A Survey on AI and ML Techniques

Rishikesh Nandankar, Rutvik Raut

rishikeshn.ce19@stvincentngp.edu.in, rutvikr.ce19@stvincentngp.edu.in

St. Vincent Pallotti College of Engineering & Technology, Gavsi Manapur, Wardha Road, Nagpur – 441108, India.

Abstract: **Nowadays due to the exploitation of a large amount of data, there is a need to manage and store this data. AI/ML are the technologies used to do so. These methodologies use some techniques to perform operations on the data. AI technology helps a machine to get self intelligent and perform tasks like humans whereas ML technology makes the machine learn from past experiences using AI. The techniques used by AI/ML are explained and compared in this paper.**

1. Introduction

Current day, AI/ML is the most popular technology known by every person in the software industry. In current times data is being generated in large amounts. To manage those data and to make work easier and manageable these technologies are coined. These methodologies are implemented in the industry to collect and manipulate data.

1.1 Artificial Intelligence (AI)

The term Artificial Intelligence (AI) is known as the capability of an electronics device to perform the duty related to Human Intelligence, this technique is widely operated in the implementation of new projects. It is a smart machine that can do tasks that typically require human intelligence. It also endeavors to replicate or simulate human intelligence into machines. In general, Artificial Intelligence is trying to make machines/computers act and think like Human beings.

1.2 Machine Learning (ML)

Machine Learning is one of the parts of Artificial Intelligence which directs the learning knowledge system through the data, surroundings, and observations, also in general it's a way to teach the machine to acquire the Human surroundings by improving their man-made Intelligence.

1.2.1 Supervised Learning

Supervised learning is an approach to Machine Learning in which we train machines using labeled data. It means that we need a supervisor or an instructor to inspect the machine. It helps you to collect data or generate data from past experiences. It is used to solve real-world problems and based on past experience the machine optimizes its performance of the machine.

1.2.2 Unsupervised Learning

Unsupervised learning is an approach to Machine learning in which there is no need for a supervisor or an instructor to inspect the machine. It is the automated machine that discovers its data. The data in this technique is unlabeled. It can perform more complex tasks than supervised learning. In unsupervised learning, the data

is not much more accurate than in supervised learning.

Table 1 - Difference between Supervised and Unsupervised Learning

	Supervised	Unsupervised
Types	It is used for two types of problems: Regression and Classification.	It is used for two types of problems: Clustering and Association.
Classes	In this technique, classes are known.	In this technique, classes are unknown.
Feedback	Direct feedback is accepted by the machine.	No feedback is taken by the machine.
Precision	It gives precise results.	It gives less precise results.
Application	Pattern Recognition	Fraudulent detection

1.3 Importance

Processing of data became easier due to the implementation of AI & ML into data industries. All industries are facing flooding of data and difficulty to store, process, and evaluating the data. It is also used to automate the machines to work on their own using some intelligence. It has been used for the past few years in medical technologies like robots performing operations under the guidance of doctors. In the automobile industry, Intelligent robots are used to operate bulky and risky operations which normal human beings cannot perform. In the Industrial sector where there is a lot of demand, these technologies are adaptable to any situation according to market conditions. As it predicts accurate outcomes, these are being popular these days.

Fig. 1 Fields Of AI
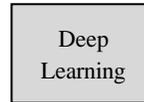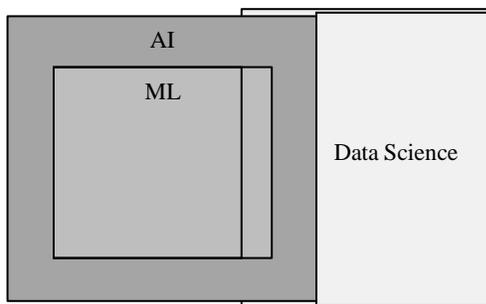
2. Techniques of AI/ML

AI and ML use some techniques to make the machine intelligent and work according to human needs. In past times, these techniques have progressed to an extent to make their implementation simple in technology.

2.1 Artificial Intelligence (AI)

Some of the techniques of AI are listed below:

2.1.1 Machine Learning:

In this aspect of AI, machines can learn from their past experiences rather than programming them explicitly. There are some algorithms like Unsupervised Learning, Supervised Learning, and Reinforcement Learning. Whereas in Unsupervised Learning without any guidance the information is processed in the machine. While in supervised Learning there is the need for guidance over the machine to perform tasks and Reinforcement Learning is an intelligence of a machine that makes decisions according to the environment performed by a software agent.

2.1.2 Natural Language Processing (NLP):

Natural language processing is one of the techniques of AI which helps the computer to understand human language. The data to be considered can be in the form of text or audio (written or spoken). The

data in one form can be converted into another form at both ends.

2.1.3 Automation and Robotics:

It uses mathematical expressions and algorithms to learn from past experiences. It is also used to perform repetitive tasks and also it enhances the productivity and effectiveness of the solution.

2.1.4 Machine vision:

It is the automation of the system which automatically inspects and analyzes the data. In other ways, we can say that machine vision is the brain behind all processes performed by the system.
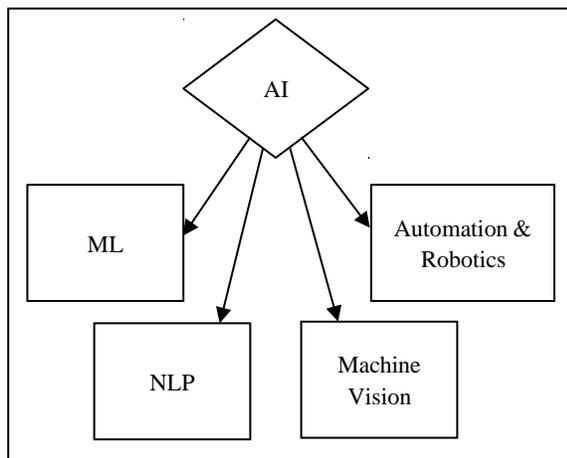

Fig. 2: Techniques in AI

2.2 Machine Learning (ML)

Some of the techniques of ML are listed below:

2.2.1 Regression:

It is the mathematical way of representing continuous data based on two or more data parameters and predicting the result.

2.2.2 Classification:

Classification of data categorizes the data into different classes using a label and predicts the output.

2.2.3 Clustering:

It is a technique of grouping out similar data points into different clusters i.e. dividing data points one having similar properties and the other having dissimilar properties from the first group.

2.2.4 Dimensionality Reduction:

It is the technique used to reduce the number of input variables in the data. It helps to keep the machine learning dataset structure simpler considering the number of parameters in the dataset.

2.2.5 Decision Trees:

A decision tree is a technique; covers both regression and classification is a flowchart that contains the root, branches, and leaf nodes. It is useful for taking implicit and explicit decisions in the system.

2.2.6 Neural Networks:

This is an artificial network of several neurons arranged in different layers used to translate data from one form to another form.
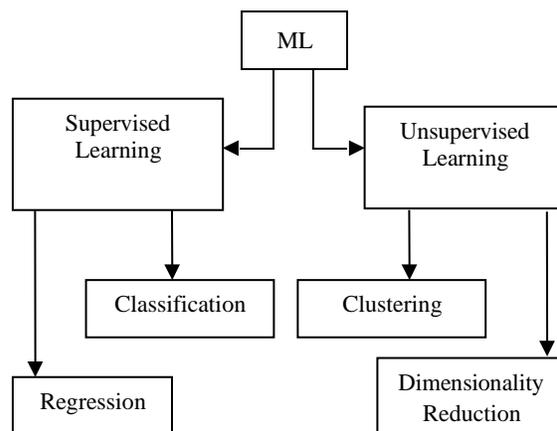
Fig. 3 Techniques in ML

Table 2 - Difference between Classification and Regression

	Classification	Regression
Input	Predicts the category of the input data.	Predicts continuous values of the input data.
Mapping Function	Generate approximate mapping function from input data to forecast output data.	Generate mapping function based on input data and output data.
Output	Output data is discrete.	Output data is constant.
Types	Multi-class classifier and Binary classifier	Non-linear and linear regression
Algorithms	i. Decision Tree classification ii. Random Forest classification iii. K-nearest neighbor	i. Simple linear regression ii. Multiple Linear regression iii. Polynomial regression
Examples	Spam detection in Gmail.	Analysis of business data.

Table 3 - Difference between Clustering and Dimensionality Reduction

	Clustering	Dimensionality Reduction
Basic Definition	It is the grouping of data.	It is the contraction of data.
Function	It finds similarities between the clusters in the data.	It reduces the useless data in the dataset.
Data Loss	Fewer chances of data loss.	More chances of data loss.
Techniques Used	K-means clustering, hierarchical clustering, etc.	Factors-based reduction and projection-based reduction.

3. Conclusion:

In this paper, the similarities and dissimilarities between the different techniques of AI as well as ML-like supervised unsupervised based on types, classes, feedback, precision, and application; classification and regression based on input, mapping function, output, types, algorithms, and some examples; clustering and dimensionality reduction based on its basic definition, function, data loss, and techniques used. We have also covered basic definitions of techniques used in AI and ML.

4. References:

https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning

https://www.guru99.com/supervised-vs-unsupervised-learning.html

N. Polyzotis, S. Roy, S. E. Whang, and M. Zinkevich, "Data lifecycle challenges in production machine learning: A survey," SIGMOD Rec., vol. 47, no. 2, pp. 17–28, Jun. 2018.

Vadari, Sekhar and Desik, P.H. Anantha, The Role of AI/ML in Enhancing Knowledge Management Systems (October 11, 2021). The IUP Journal of Knowledge Management, Vol. 19, No. 2, April 2021, pp. 7-31, Available at SSRN: https://ssrn.com/abstract=3940112

Agarwal, A., Gans, J., & Goldfarb, A. (2018). Prediction machines: The simple

economics of artificial intelligence. Harvard Business Press, Boston, MA, USA.

Benk, M., & Ferrario, A. (2020). Explaining Interpretable Machine Learning: Theory, Methods, and Applications (SSRN Scholarly Paper ID 3748268). Social Science Research Network. https://papers.ssrn.com/abstract=3748268

Doshi-Velez, F., & Kim, B. (2017). Towards A Rigorous Science of Interpretable Machine Learning. ArXiv:1702.08608 [Cs, Stat].

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255–260.

Miller, T. (2017). Explanation in Artificial Intelligence: Insights from the Social Sciences. ArXiv:1706.07269 [Cs].

Mitchell, T. M. (1997). Machine Learning (1 edition). McGraw-Hill Education.

Russell, S., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach. Upper Saddle River, NJ, USA: Prentice Hall Press.