



Research Article

RESEARCH ON APPLICATION OF ARTIFICIAL INTELLIGENCE IN DATA MINING

Rajeswari J* and Sasikala B

Department of Computer Science (Artificial Intelligence) Sri GVG Visalakshi College for Women,
Udumalpet

ARTICLE INFO

Article History:

Received 14th January, 2022

Received in revised form 29th

February, 2022

Accepted 05th March, 2022

Published online 28th April, 2022

Key words:

Mining, Intelligence, Techniques, Machine Learning, Scope, Limits.

ABSTRACT

Data mining is a relatively new and fast-growing field. Machine learning, statistics, database research, high-performance computing, and commerce are among the areas that provide ideas and resources. This explains why data mining is such a dynamic, diverse, and fast evolving discipline. In huge and expanding data sets, data mining techniques are employed to identify patterns, structure, regularities, and singularities. In other words, Data Mining is the process of analysing hidden patterns of data from various perspectives for categorization into useful data, which is collected and assembled in specific areas such as data warehouses, efficient analysis, data mining algorithms, assisting decision making, and other data requirements, ultimately resulting in cost-cutting and revenue generation. Data mining is the process of automatically examining enormous amounts of data for patterns and trends that go beyond simple analysis[2]. Data mining estimates the probability of future events by utilising advanced mathematical algorithms for data segments. Data mining is also known as data knowledge discovery (KDD). By outsourcing data mining, all of the work may be completed more quickly and at a lower cost. Specialized businesses can also take advantage of new technology to acquire data that would otherwise be impossible to locate manually. Although there is a wealth of material available on multiple platforms, there is a scarcity of expertise.

Artificial intelligence (AI) is a field of science concerned with the development of intelligent machines. These machines are referred to as intelligent because they possess human-like reasoning and decision-making abilities. Artificial intelligence is the study of developing intelligent machines that can perform tasks similar to humans. It does not rely on learning or feedback; instead, it contains control systems that are explicitly designed. By calculating, AI systems come up with solutions to issues on their own. AI systems use the data mining technique to create solutions from mined data. Artificial intelligence is built on the foundation of data mining. Data mining is a subset of programming codes that contain the information and data that AI systems require. Machine Learning is a key part of Artificial Intelligence[1]. This means that AI's intelligent behaviour is based on machine learning algorithms. A computer is considered to learn from a task if the error lowers over time and the performance is as expected.

Copyright©2022 Rajeswari J and Sasikala B. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Data mining, also known as knowledge discovery in databases, is the nontrivial extraction of implicit, previously unknown, and potentially usable information from data in artificial intelligence and machine learning. The widespread availability of huge volumes of data, as well as the need to transform such data into valuable information and knowledge, has sparked interest in data mining. Risk monitoring, company management, manufacturing control, market analysis, engineering, and science discovery are all possible applications of the knowledge gathered. Data mining is proving to be an excellent tool for pursuing new paths for automatically examining, visualising, and uncovering patterns in data that aid decision-making[4]. Data mining is a technique for identifying

patterns in data that go beyond simple examination. Many disciplines employ modern data mining techniques to tackle association, classification, segmentation, diagnosis, and prediction problems (association rules, decision trees, Gaussian mixture models, regression algorithms, neural networks, support vector machines, Bayesian networks, and so on). One sample definition is based on a comparison of computer devices' intellect to that of humans. Another meaning is concerned with machine performance, which has traditionally been considered to be under the domain of intelligence.

Artificial intelligence in data mining techniques are widely used to address categorization, planning, prediction, optimization, diagnosis, computing, collecting and analysing

*Corresponding author: **Rajeswari J**

Department of Computer Science (Artificial Intelligence) Sri GVG
Visalakshi College for Women, Udumalpet

consumer data, glean insights into what customers want and need, and acting on those insights in a variety of domains[3]. Data mining and big data can be found practically everywhere, and it is critical to protect the data that is generated in large quantities so that nothing is lost. This type of data is frequently processed using artificial intelligence. Artificial Intelligence and its subbranches (such as Machine Learning, Deep Learning, and Neural Networks) are algorithmic in nature.

Artificial Intelligence and Data Mining

Artificial intelligence is the study of developing intelligent machines that can perform tasks similar to humans. It does not rely on learning or feedback; instead, it contains control systems that are explicitly designed. By calculating, AI systems come up with solutions to issues on their own. AI systems use the data mining technique to create solutions from mined data. Artificial intelligence is built on the foundation of data mining[6]. Data mining is a subset of programming codes that contain the information and data that AI systems require. Machine learning is about extracting knowledge from the data. It can be defined as,

“Machine learning is a subfield of artificial intelligence, which enables machines to learn from past data or experiences without being explicitly programmed”.

Without being explicitly coded, machine learning allows a computer system to generate predictions or make decisions based on historical data. Machine learning makes use of a large amount of structured and semi-structured data in order for a machine learning model to produce reliable results or make predictions based on it.

Machine learning is based on an algorithm that learns on its own with the use of previous data. It only works for restricted domains; for example, if we create a machine learning model to detect dog pictures, it will only return results for dog pictures; however, if we add fresh data, such as a cat picture, it will become unresponsive. Machine learning is utilised in a variety of applications, including online recommender systems, Google search engines, email spam filters, and Facebook auto friend tagging suggestions, among others.

Data mining and Machine Learning are both scientific disciplines[7]. Despite the fact that these names are sometimes used interchangeably, there are some significant differences between them.

Scope: Data mining is a process that uses patterns and data visualisation tools to figure out how different attributes in a data set are related to one another. The purpose of data mining is to determine the link between two or more attributes in a data set and then utilise that information to anticipate outcomes or actions.

Machine Learning is used to make predictions about the outcome, such as a price estimate or an approximation of time length. It automatically learns the model over time as it gains experience. It gives immediate feedback.

Function: Data mining is a technique for extracting meaningful information from large amounts of data. Machine Learning, on the other hand, is a method of iteratively providing a machine with a training dataset to improve complex algorithms.

Uses: Data mining is more commonly utilised in the research field, whereas machine learning is more commonly used in making product, price, and time recommendations.

Concept: The idea behind data mining is to use ways to extract information and identify trends and patterns.

Machine Learning is based on the idea that machines can learn from existing data and develop on their own. Machine learning employs data mining techniques and algorithms to create models based on the logic of data that forecast future outcomes. Math and programming languages are used to create the algorithms.

Method: Data mining is used by Machine Learning to refine its algorithms and change its behaviour in response to future inputs. As a result, data mining serves as a source of input for machine learning.

Machine learning algorithms will continuously run and improve the performance of the system automatically, and also analyze when the failure can occur. When there is some new data or change in the trend, the machine will incorporate the changes without the need to reprogram or any human interference.

Nature: Machine learning differs from data mining in that it learns on its own, whereas data mining requires human intervention to use procedures to retrieve data.

Learning Capability: Machine Learning is a step forward of data mining because it employs the same principles to learn and adapt to changes automatically. It's a lot more precise than data mining. Data mining is a manual process because it requires a human to initiate the analysis.

Implementation: Data mining entails creating models that are then used to use data mining techniques. Models such as the CRISP-DM model are developed. For knowledge discovery, the data mining method employs a database, a data mining engine, and pattern evaluation.

Machine Learning algorithms are used in artificial intelligence, neural networks, neuro-fuzzy systems, and decision trees, among other applications. To forecast outcomes, machine learning employs neural networks and automated algorithms.

Accuracy: Data mining accuracy is determined by how data is acquired. Data mining generates precise results that machine learning can use, allowing machine learning to create superior outcomes.

Data mining may miss crucial links because it necessitates human interaction. Data Mining approaches have been shown to be less accurate than machine learning algorithms.

Applications: Machine learning methods require data to be given in a consistent format, which limits the number of algorithms available. To use machine learning to evaluate data, data from various sources should be converted from native format to a standard format that the machine can understand. It also requires a large amount of data for accurate results. This is an overhead when compared to data mining.

Examples: Data mining is used in identifying sales patterns or trends while machine learning is used in running marketing campaigns.

Classification of Data Mining Techniques:

The widespread availability of huge volumes of data, as well as the need to transform such data into valuable information and knowledge, has sparked interest in data mining. Risk monitoring, company management, manufacturing control, market analysis, engineering, and science discovery are all possible applications of the knowledge gathered.

In general, three types of data mining techniques are used: association, regression, and classification.

Association analysis

The finding of association rules revealing attribute-value conditions that occur frequently together in a particular collection of data is known as association analysis. The technique of association analysis is commonly utilised to determine the relationship between specific products in shopping carts.

Regression analysis

Regression analysis builds models that use the analysis of independent factors to explain dependent variables. Correlating the product pricing and the typical customer income level, for example, can be used to predict a product's sales performance.

Classification and prediction

The process of building a collection of models to predict the class of objects whose class label is unknown is known as classification. If-then rules, decision trees, or mathematical formulas can all be used to express the derived model.

A decision tree is a tree structure that looks like a flowchart, with each node representing a test on an attribute value, each branch representing the test's outcome, and each tree leaf representing a class or class distribution. Classification rules can be created from decision trees.

Data items' class labels can be predicted via classification. The term "prediction" refers to the process of identifying distribution trends based on available data. The data mining process consists of an iterative sequence of the following steps:

- Data coherence and cleaning to remove noise and inconsistent data
- Data integration such that multiple data sources may be combined
- Data selection where data relevant to the analysis are retrieved
- Data transformation where data are consolidated into forms appropriate for mining
- Pattern recognition and statistical techniques are applied to extract patterns
- Pattern evaluation to identify interesting patterns representing knowledge
- Visualization techniques are used to present mined knowledge to users

The classification of machine learning methods

Rule induction

The goal of rule induction is to create a decision tree or a set of decision rules from the training data. The fundamental advantage of rule induction is that it can analyse big data sets

quickly and is ideal for classification and prediction applications. The results are simple to understand and implement.

Neural networks

The neural network is made up of processing nodes that look like neurons in the human brain. To construct a multi-layer network structure, the input node is connected to the output node via a hidden node. The neural network is trained on historical sample data over and over again. The most significant benefit of a neural network is its ability to accurately foresee complicated issues.

Genetic algorithms

A genetic algorithm is a combinatorial optimization method that is based on the process of biological evolution. The core concept is that the fittest and best or better individuals will survive. Reproduction, hybridization, and mutation are all part of the operation process. Genetic algorithms have the advantage of being simple to combine with other systems.

Inductive logic programming

To define and explain concepts, inductive logic programming use first-level attribute logic. It begins with defining positive and negative examples, followed by a ranking of the new examples. This technique has a powerful conceptual description mechanism and is capable of expressing complex relationships.

Classification of machine learning tasks in data mining

Classification

A classification model is created using the training data set. The categorization model can then divide the data into many groups automatically[8]. KNN classification algorithm, naive bayes classification method, decision tree, artificial neural network, and support vector machine are examples of machine learning classification algorithms.

Regression analysis

The relational expression between variables and variables is obtained by evaluating the data and applying statistical methods. These built-in laws are utilised to forecast and estimate future trends. Regression trees, artificial neural networks, linear regression, and logic regression are all used to create regression models.

Association rules

Among transactional data, there are rules of association. Frequent itemsets can be obtained by mining the relationship between transactional data. The chance of particular transactions occurring at the same time can be predicted using this information. Apriori is a well-known association rule mining algorithm.

Clustering

Multiple data without category labels are aggregated in a number of different clusters using the mining process, so that data items within the cluster are similar to each other and data objects between clusters are dissimilar. K-means clustering is a well-known clustering algorithm.

CONCLUSION

In machine learning, data mining, data science, and deep learning, data becomes the most significant aspect. Today's environment requires a great deal of data analysis and understanding. As a result, firms must decide whether or not to spend time, effort, and money on these analysis tools. Data mining using machine learning technology has been used in a variety of industries, including finance, retail, insurance, and telecommunications. Financial analysts, for example, utilise data mining to develop prediction models to discover the patterns that have caused market volatility in the past, boosting their ability to predict market volatility. Because data is developing at such a rapid rate, these methods must be able to incorporate new data sets quickly and forecast valuable analyses.

Machine learning can assist us in fast processing data and automatically delivering findings in the form of models. Salespeople in the retail industry can use data mining to create predictive models to figure out who is most likely to reply to letters, resulting in increased sales. When using data mining technologies, businesses should be aware of the benefits and drawbacks of various technologies and processes, and choose the best technology for the environment and task at hand. Data mining techniques use previous data to create patterns and trends that can be used to forecast future results. These results come in the shape of graphs, charts, and other visual representations. As corporate processes improve, these technologies will explode in popularity. As a result, businesses will be able to automate laborious processes, enhance sales and profitability, and improve client retention.

References

1. AnZengbo, ZhangYan. The Application Study of Machine Learning[J]. Journal of Changzhi University, 2007, 24(2):21-24.

2. ChenXiao. Application of machine learning algorithm in data mining [J]. Modern Electronics Technique, 2015,38(20):11-14.
3. HeQing. A Survey of Machine Learning Algorithms for Big Data[J]. Pattern Recognition and Artificial Intelligence, 2014(4):327 -336.
4. LiYun. Application of machine learning algorithms in data mining[D]. Beijing: Beijing University of Posts and Telecommunications, 2014.
5. R.Groth.HouDi. Data Mining - Building Competitive Advantages of Enterprises[M]. Xi'an:Xi'anJiaotong University press, 2001.
6. WangXiao. Research on Trends of Machine Learning Algorithms in Big Data Environments[J]. Natural sciences journal of harbin normal university, 2013 (4): 48-50.
7. YangJingfang. The application of machine learning algorithm in data mining[J]. Electronic Technology & Software Engineering, 2018(04):191.
8. Zhang Shaocheng. Research and Application of Machine Learning in Data Mining Based on Big Data[J]. Journal of Liaoning university Natural Sciences Edition, 2017, 44(1):15-17.
9. ZhaoYijun, ShangMengjiao. The characteristics of data mining as a cross discipline[J]. Times Finance, 2017(03):263-264.
10. ZouYi. Overview of Data mining technology[J]. Information & Communications, 2016(12):164-165.

How to cite this article:

Rajeswari J and Sasikala B (2022) 'Research on Application of Artificial Intelligence In Data Mining', *International Journal of Current Advanced Research*, 11(04), pp. 588-591. DOI: <http://dx.doi.org/10.24327/ijcar.2022.591.0129>
