

MODY UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF ENGINEERING AND TECHNOLOGY



TECHNICAL SEMINAR PRESENTATION

CS-14.474

SUBMITTED TO –

Dr. Sanjeev Patwa

Dr. Prateek Bhanti

SUBMITTED BY –

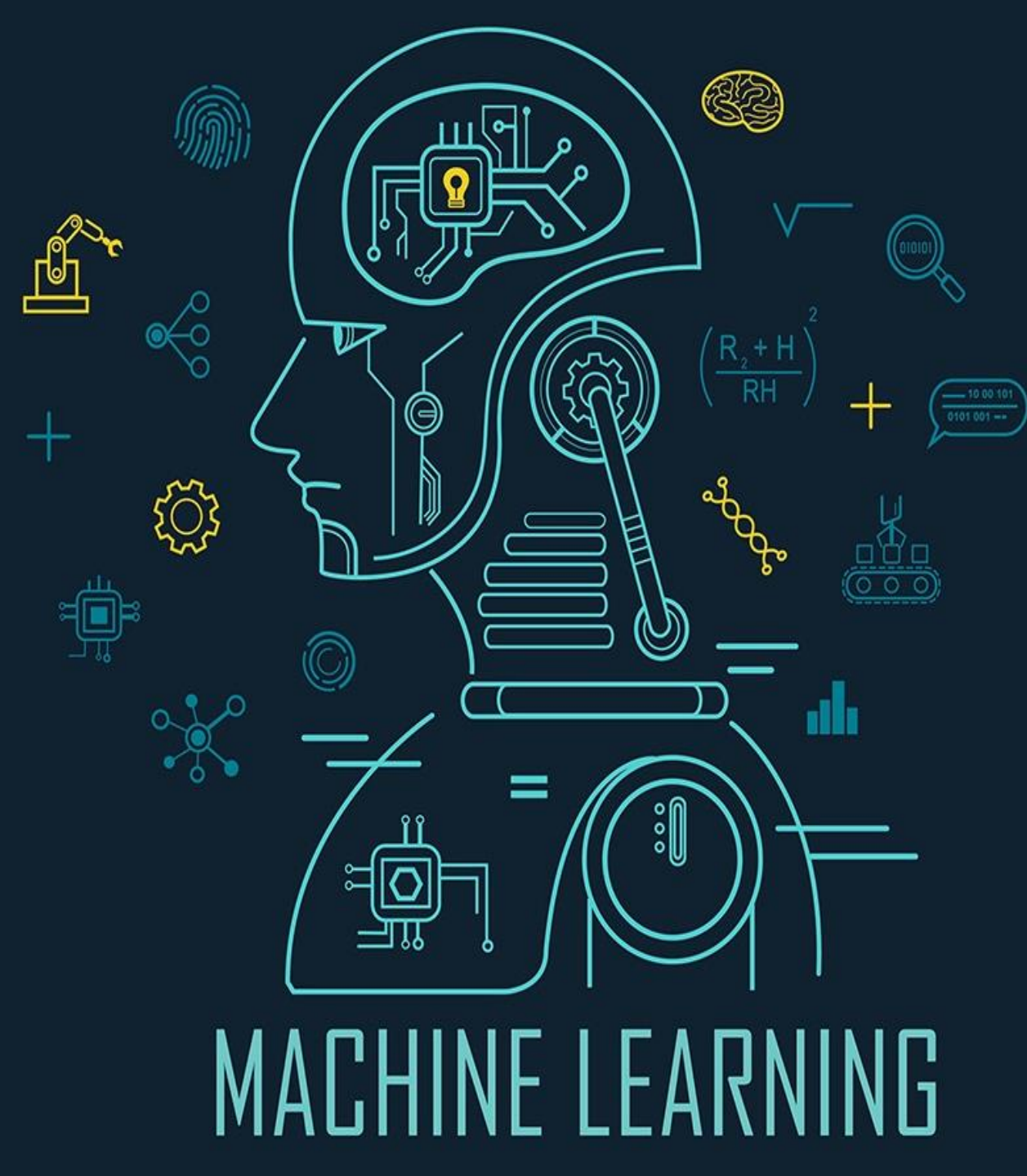
Akshita Kanther(180161)

Hiral Jain(180167)



MACHINE LEARNING (ML) - ALGORITHMS AND APPLICATIONS



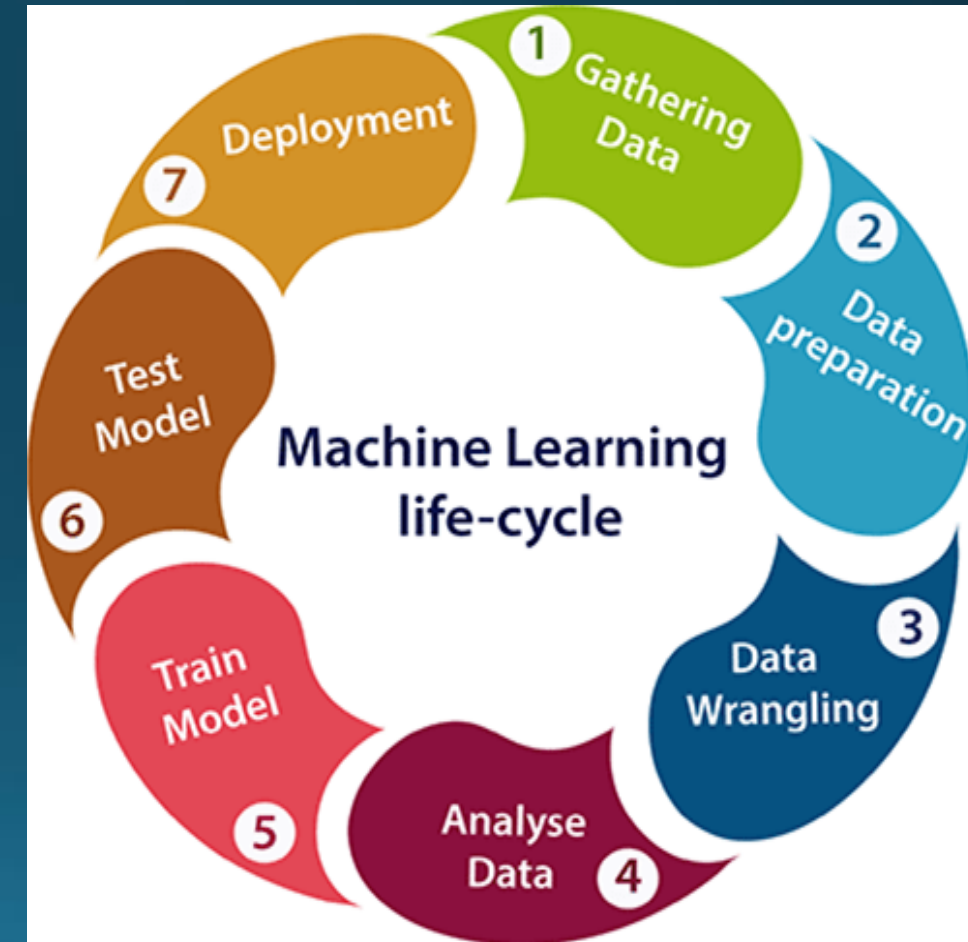


INTRODUCTION

- Machine Learning(ML) is a branch of computer science where we develop algorithms that make a machine learn to do something without actually making computations about it.
- The basic premise of machine learning is to build algorithms that can receive input data and use statistical analysis to predict an output while updating outputs as new data becomes available.
- The primary goal of ML is to observe the given problem or a dataset and determine the hidden patterns and relationships and extract details/information which helps in data-driven decision making and provides insights.
- Machine learning is an application of artificial intelligence that involves algorithms and data that automatically analyze and make decision by itself without human intervention.

ML LIFECYCLE

1. Data Gathering is the first step of the machine learning life cycle. The goal of this step is to identify and obtain all data-related problems.
2. After collecting the data, we need to prepare it for further steps. Data preparation is a step where we put our data into a suitable place and prepare it to use in our machine learning training.
3. Data wrangling is the process of cleaning and converting raw data into a useable format. It is the process of cleaning the data, selecting the variable to use, and transforming the data in a proper format to make it more suitable for analysis in the next step.
4. The aim of this step is to build a machine learning model to analyze the data using various analytical techniques and review the outcome
5. Now the next step is to train the model, in this step we train our model to improve its performance for better outcome of the problem.
6. Once our machine learning model has been trained on a given dataset, then we test the model. In this step, we check for the accuracy of our model by providing a test dataset to it. Testing the model determines the percentage accuracy of the model as per the requirement of project or problem.
7. The last step of machine learning life cycle is deployment, where we deploy the model in the real-world system.

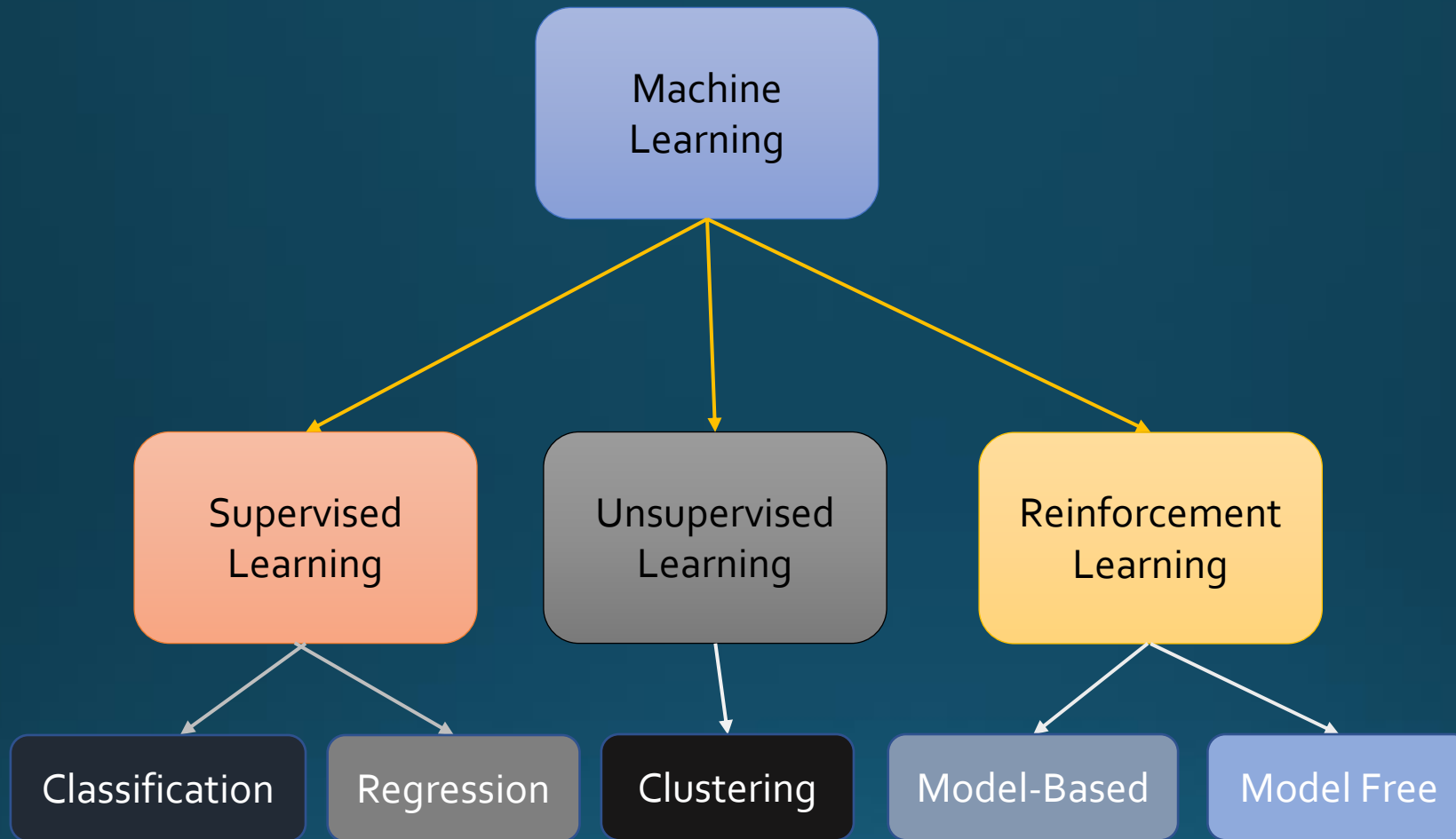


TYPES OF MACHINE LEARNING

Machine learning can be divided into three main categories:-

- **Supervised learning** -Supervised learning is a type of machine learning that uses labeled data to train machine learning models. In labeled data, the output is already known. The model just needs to map the inputs to the respective outputs.
- **Unsupervised learning** - Unsupervised learning is a type of machine learning that uses unlabeled data to train machines. Unlabeled data doesn't have a fixed output variable. The model learns from the data, discovers the patterns and features in the data, and returns the output.
- **Reinforcement learning** - Reinforcement Learning trains a machine to take suitable actions and maximize its rewards in a particular situation. It uses an agent and an environment to produce actions and rewards. The agent has a start and an end state. But, there might be different paths for reaching the end state, like a maze. In this learning technique, there is no predefined target variable.





- Support Vector Machine(SVM)
- Naïve Bayes
- Random Forest
- Decision Tree
- Linear Regression
- Logistic Regression

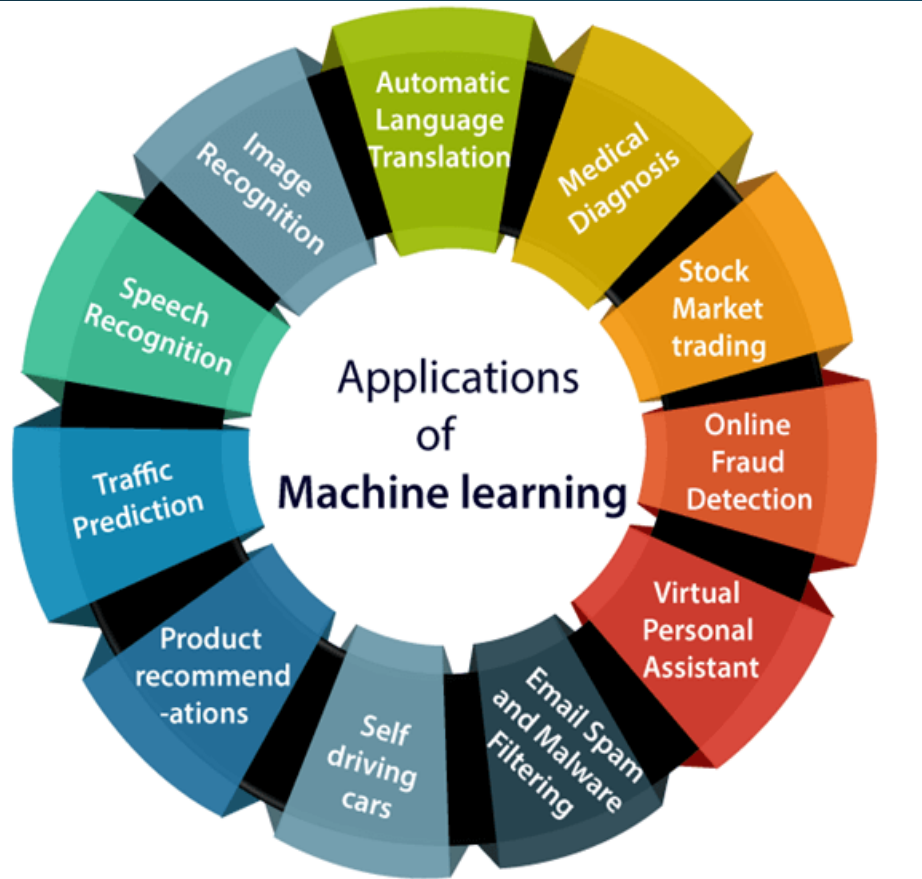
- K-means
- Apriori

ML ALGORITHMS

1. **Naive Bayes Classifier (NBC)** - The Naïve Bayes classifier is based on Bayes' theorem and classifies every value as independent of any other value. It allows us to predict a class/category, based on a given set of features, using probability.
2. **K-Means Clustering** - The K Means Clustering algorithm is a type of unsupervised learning, which is used to categorize unlabelled data, i.e. data without defined categories or groups. The algorithm works by finding groups within the data, with the number of groups represented by the variable K. It then works iteratively to assign each data point to one of K groups based on the features provided.
3. **Support Vector Machine (SVM)** - Support Vector Machine algorithms are supervised learning models that analyze data used for classification and regression analysis. They essentially filter data into categories, which is achieved by providing a set of training examples, each set marked as belonging to one or the other of the two categories. The algorithm then works to build a model that assigns new values to one category or the other.
4. **Apriori Algorithm** - The Apriori machine learning algorithm is an unsupervised algorithm used frequently to sort information into categories. depicts association rules from the input data set. Association rule entails that on occurrence of item 'X', item 'Y' happens definitely, similar to "if-then" rule. For instance, ***IF*** a person purchases bike, ***THEN*** certainly they will purchase helmet.

5. **Linear Regression (LR)** - Linear regression is the most basic type of regression. Simple linear regression allows us to understand the relationships between two continuous variables.
6. **Logistic Regression** - Logistic regression focuses on estimating the probability of an event occurring based on the previous data provided. It is used to cover a binary dependent variable, that is where only two values, 0 and 1, represent outcomes.
7. **Decision Tree** - A decision tree is a flow-chart-like tree structure that uses a branching method to illustrate every possible outcome of a decision. Each node within the tree represents a test on a specific variable and each branch is the outcome of that test.
8. **Random Forest** - Random forests or 'random decision forests' is an ensemble learning method, combining multiple algorithms to generate better results for classification, regression and other tasks. Each individual classifier is weak, but when combined with others, can produce excellent results.

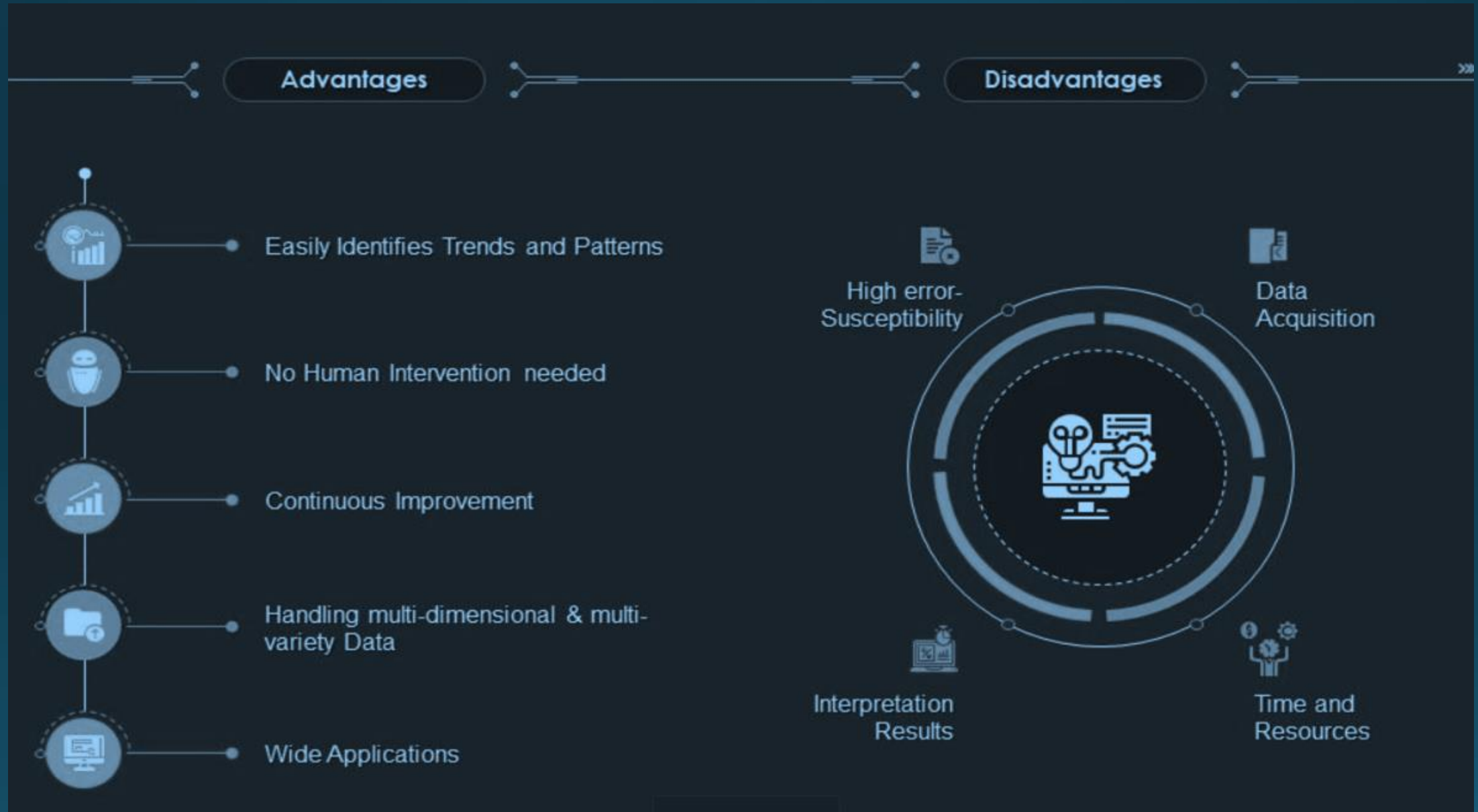
ML APPLICATIONS



1. **Spam Detection:** Received emails are identified whether they are spam if yes then they are not shown to the user with regular emails in their inboxes. All these spam emails are maintained in a separate folder. This is done by learning how to recognize a spam mail by identifying the characteristics from newly received ones.
2. **Face Detection:** Identifying faces from a given number of photos and tagging them automatically if next time the same face is detected for a new photo. Currently, Facebook is popularly known for using this technique where when we upload a picture some suggestions of friends are automatically detected and shown to the user whose friends are in the picture. For example, Google photos separate all the pictures according to people in it.
3. **Credit Card Fraud Detection:** According to customers past transactions, if there are any inappropriate purchases made then the customer is warned immediately about the condition.
4. **Digit Recognition:** Machine's camera detects postal codes that are handwritten and arranges all the letters according to the geographical locations they have to reach. The machine is trained to learn handwritten numbers and transforms into digital signs.
5. **Speech Understanding:** Deals with listening to a speech by the user, understand user intentions process what machine has understood. Machine is expected to follow instructions from the users accurately. "Cortana" by Windows, "Siri" by Apple and "Okay Google" by Google are popular and successfully implemented applications of this technique.

6. **Product Recommendation:** With all the data of a customer's past purchases or interests online, the machine will recommend some products which would attract customers to view them and maybe even purchased. Flipkart, Amazon and many other e-commerce sites have been implanting this technique where we receive only recommended products as advertisements.
7. **Medical Diagnosis:** For detecting diseases more accurately, hospitals these days are using machines that could decide whether a person is affected with any diseases for symptoms he/her has, with the help of complete data about all diseases and symptoms. IBM designed a system with 95% precision in predicting the cancerous images in contrast to 75%-84% precision by doctors.
8. **Customer Segmentation:** With the help of past behaviour patterns, the ML algorithms tries to predict the number of users who will choose the paid versions from the trial versions. Amazon Prime has implemented this technique. Chat smarter with Allo: Allo is a messenger application by Google which will learn from its users, how they are responding to what kind of messages and recommend the user with a response that he /she has thought of typing.
9. **Financial Service:** Companies can identify the company insights of financial sector data and can overcome the occurrence of financial fraud. It is used to identify the opportunities for investment and trade. We can also prevent the financial risks prone institutions by using cyber surveillance and take necessary actions to prevent fraud.
10. **Transportation:** By the travel history and pattern of travelling in various routes, machine learning helps in Transportation Company to predict the problem in routers and advice their customers to choose a different route. Transportation firmly uses machine language to carry out data analysis and data modelling.

CHALLENGES AND LIMITATIONS OF MACHINE LEARNING



A hand is shown in the lower-left foreground, pointing its index finger towards a glowing, rectangular digital interface. The interface is set against a dark blue background filled with numerous small, bright, out-of-focus light points, some of which are blue and others white. The interface itself has a bright, yellowish-white glow emanating from its center, suggesting a selection or activation point. The overall aesthetic is futuristic and technological.

CONCLUSION

In this paper, we have conducted a comprehensive overview of machine learning algorithms and applications. According to our goal, we have briefly discussed how various types of machine learning methods can be used for making solutions to various real-world issues. A successful machine learning model depends on both the data and the performance of the learning algorithms. We also discussed several popular application areas based on machine learning techniques to highlight their applicability in various real-world issues. Finally, we have summarized and discussed the challenges faced and the potential research opportunities and future directions in the area. Therefore, the challenges that are identified create promising research opportunities in the field which must be addressed with effective solutions in various application areas. Overall, we believe that our study on machine learning-based solutions opens up a promising direction and can be used as a reference guide for potential research and applications for both academia and industry professionals as well as for decision-makers, from a technical point of view.

REFERENCES

- [1] https://www.researchgate.net/publication/320609700_A_Review_on_Machine_Learning_Algorithms_Tasks_and_Applications
- [2] <https://www.ijitee.org/wp-content/uploads/papers/v8i11S2/K106909811S219.pdf>
- [3] Top 10 Machine Learning Algorithms, <https://www.dezyre.com/article/top-10-machinelearning-algorithms/202>
- [4] Alpaydin, E. Introduction to Machine Learning (The MIT Press, Cambridge, MA, 2014).
- [5] Singh, S., Kumar, N. and Kaur, N., 2014. Design Anddevelopment Of Rfid Based Intelligent Security System. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume, 3.
- [6] Xin-She Yang, Su Fong Chien, and Tiew On Ting, "Computational Intelligence and Metaheuristic Algorithms with Applications", The Scientific World Journal, Vol. 2014, pp. 1-4, 2014
- [7] Kumar, N. and Gupta, S., 2016. Offline Handwritten Gurmukhi Character Recognition: A Review. International Journal of Software Engineering and Its Applications, 10(5), pp.77-86.
- [8] Introduction to Natural Language Processing (NLP): What is NLP?, <https://blog.algorithmia.com/introduction-natural-language-processing-nlp/>
- [9] What are some interesting possible applications of machine learning? <https://www.quora.com/What-are-some-interesting-possible-applications-of-machinelearning>
- [10] What are some real-world examples of applications of machine learning in the field? <https://www.quora.com/What-are-some-real-worldexamples-of-applications-of-machine-learning-in-thefield>
- [11] Sharma, D., Pabby, G. and Kumar, N., Challenges Involved in Big Data Processing & Methods to Solve Big Data Processing Problems. IJRASET, 5(8), pp.841-844.