Dots of Machine Learning and Applications

Annapurna Neeralgi[1] and B. Fakrudin[2]
[1]Assistant Professor of Computer Science
[2]Professor of Plant Biotechnology & Dean
DSLD College of Horticultural Engineering and Food Technology,
University of Horticultural Science, Devihosur-Haveri, Karnataka, India PIN-581110
*Corresponding Author: anuneeralgi10@gmail.com

In a rapidly changing world, machine learning (ML) has become a great resource to effectively convert data into knowledge and assist human in making intelligent actionable decisions. This shift is revolutionizing how we interact with the world, make decisions, and solve complex problems.

Google search engine collates related information from across the internet and provides it against a query. This is possible because the 'Google search' uses machine learning. Facial recognition on Instagram and Facebook is being done with the help of ML algorithms, and helps us to tag our friends automatically, when we buy a product through the online marketplace, depending on our purchase we get other related recommendations, these recommendations are through ML. Apart from this, ML plays a very important role in health care to detect and diagnose various diseases, in weather predictions, in the financial sector in fraud detection, algorithmic trading, in the transport sector to predict traffic conditions, self-navigation, autopilots etc., Agriculture sector price predictions, disease and pest predictions.

What is Machine Learning?

Machine learning (ML) is a subset of artificial intelligence (AI) that involves the development of algorithms and statistical models that enable computers to perform tasks without explicit instructions. Instead, these systems learn from and make decisions based on data. The primary goal of machine learning is to allow computers to learn from experience, identify patterns, and make predictions or decisions with minimal human intervention.

Artificial intelligence (AI) is the field of computer science that focuses on creating systems capable of performing tasks that typically require human intelligence. These tasks include reasoning, learning, problem-solving, perception, language understanding, and decision-making. ML is a subset of Artificial Intelligence; it uses various algorithms that learn from data to make predictions or decisions. Deep Learning (DL), a further subset of ML, uses neural networks with many layers (hence "deep") to analyse various factors of data. Further, ML is an application of artificial intelligence that provides systems with the ability to learn on their own and improve from experience without being programmed externally.

Understanding Machine Learning

Data like text, numbers and images are the key factors for ML. The ML uses this data to teach the computer to understand the different patterns in the given data and make decisions, this is done with the help of different machine learning algorithms, for different specific tasks. There are three different types of Machine Learning;

Supervised Learning: This is like teaching a kid to recognize something by showing flash cards. Here, the computer is given examples called training data with the correct answers called labels.

Example: Bunch of photos of tomatoes and apples with each of them labelled as "tomato" and "apple". The computer learns to recognize the pattern that distinguishes tomato and apple.

The important algorithms used are Linear Regression, Decision Tree, Random Forest, Support Vector Machine, K-Nearest Neighbour, etc.,

Unsupervised Learning: Here, the data set is without the label, the ML uses an algorithm to group them into different clusters based on the patterns and features. Bunch of photos of animals without labels, the computer might group them into clusters based on similarities, like all cats in one group and all dogs in another. The important algorithms used are KMeans Clustering, Principal Component Analysis (PCA)

Reinforcement Learning: Here the computer learns on its own. The algorithm learns by interacting with its environment, receiving rewards for desirable actions. AI playing games like chess, PUBG, and self-driving cars. The important algorithms used are Q-Learning, Deep Q-Networks (DQN)

How Machines Learn

Features are the important pieces of information from the data. For example, in animal photos, features could be the shape of ears or the colour of fur, in fruit photos the shape and colour are

the features. The machine learning models are the mathematical formulas or set of rules, that computers use to make decisions/ predictions based on the features in the data. Training is the process of teaching the model using a dataset. The computer adjusts its model to better match the data. During the training process, the computer understands the patterns/similarities in the given data. The knowledge gained during the training process is then applied to the unseen data for predictions during the testing process.

The steps involved in ML-Algorithm

- **Collect Data**: Gather the data needed, like photos, text, or numbers.
- **Preprocess Data**: Clean the data and choose the important features.
- **Choose a Model**: Pick an appropriate ML algorithm based on the task (e.g., supervised learning for labelled data).
- **Train the Model**: Teach the model using training data set.
- **Test the Model**: Evaluate the model with new data to ensure it works well.
- **Make Predictions**: Use the model to make predictions or decisions on new data.

Step-by-step example with Fruits to predict whether the given fruit is an apple or orange

- **Collect Data:** Take 100 pictures of apples and 100 pictures of oranges.
- **Preprocess Data:** Resize all pictures to 100x100 pixels.
- **Extract Features:** Identify features like color and shape.
- **Choose a Model:** Select a decision tree model.
- **Train the Model:** Show the model 80 pictures of apples and 80 pictures of oranges, telling it which is which.
- **Test the Model:** Show the model 20 new pictures of apples and 20 new pictures of oranges without labels and see if it can identify them correctly.
- **Make Predictions:** Use the trained model to identify new pictures of fruits.

How it works

Patterns and Rules: The model learns patterns and rules from the training data. For example, it will learn that round, orange-coloured fruits are likely to be oranges and round-shaped red colour fruits are apples.

Adjusting: During training, the model adjusts its internal rules to better match the training data. It learns from mistakes and tries to improve.

Generalizing: A good model can generalize from the examples it has seen to new, unseen examples. This means it can accurately identify new apples and oranges based on what it learnt.

By following these steps, machines learn from data and can make intelligent decisions or predictions, much like how humans learn from experience.

Use of ML across Industries:

This capability makes ML a powerful tool for tackling complex problems and driving advancements in multiple industries. This technology is crucial for enhancing efficiency, accuracy, and innovation across various sectors. In this article, we will discuss some of the industries which are making the best use of ML.

Healthcare: Of the many things machine learning can do, patient care and medical research are two fields that could greatly benefit from ML in healthcare. Algorithms can be used to examine medical images and find diseases such as cancer in their initial stages, significantly increasing the chances of recovery. Additionally, ML is used in personalized medicine to model how patients will respond to different treatments based on their genetic information and medical history.

Disease Prediction and Diagnosis:

- ***Cancer Detection:*** ML models can analyze medical imaging data (e.g., radiology images) to detect cancer and other diseases at early stages.
- ***Predicting Disease Progression:*** Using patient data to predict the course of diseases and personalize treatment plans.
- ***Histopathology:*** Analysing tissue samples to detect abnormalities and diagnose diseases.
- ***Medical Imaging:*** Enhancing the analysis of X-rays, MRIs, and CT scans.

Finance: ML tools are security and decision-making enhancers. Detecting fraud: ML-based systems can detect things like unusual transaction patterns, which are a sign of potentially fraudulent activity. ML algorithms are also used to help improve credit scoring, risk modelling, and algorithmic trading by processing larger amounts of data to create more precise predictions.

Retail: For retail - ML revolves around understanding consumer behaviour and choices. These systems also provide recommendations based on your purchase history and browsing activities that are right for you as a customer.

Food Industry: ML in food industries plays a crucial role in improving various processes, and enhancing product quality and food safety. ML algorithms are being used for inspecting and grading food products based on quality parameters like colour, shape, size and defects in the products. For example: Image processing can be used for finding defective fruits and vegetables. Predictive maintenance in food processing plants can prevent equipment failures, reducing downtime and ensuring continuous production. Machine learning models are employed to detect contaminants and pathogens in food products, ensuring compliance with safety standards.

Space Research: Machine learning (ML) is playing a significant role in advancing space research by enabling more efficient data analysis, improving the accuracy of predictions, and facilitating the discovery of new insights. ML algorithms are used to process and analyze large volumes of astronomical data from telescopes and space missions, helping to identify celestial objects and phenomena. ML techniques are used to enhance and interpret satellite images, enabling better monitoring of Earth's environment and climate.

Smart Farming

Smart farming is being revolutionized by machine learning (ML), innovative solutions guarantee sustainability, enhance productivity, and optimize agricultural operations. machine learning algorithms evaluate data from sensors, drones, and satellite imagery and accurate suggestions on irrigation, fertilization, and pest management are provided. Various Image processing techniques and sensory data are fed to ML models to diagnose the pests, diseases and nutrient deficiencies in the crops. The historical data analysis is helping to predict yield, market demand etc.

Accurate weather forecasts and climate analysis are provided by ML models, which assist farmers in risk mitigation and agricultural activity planning. ML-driven applications are being used to optimise resources like water, moisture, soil, fertilizers, pesticides etc., and help the farmer reduce the cost of cultivation. Autonomous tractors and harvesters equipped with AI and machine learning algorithms can perform tasks with high precision and efficiency, reducing the need for manual labour.

Machine learning (ML) has numerous applications in the biological sciences, providing powerful tools for analysing complex biological data, predicting outcomes, and uncovering patterns that traditional methods may not easily detect. Some key areas where machine learning is making significant impacts in biology:

Genomics and Genetics

- **Gene Expression Analysis**

Prediction of Gene Function: ML models can predict the function of genes based on patterns in gene expression data.

Differential Expression Analysis: Identifying genes that are differentially expressed under different conditions (e.g., healthy vs. diseased states).

- **Variant Calling and Genome Annotation**

Variant Detection: Machine learning algorithms can improve the accuracy of detecting genetic variants from sequencing data.

Functional Annotation: ML can be used to predict the functional impact of genetic variants, such as whether a mutation is likely to be benign or pathogenic.

Proteomics

- **Protein Structure Prediction**

AlphaFold: An AI system developed by DeepMind that predicts the 3D structure of proteins with high accuracy.

Predicting Protein-Protein Interactions: ML models can predict interactions between proteins, which is crucial for understanding cellular processes.

- **Protein Function Prediction**

Using sequence data and structural information, ML models can predict the function of unknown proteins.

Drug Discovery and Development

- **Drug Screening and Design**

Virtual Screening: ML algorithms can screen vast libraries of compounds to identify potential drug candidates.

De Novo Drug Design: Generative models can design new molecules with desired properties.

- **Predicting Drug-Target Interactions:**

Machine learning models can predict how different drugs will interact with various biological targets.

Systems Biology

- **Modelling Biological Networks:**

Gene Regulatory Networks: ML can infer networks that describe the regulatory relationships between genes.

Metabolic Networks: Predicting metabolic pathways and their interactions.

- **Single-Cell Analysis:**

Clustering and Classification: Analysing single-cell RNA-seq data to classify cell types and states.

Trajectory Inference: Mapping the developmental trajectories of cells.

Evolutionary Biology

- **Phylogenetics**

Tree Construction: ML methods can help construct and analyse phylogenetic trees to study evolutionary relationships.

Predicting Evolutionary Trends: Understanding and predicting how organisms will evolve.

Environmental Biology

- **Ecological Modelling**

Species Distribution Models: Predicting the distribution of species based on environmental factors.

Ecosystem Dynamics: Modelling interactions within ecosystems and predicting changes due to environmental pressures.

- **Microbiome Analysis**

Community Profiling: Analysing the composition and function of microbial communities in different environments.

Host-Microbe Interactions: Studying how microbes interact with their hosts and affect health.

Challenges

While machine learning offers tremendous potential, there are challenges in its application to biological data, such as:

***Data Quality and Quantity*:** Biological data can be noisy, incomplete, and heterogeneous.

Interpretability: ML models, especially deep learning models, can be difficult to interpret, which is crucial for biological applications.

Integration of Multimodal Data: Combining different types of biological data (e.g., genomic, proteomic, clinical) remains a significant challenge.

Conclusion

Machine learning is revolutionizing the biological sciences, offering new ways to understand complex biological systems and develop innovative solutions to health and environmental challenges. At the forefront of a data-driven revolution, machine learning bridges the divide between big data and intelligence. It has a significant impact on daily living, industry transformation, and global challenge resolution. To secure a future where technology and mankind coexist peacefully, we must comprehend machine learning's foundations, possibilities, and ethical implications as we continue to harness its power.

References

Rahul Rai, Manoj Kumar Tiwari, Dmitry Ivanov & Alexandre Dolgui, *Machine learning in manufacturing and industry 4.0 applications*, International Journal of Production Research, Pages 4773-4778 ,14 Aug 2021

Kazuma Uesaka, Hiroya Oka , Ryuji Kato , Kei Kanie , Takaaki Kojima , Hiroshi Tsugawa, Yosuke T oda, Takaaki Horinouchi , Bioinformatics in bioscience and bioengineering: Recent advances, applications, and perspectives, Journal of Bioscience and Bioengineering Volume 134, Issue 5, November 2022, Pages 363-373

Iqbal H. Sarker, Machine Learning: Algorithms, Real-World Applications and Research Directions, SN Computer Science (2021) 2:160 https://doi.org/10.1007/s42979-021-00592-x SN Computer Science, Review Article, 22, March 2021

Raffaele Cioffi, Marta Travaglioni, Giuseppina Piscitelli, Antonella Petrillo,and Fabio De Felice,Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions, Sustainability 2020, 12, 492; doi:10.3390/su12020492

Pinky Gupta, Research Paper on Machine Learning and its application, International Research Journal of Engineering and Technology (IRJET), Volume: 09 Issue: 03, Mar 2022, 1483: 1486.

* * * * * * * *