessing, analysis, and optimization, while also addressing the challenges and solutions associated with implementing these technologies in large organizations.

This study aims to answer the following research questions:

- RQ1: How can AI and machine learning technologies be effectively applied to analyze large datasets?
- RQ2: How can AI and machine learning be utilized to support commercial decision-making in large enterprises?

This research focuses on large enterprises operating in data-driven industries and facing challenges related to processing vast amounts of data. The study does not consider the perspective of small and medium-sized enterprises (SMEs); instead, the emphasis is on the

needs and opportunities of large organizations. AI and machine learning applications are examined specifically in the context of business optimization, including strategic planning, market analysis, and customer behavior prediction (Jamalpur et al., 2024; Johri et al., 2023).

This thesis is a literature review analyzing the utilization of AI and machine learning in processing large datasets and their application in commercial decision-making within large enterprises. The research material has been collected from scientific sources such as IEEE Xplore, ACM Digital Library, Google Scholar, and ScienceDirect. The search process included queries such as: "(AI OR artificial intelligence) AND ("machine learning" OR ML) AND (big data) AND (commercial decision-making)". The analysis focuses on studies published between 2020 and 2024, which examine the latest AI applications in business environments.

The thesis consists of four main chapters. After the introduction, Chapter 2 discusses the fundamentals of AI and machine learning and their role in analyzing large datasets. Chapter 3 explores practical applications of AI and machine learning in commercial decision-making, along with their benefits and challenges. Finally, Chapter 4 presents a summary of the key findings from the study.

2 Artificial Intelligence in Large-Scale Data Analysis

Artificial Intelligence (AI) and Machine Learning (ML) have emerged as key technologies in recent years for processing and analyzing large datasets. Big data refers to vast, complex, and often unstructured collections of information that traditional data processing methods struggle to handle efficiently (Kirola et al., 2023). AI and ML provide solutions to these challenges by enabling automated data processing, analysis, and utilization in business operations. The importance of analyzing large datasets continues to grow as organizations increasingly encounter data that is distributed across various formats and sources. As the volume, velocity, and variety of data increase, traditional methods become insufficient for meeting the demands of modern data environments. This highlights the crucial role of AI and ML in data analysis, as these technologies offer scalability and automation in data processing (Kirola et al., 2023).

AI refers to systems that can perform tasks typically requiring human intelligence, such as pattern recognition, predictive modeling, and decision-making. AI enhances the efficiency and intelligence of data utilization in business applications, including market analysis and customer behavior prediction (Jamalpur et al., 2024). In this thesis, AI refers to broader technological applications that include AI-driven solutions and systems that integrate machine learning models as part of their functionality.

Machine Learning (ML), a subset of artificial intelligence, refers to algorithms that can learn and improve their performance without explicit programming. ML enables computers to analyze vast amounts of data, make predictions, and identify hidden patterns. Machine learning models learn from historical data and generate insights that can help businesses make more effective strategic decisions (Kirola et al., 2023). In this thesis, machine learning is discussed with a focus on data analysis, modeling, and prediction techniques, which represent the practical application of AI in data processing.

Big data refers to data sets that are too large and complex for traditional database systems to handle efficiently. Managing and analyzing big data requires advanced technologies and methodologies, such as AI and ML, to extract meaningful insights for business needs (Dandekar et al., 2024).

The purpose of this chapter is to explore how artificial intelligence and machine learning can be utilized in processing and analyzing large datasets. The first subsection (2.1) examines the role of AI in data processing, while the second subsection (2.2) focuses on the solutions that machine learning offers for data analysis. These technologies play a key role in enabling large enterprises to strategically leverage insights derived from big data to support commercial decision-making.

2.1 Artificial Intelligence and Data Processing

The significance of artificial intelligence in data processing has grown substantially, particularly in the analysis of large datasets. AI enables the automation of complex processes related to data collection, processing, and analysis. Traditional data processing methods struggle to handle large and diverse datasets, which are often generated by sources such as the internet and IoT devices. AI applications, including machine learning, neural networks, and natural language processing (NLP), facilitate a deeper analysis and more effective utilization of data in business operations (Pani et al., 2021). Pani et al. (2021) highlight that AI-powered applications, particularly cloud-based solutions, are highly effective in big data analytics. Cloud services provide a scalable and cost-efficient approach to managing and analyzing large datasets, which are often distributed across multiple sources. Additionally, AI-driven technologies such as deep learning enable the real-time analysis of vast data volumes, enhancing both the accuracy and speed of decision-making in business environments.

2.1.1 Big Data Challenges and the Role of Artificial Intelligence

Managing large datasets, commonly referred to as big data, is a complex task that requires specialized technologies and methodologies. Big data is often characterized by the "5 Vs": Volume, Velocity, Variety, Veracity, and Value. These concepts define the key attributes and challenges organizations face when leveraging data for business operations (Kirola et al., 2023).

- **Volume:** The sheer scale of big data is its most defining feature. IoT devices, social media, and other digital interactions continuously generate vast amounts of data. Traditional databases and analytical methods are incapable of efficiently handling

such volumes, emphasizing the need for AI-powered technologies in data management.

- **Velocity:** Data is not only generated in massive quantities but also at high speeds. Real-time data processing poses a challenge for traditional systems. AI, however, can process data rapidly and generate insights almost instantaneously, enabling companies to make quick, data-driven decisions (Jamalpur et al., 2024).
- **Variety:** Big data is heterogeneous, consisting of structured data (e.g., database tables) and unstructured data (e.g., text, images, and videos). This diversity complicates data preprocessing and analysis, but AI effectively handles complex data structures using various advanced algorithms.
- **Veracity:** The quality of big data can be inconsistent, with uncertainties and errors affecting its reliability. Poor data quality can undermine analytical accuracy. AI enhances data accuracy and reliability by identifying anomalies and errors during preprocessing (Dandekar et al., 2024).
- **Value:** Data in its raw form has limited business significance unless it can be analyzed and transformed into actionable insights. AI plays a crucial role in unlocking the value of data, enabling the detection of hidden trends and patterns that lead to more informed business decisions (Dandekar et al., 2024; Kirola et al., 2023).

By addressing these big data challenges, AI significantly improves the efficiency and effectiveness of data-driven decision-making in large enterprises.

The enormous volume, high velocity, and diverse nature of big data make its analysis particularly challenging. Traditional analytical methods are often too slow and limited to effectively handle these characteristics, making AI-driven solutions essential. Dandekar et al. (2024) emphasize that the primary challenges in big data analytics relate to data quality, privacy, and security. Future advancements in AI and machine learning are expected to focus on overcoming these challenges by making data processing more efficient and secure. One of the most significant challenges is the fragmentation of data across multiple sources, which complicates its integration and analysis. Additionally, data quality can vary significantly, necessitating robust preprocessing methods to ensure reliable outcomes. AI can automate complex processes such as data cleaning and transformation, making it suitable for analysis and ultimately leading to faster and more accurate insights.

2.1.2 Stages and Importance of Data Preprocessing

Data preprocessing is an essential step before large datasets can be analyzed using artificial intelligence and machine learning. Without preprocessing, the quality of raw data can degrade the accuracy of the analysis and lead to incorrect conclusions. The purpose of the preprocessing process is to improve data quality by making it cleaner, more consistent, and easier to handle. This stage includes several procedures, such as data cleaning, normalization, and transformation into an analyzable format (Kirola et al., 2023). Data cleaning involves removing or correcting incorrect or missing information. This is particularly important because incomplete or inconsistent data can distort the results of machine learning models. Normalization is another key aspect of preprocessing, where different measurement units are converted to the same scale so that algorithms can process the data consistently. Preprocessing also includes data transformation, such as converting categorical data into numerical form, which facilitates analysis with machine learning models.

2.1.3 Cloud Services and Scalable Data Processing

Processing large datasets requires efficient and scalable systems. Distributed systems and parallel processing are key technologies that enable data to be processed at high volumes and speeds. Distributed databases, such as Apache Hadoop and Apache Spark, are examples of technologies that utilize parallel processing for large-scale data analysis (Kirola et al., 2023). Hadoop is a distributed computing system that allows large amounts of data to be processed across multiple parallel nodes. Its core architecture is based on the MapReduce programming model, which distributes data processing into multiple segments and efficiently combines the results. Hadoop is particularly useful in situations where data is distributed across multiple sources and systems (Dandekar et al., 2024). Spark, on the other hand, is a distributed computing framework that provides a faster and more flexible way to process data compared to traditional MapReduce. Spark utilizes in-memory processing, which enhances performance and enables real-time data analysis. This makes it especially useful for analyzing large and complex datasets (Dandekar et al., 2024).

Cloud services are another crucial factor in scalable data processing. Cloud platforms provide a flexible and cost-effective way to store, process, and analyze large datasets. One of the biggest advantages of cloud services is their ability to scale according to business needs. This

means companies can dynamically allocate resources for data analysis in real time (Pani et al., 2024). Dandekar et al. (2024) state that cloud-based AI and machine learning systems have revolutionized the way companies analyze large datasets. Cloud platforms such as Amazon Web Services (AWS) and Microsoft Azure offer AI- and ML-based services that seamlessly integrate machine learning models and data processing. Cloud computing allows businesses to reduce infrastructure costs and focus on leveraging data strategically without the burden of extensive IT maintenance.

2.2 Utilizing Machine Learning in Data Analysis

Machine learning is a subfield of artificial intelligence that focuses on algorithms capable of learning and improving their performance without explicit programming. With machine learning, computers can analyze large amounts of data, make predictions, and identify hidden patterns. Machine learning models learn from historical data and generate insights that can help businesses make strategic decisions more efficiently (Jamalpur et al., 2024).

Machine learning can be categorized into three main types: supervised learning, unsupervised learning, and reinforcement learning. Each of these categories is suited for solving different types of problems and is widely used in big data analytics.

2.2.1 Fundamentals of Machine Learning and Learning Methods

Supervised learning is one of the most commonly used machine learning methods. In this approach, the algorithm is trained using data that includes known answers or correct responses from the training data. The algorithm learns to model the relationship between input and output, allowing it to predict new results from previously unseen inputs. Examples of supervised learning include classification and regression algorithms, such as logistic regression and decision trees. Supervised learning is often used for predictive analytics, such as forecasting market trends and analyzing customer behavior (Kirola et al., 2023).

Unsupervised learning differs from supervised learning in that the data is not labeled, meaning the algorithm does not have predefined correct answers. Instead, the algorithm attempts to identify patterns or structures within the data independently. Applications of unsupervised learning include clustering methods, where data is grouped based on similar characteristics, and anomaly detection, which aims to identify outliers in large datasets. Examples of unsupervised learning techniques include k-means clustering and principal

component analysis (Kirola et al., 2023). Reinforcement learning is a learning technique in which an algorithm learns to make decisions based on rewards and penalties. The algorithm interacts with its environment and seeks to maximize cumulative rewards over time.

Reinforcement learning is particularly useful in dynamic environments where decision-making must be adaptive and flexible, such as in robotics and autonomous vehicles (Kirola et al., 2023).

2.2.2 Algorithms and Modeling for Large Datasets

Analyzing large datasets requires specialized, scalable algorithms that can efficiently and accurately process vast amounts of data. Traditional analytical methods are insufficient in these situations, making machine learning-based algorithms essential for extracting valuable insights from data. Machine learning provides a variety of algorithms and methods that can be used to analyze and optimize the utilization of large datasets across different applications.

Algorithms used for processing large datasets must be scalable to handle increasing data volumes and varying data structures. Scalability means that an algorithm can adapt to process ever-growing amounts of data without significant performance degradation. Examples of such algorithms include decision trees, clustering methods, and deep learning.

Decision trees are among the most widely used algorithms for processing large datasets due to their simplicity, interpretability, and efficiency in complex environments. A decision tree algorithm constructs a series of rule-based conditions that classify data into different categories or predict continuous values. This makes decision trees useful for both classification and regression tasks. Decision trees can be optimized for large datasets using methods such as random forests and gradient-boosted models, which enhance performance and predictive accuracy. Clustering methods, such as k-means clustering and hierarchical clustering, are particularly useful when data is not pre-labeled and there is a need to identify natural groupings or clusters within the data. Clustering enables the identification of similar groups within large datasets, which helps with targeted marketing strategies or customer behavior analysis (Johri et al., 2023).

2.2.3 The Role of Deep Learning in Analysis

Deep learning is one of the most advanced and effective methods for analyzing large and complex datasets. It is based on multi-layered neural networks that automatically learn to identify hidden patterns and structures within data. Deep learning is particularly powerful for processing complex and unstructured data, such as images, text, and audio (Pani et al., 2021). Convolutional Neural Networks (CNNs) are specialized in image and visual data analysis. Their ability to recognize visual features and patterns makes them highly useful for tasks involving two-dimensional data, such as image and video analysis.

Recurrent Neural Networks (RNNs), on the other hand, are designed to handle sequential data, such as text data or time-series analysis. RNNs are particularly useful in scenarios where previous states influence current predictions, allowing them to analyze large amounts of data and generate forecasts in complex, time-dependent environments.

Pani et al. (2021) emphasize that the use of neural networks in big data analytics has become increasingly popular because they enable the analysis of massive datasets and the identification of various relationships without requiring predefined models or explicit programming instructions.

2.2.4 Machine Learning Tools

Machine learning provides powerful tools for modeling and prediction in large-scale data analysis. These tools enable businesses to anticipate future events, make data-driven decisions, and optimize business processes. Predictive capabilities are based on learning from historical data and identifying patterns, which allow for forecasting market trends, customer behavior, and resource needs (Dandekar et al., 2024).

In addition to predictive analytics, machine learning helps with anomaly detection, which is critical for risk management and quality control. Anomalies are data points that significantly deviate from the rest of the dataset, and their detection relies on algorithms such as support vector machines (SVMs) and deep learning models. These models are particularly useful for analyzing large and complex datasets, as they can identify hidden issues and enable rapid responses (Johri et al., 2023).

Successful implementation of machine learning requires advanced tools and libraries such as TensorFlow, PyTorch, and Scikit-learn. TensorFlow, developed by Google, is an open-source machine learning framework designed especially for scalable deep learning applications. It provides a flexible environment for building and training deep learning models on large datasets. Its ability to leverage GPU computing significantly enhances processing power, making it an excellent choice for handling large and complex neural networks (Pani et al., 2021). PyTorch, developed by Meta (formerly Facebook), is another widely used deep learning library. It offers an easy-to-use interface specifically designed for flexibility and rapid prototyping. Due to its dynamic computation graph, PyTorch is a popular choice in academic research, though it is also well-suited for industrial applications that require scalable solutions for processing large datasets. Scikit-learn is a widely used library, particularly for classical machine learning algorithms such as linear regression, decision trees, and clustering methods. While Scikit-learn is not designed for deep learning in the same way as TensorFlow or PyTorch, it is an excellent tool for rapid prototyping and analysis of traditional machine learning solutions. Scikit-learn is especially effective when analyzing large but structurally simpler datasets.

2.2.5 Applications of Machine Learning

Machine learning applications can be found across nearly all industries. One widely used application is market forecasting, where machine learning models analyze historical sales data, economic indicators, and other key variables to predict future market trends. This provides businesses with a competitive advantage, enabling them to respond more quickly to market changes and optimize inventory management or supply chains proactively.

Another significant application area is customer behavior prediction. Machine learning models can analyze purchasing habits, online behavior, and demographic data to predict which customers are most likely to buy specific products or services in the future (Johri et al., 2023). Dandekar et al. (2024) emphasize that machine learning plays a crucial role in decision-making. Machine learning models allow organizations to make data-driven decisions based on accurate predictions and analysis. This reduces decision-making risks and helps businesses optimize their operations more efficiently.

Anomaly detection is a third key application area in large-scale data analysis. Anomalies are data points that significantly deviate from the observed patterns in a dataset. Detecting these

anomalies is important because they can indicate problems or unusual events that require attention. For example, detecting anomalies in financial data can help prevent fraud, while in industrial settings, it can help prevent costly equipment failures. Machine learning models such as Support Vector Machines (SVMs) and deep learning models are widely used for anomaly detection in large-scale data analysis. These models can identify complex anomalies and respond quickly to situations requiring immediate attention. Additionally, they can be optimized for handling large datasets in distributed systems, improving their suitability for big data environments.

2.2.6 Practical Examples

Several large companies have successfully utilized machine learning tools and frameworks for analyzing large datasets. For example, Google extensively uses TensorFlow in its products, such as search algorithms and targeted advertising, analyzing vast amounts of user data in real time. This enables more accurate predictions of user behavior and improves business outcomes (Google Research, 2016; Google Open Source, n.d.).

Another example is Netflix, which leverages both TensorFlow and PyTorch to develop deep learning models for improving its recommendation systems. Netflix's algorithms analyze user viewing history and other behavioral data to predict which content users are most likely to watch next. By using these machine learning models, Netflix can enhance customer satisfaction and increase user engagement with the platform (How Netflix Uses Python, 2023).

Amazon also extensively applies machine learning, particularly through its Amazon Web Services (AWS) platform. AWS offers a range of machine learning solutions for businesses looking to analyze large datasets and utilize predictive analytics in decision-making. AWS machine learning services, such as Amazon SageMaker, have helped numerous companies develop machine learning models without the need to maintain their own infrastructure (Amazon Web Services, 2023).

2.3 Summary

Artificial intelligence and machine learning enable the efficient analysis of large datasets. Chapter 2 explored how these technologies provide solutions to the complex challenges of data processing and analysis. The role of AI in data processing is particularly emphasized when analyzing vast amounts of heterogeneous and rapidly changing data. Machine learning introduces scalable algorithms and methods capable of identifying hidden patterns and predicting future trends from large and complex datasets.

This chapter highlighted that machine learning offers tools such as predictive analytics and anomaly detection, which help businesses optimize their decision-making. Supervised and unsupervised learning models enable efficient data classification and forecasting, while deep learning models, such as convolutional and recurrent neural networks, have proven to be highly effective in processing unstructured data like images and text. Machine learning tools and frameworks, including TensorFlow, PyTorch, and Scikit-learn, provide powerful solutions for building and implementing models in large-scale data environments. The use of these tools has led to significant advancements in data analysis.

Based on the content of this chapter, an answer can be provided to Research Question 1 (RQ1): How can AI and machine learning technologies be effectively applied to analyze large datasets? AI and machine learning can be effectively applied to large-scale data analysis by utilizing scalable algorithms, such as decision trees and deep learning, as well as by employing efficient preprocessing methods to clean and normalize data. Additionally, machine learning provides tools for making predictions and detecting anomalies, which help businesses make data-driven and accurate decisions. In conclusion, AI and machine learning are essential for analyzing large datasets, offering solutions to both technological and business challenges that traditional analytical methods cannot adequately address.

3 Utilizing Artificial Intelligence in Commercial Decision-Making

The rapid advancement of artificial intelligence and machine learning has begun to significantly transform commercial decision-making in large enterprises. These technologies allow companies to leverage large volumes of data more efficiently and make better, data-driven decisions. Data-driven decision-making has become a key factor in today's business environment, where competition is fierce, and companies require accurate and up-to-date information to support their decision-making processes. AI and machine learning provide businesses with the ability to analyze large datasets, predict future events, and optimize business processes. These technologies can be used to analyze customer behavior, identify market trends, and streamline supply chains. Johri et al. (2023) emphasize that big data and AI are revolutionizing business decision-making by providing companies with tools to make more precise and faster decisions, ultimately enhancing their competitiveness in global markets. This chapter examines the practical applications of AI and machine learning in decision-making within large enterprises, as well as the commercial benefits and challenges associated with these technologies.

3.1 Practical Applications in Decision-Making

3.1.1 Market Trend Analysis and Forecasting

Artificial intelligence and machine learning have significantly transformed how large enterprises analyze market trends and predict future business developments. These technologies enable companies to process vast amounts of data, identify patterns and trends, and forecast future market changes. This capability helps organizations make more accurate and faster decisions, which is particularly crucial in today's rapidly evolving business environment (Johri et al., 2023). AI and machine learning are widely used in analyzing large volumes of market data, as they allow for the automation of data-driven decision-making. Machine learning models, such as time series analysis and deep learning, can analyze historical market data, including sales figures, economic indicators, and consumer behavior,

to generate forecasts for future trends. This helps businesses anticipate market shifts and adjust their business strategies accordingly (Sun et al., 2022). Time series analysis is a commonly used method in market trend forecasting. It involves analyzing historical data to predict future developments based on past trends. Companies use time series models to forecast variables such as sales volumes or seasonal demand fluctuations. Deep learning models, such as convolutional and recurrent neural networks, are particularly effective for analyzing complex and unstructured data. These models are used when the dataset is extensive and includes sources such as news articles, social media content, or other external data, which can help identify hidden market trends. Deep learning has proven to be highly effective in predicting market developments and assisting businesses in adapting to rapidly changing conditions (Sun et al., 2022).

3.1.2 Customer Behavior Analysis and Personalization

Sun et al. (2022) present an example of how artificial intelligence can be utilized in decision-making systems. Their study demonstrates how companies can use AI-based tools to analyze vast market data and proactively identify changes in consumer behavior. Such tools provide businesses with the ability to make strategic decisions based on data analysis rather than relying solely on historical trends. For instance, companies can leverage these tools to optimize marketing campaigns, enhance customer experience, and adapt their products to rapidly changing market demands (Sun et al., 2022). AI and machine learning have transformed how companies understand and predict customer behavior. By analyzing large datasets, machine learning models can detect trends in customer behavior, allowing businesses to make precise and personalized decisions in their marketing strategies and customer experience development.

Machine learning models help businesses segment their customer base based on behavioral patterns, demographic information, and purchasing habits. This segmentation allows for more targeted marketing communication, improving both marketing efficiency and customer satisfaction. Personalization is another key application of machine learning, where companies can tailor marketing messages and offers to individual customers. Machine learning models analyze customer behavior, such as browsing history, purchasing patterns, and interactions on websites, to provide personalized recommendations based on their interests. This enhances the customer experience and increases the likelihood of purchases or continued service.

engagement. According to Johri et al. (2023), AI-powered customer behavior analysis tools help businesses better understand what customers want and need. AI can, for example, analyze customer interactions on social media, identify their interests, and provide real-time personalized recommendations. This improves customer satisfaction and fosters customer loyalty.

Predictive analytics is a key tool that allows companies to anticipate customer needs and respond to them in advance. This can include forecasting customer purchasing behavior, calculating customer lifetime value, or predicting customer churn (Johri et al., 2023). By leveraging predictive analytics, businesses can develop more effective customer strategies and enhance the customer experience. Recommendation systems are a concrete example of predictive analytics that improves customer experience. These systems analyze customer purchase data, browsing behavior, and past interactions to provide personalized product or service recommendations. This not only enhances customer satisfaction but also boosts sales and increases customer loyalty (Johri et al., 2023). Another important application area is customer churn prediction. Customer churn refers to the scenario where a customer stops using a service or switches to a competitor. Machine learning models can analyze customer behavior and detect anomalies that may indicate a potential churn risk. When such signs are identified, companies can intervene proactively by offering targeted campaigns or special discounts to encourage customers to remain loyal to the service. Predictive analytics enables companies to act proactively and ensure that customers receive the best possible experience tailored to their needs and behaviors.

3.1.3 Resource Optimization and Management

Artificial intelligence and machine learning have revolutionized resource optimization and supply chain management in large enterprises. These technologies enable companies to analyze vast amounts of data related to logistics, warehousing, and production and make fast, data-driven decisions. This enhances supply chain efficiency and allows for better resource management. The application of AI and machine learning in logistics and inventory management helps businesses anticipate future demands and optimize supply chains. Algorithms allow companies to analyze demand forecasts, inventory levels, and delivery

times, enabling them to optimize warehouse management and logistics routes. This reduces both overcapacity and undercapacity, lowering storage costs and improving delivery speed. For example, machine learning models can predict inventory needs by analyzing historical sales data and external factors such as seasonality and market trends. This allows businesses to maintain better inventory control, ensuring that products are available at the right place and time. Pani et al. (2021) emphasize that cloud-based machine learning models offer a scalable solution for real-time inventory management, optimized delivery routes, and improved supply chain efficiency. Supply chain optimization is particularly important for global companies that handle large volumes of products and information across multiple countries. AI helps companies streamline logistics, reduce delays, and improve delivery reliability. This leads to enhanced customer experience and cost savings.

Real-time data analysis is another significant AI application in supply chain management. Real-time analytics enables companies to quickly respond to supply chain disruptions and accelerates decision-making in production and logistics management. Machine learning models can analyze sensor data collected from production lines, warehouses, and logistics operations, providing real-time recommendations for responding to demand fluctuations or optimizing logistics routes. For example, many industrial companies use AI for production planning and predictive maintenance. Machine learning models can analyze real-time operational data from machinery and predict when maintenance is needed. This helps reduce unplanned downtime and improves overall production efficiency. Similarly, AI can optimize logistics routes based on changing weather conditions or delivery delays. Another crucial AI application is demand forecasting using real-time data. Companies can analyze sales and customer behavior in real time and adjust their production and logistics to meet shifting demand patterns. AI-powered solutions can anticipate sudden changes in market demand, allowing businesses to react quickly before stock shortages or oversupply issues arise (Sun et al., 2022).

3.1.4 Risk Management and Fraud Detection

The use of artificial intelligence and machine learning in business risk management and fraud detection has grown significantly. These technologies help companies identify risks and potential issues in advance by analyzing vast amounts of data, enabling quick responses and more effective risk management. Additionally, machine learning models can detect anomalies

and unusual activities, which helps prevent fraud and minimize financial losses. AI and machine learning allow businesses to analyze large volumes of operational data and identify hidden risks that may threaten their activities. Machine learning models can assess various risk factors, including financial indicators, market trends, and internal operations, providing businesses with the ability to anticipate problems before they escalate. Neural networks are widely used in risk analysis, as they can recognize complex relationships and predict future risks (Jamalpur et al., 2024). Jamalpur et al. (2024) highlight that AI models based on neural networks are particularly effective in identifying business risks, such as financial crises, market changes, and operational challenges. These models can be applied to both internal and external risk factors.

- In internal risk analysis, AI can help identify imbalances in resource allocation or bottlenecks in production processes.
- In external risk analysis, AI can monitor market developments, competitor activities, or geopolitical changes, allowing businesses to respond accordingly.

Beyond risk analysis, AI can also be used to predict supply chain disruptions and develop contingency plans. Real-time data analysis enables businesses to assess the likelihood of disruptions and plan preventive measures, which can help reduce costs and avoid significant business interruptions.

3.1.5 Supporting Strategic Decision-Making

The role of artificial intelligence and machine learning extends beyond operational activities and analytics; they also provide powerful tools for supporting strategic decision-making at the executive level. Today, companies widely use AI to assist in strategic planning and risk assessment, enabling better-informed decisions about the future direction of the business. AI utilization supports executive-level strategic decision-making by providing real-time data and analytics that influence long-term decisions. Machine learning models analyze large volumes of historical and real-time data, offering insights that help executives identify opportunities and threats in the business environment. This data-driven approach ensures that decisions are made objectively, reducing human error risks and improving predictability (Sun et al., 2022).

According to Sun et al. (2022), AI-powered decision support systems (DSS) provide executives with effective tools for strategic planning. These systems allow leadership to

evaluate different options, compare scenarios, and simulate decision impacts on the company's future development. Simulations enable businesses to test various strategic options and assess their potential effects before making critical decisions. This gives executives a clearer understanding of how different choices could impact business operations in the long term (Jamalpur et al., 2024). Machine learning models facilitate scenario simulations by analyzing market fluctuations, weather conditions, competitor actions, or geopolitical events and their impact on business operations. This allows companies to better prepare for unforeseen situations and develop strategies based on realistic, data-driven scenarios. AI serves as a valuable tool for risk management and strategic planning (Sun et al., 2022). For example, large industrial companies operating in multiple markets can use AI to assess how various economic or political scenarios might affect their business across different regions. Based on this data, companies can develop scenario-based action plans and proactively prepare for market changes. AI also helps minimize risks and enhance decision-making quality, allowing businesses to test different strategies without significant financial costs.

3.2 Commercial Benefits and Challenges

3.2.1 Efficiency Improvements Brought by AI

Artificial intelligence and machine learning allow businesses to automate various business processes, leading to significant efficiency gains and cost savings. Machine learning models analyze and optimize processes in real time, reducing manual work and minimizing errors. For example, optimizing production processes and logistics with AI can lower costs, enhance supply chain operations, and increase production capacity without requiring additional resources (Dandekar et al., 2024). Dandekar et al. (2024) emphasize that big data analytics implementation is particularly beneficial for companies that handle large volumes of data, such as retailers, logistics companies, and manufacturers. By automating processes and optimizing resource allocation, businesses can achieve substantial cost savings, as AI can identify efficiency bottlenecks within data and suggest optimal solutions to address them. This helps reduce overcapacity and improve resource utilization, leading to direct financial benefits.

3.2.2 Gaining a Competitive Advantage with AI

Companies that leverage artificial intelligence can achieve a significant competitive advantage compared to those that do not adopt this technology. AI enables more accurate predictions, better market trend analysis, and deeper customer behavior insights, giving businesses a greater ability to respond quickly to changing market conditions. This is particularly crucial in highly competitive industries, where companies must continuously innovate and adapt (Johri et al., 2023). Johri et al. (2023) emphasize that integrating AI into business decision-making and analytics allows companies to differentiate themselves from competitors. Businesses that effectively utilize AI can make faster and more accurate decisions, identify market opportunities ahead of competitors, and offer personalized solutions to customers. This enhances the customer experience and strengthens customer loyalty, ultimately increasing a company's long-term competitiveness.

AI-driven decision-making also reduces the risk of human errors and subjective biases, thereby improving the quality of decision-making. Data-driven decision-making is based on objective information analyzed by AI, enabling businesses to make fact-based and precise

choices. This is particularly important in strategic decision-making, where decisions significantly impact a company's long-term success. AI-powered decision support systems help minimize subjectivity and cognitive biases in decision-making. Machine learning models can analyze large amounts of historical and real-time data, allowing for well-founded and data-driven decisions. This reduces uncertainty and enhances decision-making accuracy, helping businesses develop more sustainable and low-risk strategies.

3.2.3 Technological Challenges and Legacy Systems

While artificial intelligence and machine learning offer significant commercial benefits, their adoption and implementation do not come without challenges. Technological limitations, lack of employee expertise, ethical and legal concerns, and organizational culture changes can present major obstacles to AI implementation in large enterprises (Dandekar et al., 2024). One of the most significant technological challenges in adopting AI in large companies is integrating it into existing, often outdated, IT systems. These so-called legacy systems are often critical to daily business operations, and replacing or upgrading them can be complex and costly. Legacy systems were not originally designed to support modern AI and machine learning solutions, leading to significant technical challenges (Kirola et al., 2023).

First, technological compatibility issues are common. Legacy systems may use outdated programming languages, database solutions, and infrastructures that are not compatible with AI-powered applications. This makes data sharing and integration difficult, which is critical for AI systems to function properly. For example, real-time data analysis and decision-making require AI solutions to access up-to-date data, which can be challenging with older systems. Second, cybersecurity and data privacy concerns are heightened during integration. When AI solutions are integrated into existing systems, it is essential to ensure that security standards are met and that sensitive data remains protected (Dandekar et al., 2024). Legacy systems may not comply with modern security protocols, exposing the company to risks. This often requires thorough system evaluations and potential upgrades to meet current cybersecurity requirements.

Third, organizational challenges, such as employee resistance and change management, can slow down the integration process. Employees may be accustomed to certain systems and workflows, and the introduction of new AI solutions can cause uncertainty or resistance to

change. It is crucial to clearly communicate the reasons and benefits of the transition, as well as provide the necessary training and support for employees (Dandekar et al., 2024).

Additionally, financial factors pose a major obstacle. Integration may require significant investments in hardware, software, and expertise. Companies often need to upgrade their infrastructure, acquire new technologies, and possibly hire external experts to implement the integration. These costs can be substantial, especially if the company's systems are heavily outdated. In summary, integrating AI into existing systems is a complex process that requires careful planning, resources, and change management. Companies must develop strategies that allow for a smooth integration, minimizing business disruptions while maximizing the benefits of new technology (Dandekar et al., 2024).

3.2.4 Infrastructure Requirements and Environmental Impact

The implementation of AI solutions, especially when dealing with large-scale datasets, requires significant computing resources. Machine learning models, such as deep learning neural networks, can contain millions of parameters, which must be trained on large datasets. This training process demands high computational power, meaning the use of high-performance processors such as GPU- or TPU-based systems (Kirola et al., 2023).

Furthermore, the continuous use of AI in production environments requires substantial computing power for real-time data processing. This can result in high operational costs, particularly when using cloud services, where computing power is billed based on usage. Another critical factor is storage capacity. Large datasets require vast amounts of storage, and data must be readily accessible for analysis. This often necessitates investments in high-speed storage solutions, such as SSD drives or distributed data warehouses. Additionally, data backup and recovery systems must be upgraded to accommodate the increasing storage needs and security requirements.

Network infrastructure capacity is also a crucial factor. Transferring and processing large datasets requires fast and reliable network connections. This often means investing in higher-speed networking equipment and connections, as well as optimizing network performance. Without sufficient network capacity, data processing can experience significant slowdowns, negatively affecting AI solution efficiency (Dandekar et al., 2024). Additionally, energy efficiency and environmental impact must be considered. Large-scale data centers consume vast amounts of energy, increasing operational costs and environmental footprint. Companies

must consider sustainable development strategies and explore options for energy-efficient solutions or renewable energy sources to mitigate these effects.

3.2.5 Challenges and Strategies for AI Adoption

Effectively utilizing AI and machine learning requires specialized expertise, which may be limited within organizations. There is a global shortage of data analysts, data scientists, and machine learning engineers, making recruitment difficult and increasing salary costs (Dandekar et al., 2024). Additionally, many employees lack the necessary skills to develop or implement AI solutions, requiring significant investments in training. This applies not only to technical staff but also to business leaders, who need to understand AI's capabilities and limitations to make informed decisions (Kirola et al., 2023).

The adoption of AI solutions requires cultural changes within organizations and strong change management. Leadership commitment is crucial to ensure that new technologies are integrated into operations and that employees can adapt smoothly to new processes. Employee concerns about job displacement or changing roles may lead to resistance, highlighting the importance of open communication and training. Jamalpur et al. (2024) emphasize that addressing employee fears and managing organizational change are key challenges in AI implementation. Overcoming these challenges requires a structured and consistent approach to ensure successful AI adoption.

For successful AI adoption, companies should leverage phased implementation, partnerships, and continuous skill development. A phased implementation approach, where AI solutions are introduced through pilot projects, helps mitigate technological risks and manage costs. This strategy allows businesses to test and refine technologies before rolling them out on a larger scale (Kirola et al., 2023).

Many organizations lack sufficient in-house expertise for AI adoption, making strategic partnerships with external experts or consulting firms a viable solution. Outsourcing the design and implementation of AI solutions allows companies to leverage the latest technologies and best practices, accelerating adoption and helping navigate technological and regulatory challenges (Sun et al., 2022). Training and skill development are essential for the long-term success of AI utilization. The continuous advancement of technology requires

employees to engage in lifelong learning, necessitating investments in skill development across all organizational levels. This includes both technical skills and leadership skills, ensuring that executives can effectively guide AI-driven strategies (Dandekar et al., 2024). By investing in training, companies can ensure that their workforce can fully utilize AI solutions and adapt to rapid technological changes.

3.2.6 Ethics and Data Privacy

The use of artificial intelligence raises significant ethical and data privacy concerns, which businesses must carefully address. It is essential to establish clear ethical principles and guidelines to ensure that AI solutions are fair, transparent, and compliant with regulations. This includes ensuring lawful handling of personal data, avoiding discrimination and biases, and improving model explainability. Particularly in Europe, GDPR regulations impose strict limits on the collection and use of personal data, which can add complexity to AI adoption. Data privacy concerns and the "black box" problem of AI models pose major challenges. Complex machine learning models, such as deep learning networks, can be difficult to interpret and explain, leading to trust issues among customers and stakeholders. Building transparency is key to strengthening trust, and Explainable AI (XAI) methods can help make AI-driven decisions more understandable. Dandekar et al. (2024) and Jamalpur et al. (2024) emphasize the importance of data ethics and transparency in AI adoption. Companies must develop practical guidelines that are clear and easy to apply across all organizational levels. By implementing transparent processes, organizations can not only mitigate risks associated with AI misuse but also build trust both internally and externally.

3.3 Summary

Chapter 3 examined how artificial intelligence and machine learning have significantly transformed commercial decision-making in large enterprises. The key findings indicate that AI helps businesses make more efficient, data-driven decisions, improving competitiveness, reducing costs, and increasing decision-making accuracy.

With AI-driven solutions, companies can:

- Analyze market trends
- Optimize supply chains

- Predict customer behavior
- Manage risks using real-time data analysis

The research question TK2: "How can AI and machine learning support commercial decision-making in large enterprises?" was addressed through the analysis in Chapter 3.

AI enables large enterprises to:

- Optimize resources
- Enhance business processes
- Deliver better customer experiences through predictive analytics
- Make more accurate strategic decisions using simulations and data analysis

This strengthens decision-making quality, reduces risks and unnecessary costs, and improves overall efficiency. The key conclusions suggest that AI enhances decision-making speed and accuracy, which is particularly valuable for businesses operating in rapidly changing market conditions. Companies that adopt AI gain a competitive advantage, as AI improves their ability to anticipate future business environment changes and evolving customer needs.

4 Summary

This literature review examined the role of artificial intelligence and machine learning in analyzing large datasets and their impact on commercial decision-making in large enterprises. The key findings indicate that AI and machine learning enable efficient processing of vast amounts of data and support strategic decision-making by providing in-depth analyses and forecasts. These technologies facilitate automated data preprocessing, analysis, and business application. For example, neural networks allow businesses to simulate various business scenarios and assess their impact (Jamalpur et al., 2024). Additionally, AI-driven decision support systems help optimize resource utilization and reduce business risks (Sun et al., 2022).

AI accelerates decision-making by analyzing large datasets in real time, enabling more precise and timely decisions. Johri et al. (2023) found that AI implementation in business reduces human errors and improves operational efficiency. Through predictive analytics, businesses can anticipate market trends and customer behavior, supporting strategic decision-making and resource optimization. Moreover, AI helps detect and manage anomalies, such as unusual events in large datasets, enhancing risk management and response capabilities.

However, leveraging AI and machine learning requires significant investments in infrastructure, computing power, and workforce expertise. Kirola et al. (2023) emphasize the importance of parallel processing and distributed databases in handling large datasets. At the same time, a lack of technological expertise among employees remains a major barrier to effective AI adoption. Companies must invest in continuous training and provide employees with opportunities to learn AI tools, such as TensorFlow and PyTorch, to ensure efficient implementation and utilization of AI technologies.

Additionally, data privacy, ethics, and regulatory compliance present significant challenges. Dandekar et al. (2024) highlight that businesses must adhere to strict ethical standards and ensure that their operations comply with legal frameworks. Doing so strengthens customer trust and helps companies avoid potential legal issues.

AI and machine learning are indispensable technologies in modern business, but their successful implementation requires a strategic and responsible approach. Companies should

prioritize continuous training, as developing employees' technological expertise is crucial. It is essential to offer regular training programs and establish learning initiatives that focus on the effective use of AI and machine learning tools, such as TensorFlow and PyTorch. At the same time, fostering collaboration is key. Companies should actively participate in AI ecosystems, including research programs and startup partnerships, as these can accelerate innovation and introduce new solutions.

Data privacy and ethics must be a priority. Businesses should establish clear guidelines for data ethics and ensure that AI is used responsibly and transparently. This not only minimizes risks but also strengthens the company's reputation and builds trust among stakeholders. Additionally, investing in scalable technologies, such as cloud services provided by AWS, Google Cloud AI, and Azure Machine Learning, is essential. These platforms offer cost-effective solutions for processing large datasets and facilitate AI adoption within organizations.

In the future, companies should focus on leveraging AI and machine learning in new business areas, such as sustainable development and enhancing customer experience. Technological advancements, including generative AI and multimodal models, present new opportunities that businesses should explore and experiment with. Finally, it is important to recognize that AI and machine learning are not just technological tools, but strategic investments that can transform business operations and enhance competitiveness. When used responsibly and with a long-term perspective, AI and machine learning provide significant value to both businesses and their customers.

References

1. M. Kirola, M. Memoria, M. Shuaib, K. Joshi, S. Alam and F. Alshanketi, "A Referenced Framework on New Challenges and Cutting-Edge Research Trends for Big-Data Processing Using Machine Learning Approaches," 2023 International Conference on Smart Computing and Application (ICSCA), Hail, Saudi Arabia, 2023, pp. 1-5, doi: 10.1109/ICSCA57840.2023.10087686,
2. M. Dandekar, S. Lote and P. Dandekar, "Implementing the Power of Big Data Analytics," 2024 IEEE 3rd International Conference on Electrical Power and Energy Systems (ICEPES), Bhopal, India, 2024, pp. 1-5, doi: 10.1109/ICEPES60647.2024.10653580.
3. Subhendu Kumar Pani; Somanath Tripathy; George Jandieri; Sumit Kundu; Talal Ashraf Butt, "Applications of Machine Learning in Big-Data Analytics and Cloud Computing," in Applications of Machine Learning in Big-Data Analytics and Cloud Computing , River Publishers, 2021, pp.i-xxxii.
4. Amazon Web Services. (2023). Introducing the Next Generation of Amazon SageMaker: The Center for All Your Data Analytics and AI. Retrieved from: <https://aws.amazon.com/blogs/aws/introducing-the-next-generation-of-amazon-sagemaker-the-center-for-all-your-data-analytics-and-ai/> (30.12.2024)
5. Google Research. (n.d.). TensorFlow: A System for Large-Scale Machine Learning. Retrieved from: <https://research.google/pubs/tensorflow-a-system-for-large-scale-machine-learning/> (30.12.2024)
6. Google Open Source. (n.d.). TensorFlow Open Source Project. Retrieved from: <https://opensource.google/projects/tensorflow> (30.12.2024)
7. How Netflix Uses Python. (2023). TechRepublic: How Netflix Uses Python, Streaming Giant Reveals Its Programming Language, Libraries, and Frameworks. Retrieved from: <https://www.techrepublic.com/article/how-netflix-uses-python-streaming-giant-reveals-its-programming-language-libraries-and-frameworks/> (30.12.2024)
8. S. Johri, K. V. Rawal, B. Aishwarya, N. Singh, A. M. Shaaker and R. V, "Big Data and Artificial Intelligence: Revolutionizing Business Decision-Making," 2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), Gautam Buddha Nagar, India, 2023, pp. 1689-1693, doi: 10.1109/UPCON59197.2023.10434500.
9. B. Jamalpur, P. Sarkar, M. H. Krishna, M. Lourens, K. S. Lakshmi and N. Singh, "Strategic Decision Making in the AI Age: Examining Neural Network Models in Business Management," 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India, 2024, pp. 1496-1501, doi: 10.1109/IC3SE62002.2024.10593321.
10. Q. Sun, G. Chen and S. He, "Enterprise Development Decision Support System Based on Artificial Intelligence Technology," 2022 International Conference on Artificial Intelligence of Things and Crowdsensing (AIoTCs), Nicosia, Cyprus, 2022, pp. 245-249, doi: 10.1109/AIoTCs58181.2022.00044.