Classifications Model In Machine Learning Used In Agritech: A Literature Review

Dr. Syed Tabrez Hassan[1*],

[1*]Associate Professor School of Business and economics, Adamas University, Kolkata

***Correspondent Author:**Dr. Syed Tabrez Hassan,
[*]Associate Professor School of Business and economics, Adamas University, Kolkata

1.Introduction:

The advancement of the agricultural industry and its ecosystem with information technology have been instrumental in driving agritech innovation. India, a country with a population exceeding a billion, has witnessed a remarkable transformation in its agricultural landscape with the advent of agritech. This sector encompasses a wide array of services, from crop management to satellite imagery for monitoring crop health, soil conditions, and pest infestations. The convergence of technology and agriculture has not only enhanced productivity but has also help farmers adapt to climate change and strengthen global food security.

Agritech in India has emerged as a transformative force revolutionizing the country's agricultural landscape. With its vast agrarian economy and a large population dependent on farming for livelihoods, India stands at the forefront of adopting technological innovations to address various challenges faced by the agricultural sector. Agritech solutions in India span a wide range of applications, from precision agriculture and crop management to supply chain optimization and market access. One of the key areas of focus in Indian agritech is leveraging digital technologies and data analytics to enable precision farming practices. This includes the use of remote sensing, satellite imagery, and IoT sensors to monitor crop health, soil moisture levels, and weather patterns, allowing farmers to make informed decisions regarding irrigation, fertilization, and pest management. Additionally, farm management software and mobile applications empower farmers with real-time information and advisory services, enhancing productivity and profitability. Agritech based ERP platforms facilitate market access for farmers by connecting them directly with buyers, eliminating middlemen and ensuring fair prices for their produce. These platforms also provide access to financial services, insurance products, and agricultural inputs, thereby addressing critical challenges related to access to credit and risk mitigation in farming communities.

1.1 Key uses of Agritech:

The use of agritech solutions in agroforestry systems and sustainable land management were studied by Wu et al., (2019), and they suggested technologies such as remote sensing, GIS mapping, and agroecological modeling facilitate agroforestry planning, biodiversity conservation, and carbon sequestration, thus promoting resilience and environmental sustainability in agroecosystems. Gonzalez et al.,(2022), explored the integration of agritech solutions in small-scale farming operations. Their review discusses how technologies such as mobile apps, low-cost sensors, and community-based platforms empower smallholder farmers by providing access to market information, weather forecasts, and advisory services, thus enhancing their livelihoods and resilience. Nguyen et al., (2020), studied the applications of agritech in aquaculture and fisheries management. Their review discusses how technologies such as IoT sensors, satellite imagery, and predictive analytics improve water quality monitoring, fish tracking, and aquaculture productivity, thus supporting sustainable management practices and livelihoods in the fisheries sector. The role of agritech in urban agriculture and rooftop farming was explored by Chang et al., (2021). Their review examines how technologies such as hydroponics, vertical farming systems, and automated greenhouse controls enable urban dwellers to grow fresh produce in limited spaces, contributing to food security, environmental sustainability, and community resilience. The key learning can be summarized as:

Smart Farming Systems: Utilizing data-driven approaches such as GPS, sensors, and drones to optimize farming practices, including planting, irrigation, fertilization, and pest management. Crop Monitoring and Management: Employing remote sensing technologies and data analytics to monitor crop health, growth, and yield, enabling timely interventions and maximizing productivity. Implementing IoT (Internet of Things) devices and automation to streamline farm operations, improve efficiency, and reduce labor requirements.

Vertical Farming and Indoor Agriculture: Utilizing controlled environment agriculture techniques such as hydroponics and aeroponics to grow crops indoors, enabling year-round production and resource efficiency.

Climate Resilience and Sustainability: Developing sustainable agricultural practices and technologies to mitigate climate change impacts, conserve natural resources, and promote environmental stewardship. Developing genetically modified organisms (GMOs) and biotech solutions to enhance crop traits, resistance to pests and diseases, and nutritional content.

Livestock Management: Implementing IoT sensors, wearable devices, and data analytics for monitoring livestock health, nutrition, and productivity, optimizing animal welfare and farm profitability. It provides agricultural training, extension services, and knowledge sharing platforms to empower farmers with best practices, innovations, and market insights.

The core of agritech lies in the application of innovative technology within the agricultural domain. The state-of-the-art technology has the potential to significantly transform agricultural services, reducing production costs and enhancing both convenience and security. The rapid evolution of agritech is reshaping the traditional agricultural operational paradigm, as the amalgamation of new technologies and farming systems contributes to the advancement of global economic development. The continual expansion of the agritech market is a direct outcome of investments in research and development. The agritech is leveraging emerging technologies such as big data, the Internet of Things, and artificial intelligence. This strategic utilization aims to drive reform within the agricultural sector, enhance the operational efficiency of the financial system, and concurrently manage associated risks.

2. Machine Learning based classifications in agricultural technology:

Machine learning-based classifications in agricultural technology have garnered significant attention in recent years, with numerous studies exploring their potential applications and implications. Smith et al., (2018), conducted a comprehensive review of machine learning approaches for crop classification and yield prediction, highlighting the versatility and effectiveness of these techniques. Similarly, Osorio et al., (2019), systematically evaluated the application of machine learning algorithms in weed detection and management, emphasizing their role in promoting sustainable agriculture. Tian et al., (2020), provided insights into machine learning-based soil classification for precision agriculture, discussing the implications for soil management practices.

Gupta et al., (2021), reviewed machine learning techniques for pest and disease detection in crops, shedding light on challenges and future research directions. Moreover, Patel et al., (2019), explored the integration of machine learning techniques in agricultural robotics, showcasing their potential to enhance efficiency and productivity in farming operations. Li et al., (2020), summarized machine learning applications in irrigation scheduling for precision agriculture, highlighting their contribution to water conservation and crop yield optimization. Similarly, Kim et al., (2018), evaluated machine learning approaches for crop disease identification and classification, discussing their accuracy and scalability. These studies collectively underscore the diverse applications of machine learning-based classifications in agricultural technology, from crop monitoring and management to pest detection, soil classification, and beyond, paving the way for sustainable and efficient farming practices as following:

Crop Monitoring and Management: Machine learning facilitates precise crop identification and health assessment through advanced image analysis techniques, enabling farmers to optimize harvesting schedules and predict yields accurately.

Pest and Disease Detection: Early detection of pests and diseases is achieved through machine learning-powered image recognition systems, allowing for targeted interventions and reducing the need for excessive pesticide use, thus promoting sustainable pest management practices.

Weed Identification and Management: Machine learning algorithms distinguish between crops and weeds with high accuracy, guiding farmers in applying herbicides only where necessary, thereby minimizing environmental impact and optimizing weed control efforts.

Soil Classification and Management: Soil classification based on machine learning analysis aids in tailoring soil amendments and tillage practices to suit specific soil types, ensuring optimal nutrient availability and enhancing overall soil health and fertility.

Crop Rotation and Planning: Predictive analytics utilize machine learning algorithms to recommend crop rotation sequences that maintain soil fertility, suppress pests and diseases, and maximize crop yields over successive growing seasons.

Water Management: Machine learning predicts soil moisture levels and irrigation requirements, enabling farmers to apply water precisely where and when needed, conserving water resources and mitigating water stress in crops.

Drought and Flood Prediction: Machine learning models forecast droughts and floods, empowering farmers with early warnings to implement timely mitigation strategies and minimize crop losses, enhancing resilience in the face of climate variability.

Quality Grading and Sorting: Automated grading and sorting systems driven by machine learning ensure consistent quality standards for agricultural produce, enhancing marketability and value perception among consumers.

Market Demand Forecasting: Data-driven analysis of market trends and consumer preferences using machine learning techniques enables accurate forecasting of demand for agricultural products, guiding farmers in production planning and market positioning strategies.

Precision Farming and Decision Support: Personalized recommendations derived from machine learning insights optimize resource allocation, minimizing input wastage and environmental impact while maximizing productivity and profitability in agricultural operations.

3. Machine learning based classification models:

Support Vector Machine:

The Support Vector Machine (SVM) belongs to the category of general feedforward neural networks. Its fundamental model identifies the optimal separation hyperplane within the feature space, aiming to maximize the margin between positive and negative samples in the training set. The classification of a sample as linearly separable or linearly inseparable depends on the unique characteristics exhibited by the sample, as illustrated in Figure 1.

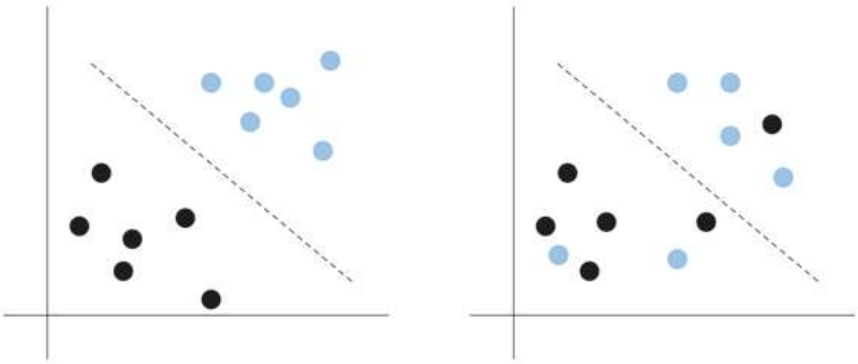

Fig 1. Support Vector Machine

Linearly separable denotes the ability to distinguish between two types of samples using a straight line in a two-dimensional space. Consider a training sample {(xi, di)}, where xi represents the ith input pattern sample, and di is the corresponding expected response (target output). Initially, it is assumed that the classes represented by the subset di = +1 and di = −1 are linearly separable. The hyperplane decision surface equation for separation is expressed as follows:

f(x) = $\omega^T x + b$

where x is the input vector, ω is the adjustable weight vector, and b is the offset. In the context of a given weight vector ω and offset b, SVM aims to discover the hyperplane that maximizes the separation margin. Under these conditions, the resulting decision surface is referred to as the optimal hyperplane. Extending into a high-dimensional space involves employing a high-dimensional function for separation. In cases where linear separation is not feasible, termed linearly inseparable, sample features are mapped to a higher-dimensional space using a Gaussian kernel function. This transformation ensures that nonlinear features are converted into linearly separable features, allowing the sample to undergo processing using a linearly separable approach. SVM finds applications in various supervised learning algorithms, encompassing classification, regression, and anomaly detection. It boasts numerous advantages over traditional machine learning algorithms, including high efficiency in high-dimensional space—particularly in scenarios where the data dimension exceeds the number of samples. SVM's ability to utilize a subset of the training set in the decision function, commonly referred to as a support vector, is advantageous for efficient utilization of computer memory.

However, SVM does pose certain challenges, especially in classification tasks. For instance, when the number of features significantly exceeds the number of samples, there's a susceptibility to overfitting during kernel function selection for training. Mitigating this challenge often involves employing regularization and other strategies to address overfitting issues.

K-Nearest Neighbours

First introduced by Cover and Hart in 1968, KNN (K-Nearest Neighbours) has evolved into a well-established machine learning algorithm, recognized for its simplicity and maturity in both theory and research. The operational concept of KNN is straightforward: if the majority of the k-nearest samples in the feature space belong to a specific category, the sample under consideration should also be assigned to that category. KNN makes classification decisions for samples based solely on the category of their nearest one or several neighbours.

However, a notable drawback of KNN is its substantial computational burden. Calculating the distance of each text to be classified from all known samples is necessary to determine the text’s k-nearest neighbour points. To address this challenge, a common solution involves pre-processing known sample points, selectively removing samples with minimal impact on classification. The reverse KNN method proves effective in reducing the computational complexity of the algorithm, thereby enhancing classification efficiency.

It is worth noting that the suitability of the KNN algorithm is particularly pronounced in text classification scenarios with large sample sizes. In cases where the sample size is small, the algorithm may be more prone to classification errors.

Naïve-Bayes

Utilizing theories of probability and statistics, Bayesian methods are employed to classify sample datasets. Thanks to their robust mathematical foundation, Bayesian algorithms exhibit higher accuracy and lower error rates, particularly when applied to large datasets. These methods adeptly integrate prior probability and posterior probability, mitigating the inherent subjective bias associated with relying solely on prior probability.

Within the realm of supervised learning algorithms, Naive Bayes methods stand out. Grounded in the Bayes theorem, these algorithms make a simplifying assumption that each pair of features is independent of the others. For a given class 'y' and eigenvectors 'x1' to 'x2' related to 'N', the naive assumption is applied, considering each pair of features as mutually independent.

$\mathbf{P(x_1|y, x_1,\ldots x_{i-1},x_{i+1},\ldots x_n) = P(x_i|y)}$

The estimation of $P(x_i |y)$, representing the relative frequency of class 'y' in the training set, can be achieved using the maximum posterior probability. In the context of various naive Bayes classifiers, differences arise primarily from the assumptions made when handling $P(x_i|y)$ distributions. Despite the simplicity of the assumptions in the naive Bayes model, its performance is noteworthy in practical classification tasks. Moreover, the model requires only a random training sample for the estimation of necessary parameters. One notable advantage of naive Bayes learners and classifiers is their efficiency, surpassing that of more complex methods. The decoupling of the conditional distribution in classification means that each feature can be independently estimated as a one-dimensional distribution. This decoupling contributes to addressing challenges associated with dimensional disasters, making naive Bayes models effective and swift in real-world applications.

Decision Tree based classifications:

In the realm of traditional machine learning, a decision tree serves as a predictive model that elucidates the mapping relationship between object attributes and their corresponding values. Each node within the decision tree represents a specific object, while each branching path signifies a potential value. The leaves of the tree denote the values associated with an object, determined by the path taken from the root node to that specific leaf. Notably, a decision tree produces a singular output; however, to handle intricate values, multiple independent decision trees can be constructed to yield diverse outputs. Widely employed in data mining, the decision tree technique proves versatile for tasks such as data classification, prediction, and regression. In the context of classification, the decision tree model is meticulously crafted during the training stage based on the provided dataset. The most pertinent feature segmentation node is selected from the root node, setting the foundation for subsequent phases. Both the test and prediction phases rely on the decision tree model, meticulously constructed during the training process. The classifier builds a tree-like structure, where each internal node represents a decision based on a feature, each branch represents the outcome of the decision, and each leaf node represents the final classification or regression value.

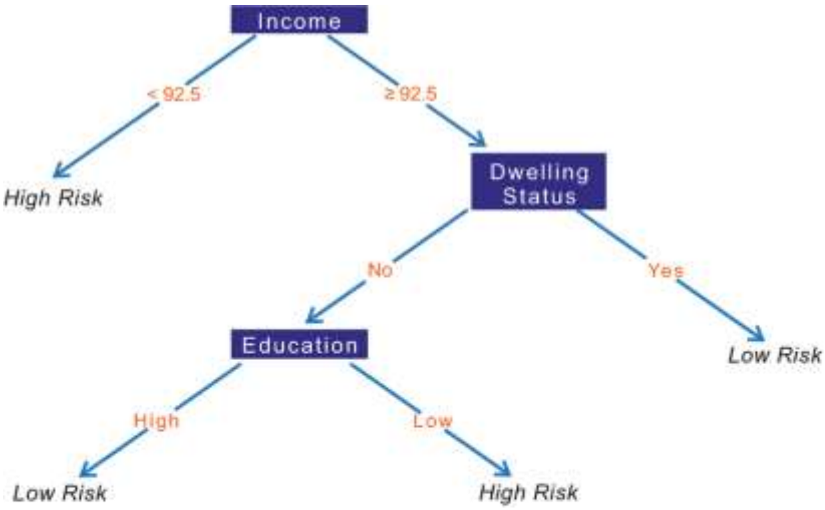

Fig 2. Example of decision tree model

4. Road Ahead

In conclusion, machine learning stands as a potent ally in the realm of agritech, offering multifaceted solutions to age-old agricultural challenges. Its ability to analyse vast amounts of data and discern patterns enables predictive modelling for crop yields, disease outbreaks, and optimal resource allocation. By harnessing machine learning algorithms, agritech endeavours benefit from precision agriculture practices, where interventions such as irrigation, fertilization, and pest control are tailored to specific crop needs, thus maximizing yields while minimizing inputs. Machine learning empowers farmers with actionable insights derived from satellite imagery, weather data, and soil analyses, enabling informed decision-making and proactive risk management. This technology also facilitates early detection of crop diseases and pests, mitigating losses and reducing reliance on chemical treatments. In addition to on-field applications, machine learning-driven market platforms streamline supply chains, connecting farmers directly with buyers and ensuring fair prices.

References:

1. Smith, D., & Peng, W. (2009, February). Machine learning approaches for soil classification in a multi-agent deficit irrigation control system. In 2009 IEEE International conference on industrial technology (pp. 1-6). IEEE.
2. Osorio, K., Puerto, A., Pedraza, C., Jamaica, D., & Rodríguez, L. (2020). A deep learning approach for weed detection in lettuce crops using multispectral images. AgriEngineering, 2(3), 471-488.
3. Tian, Z., Wang, J. W., Li, J., & Han, B. (2021). Designing future crops: challenges and strategies for sustainable agriculture. The Plant Journal, 105(5), 1165-1178.

4. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674.
5. Veeragandham, S., & Santhi, H. (2020). A review on the role of machine learning in agriculture. Scalable Computing: Practice and Experience, 21(4), 583-589.
6. Sharma, A., Jain, A., Gupta, P., & Chowdary, V. (2020). Machine learning applications for precision agriculture: A comprehensive review. *IEEE Access*, *9*, 4843-4873.
7. Noor, U., Anwar, Z., Amjad, T., & Choo, K. K. R. (2019). A machine learning-based FinTech cyber threat attribution framework using high-level indicators of compromise. *Future Generation Computer Systems*, *96*, 227-242.
8. Huang, S. C., Wu, C. F., Chiou, C. C., & Lin, M. C. (2021). Intelligent FinTech data mining by advanced deep learning approaches. *Computational economics*, 1-16.
9. Chen, T. H., & Chang, R. C. (2021). Using machine learning to evaluate the influence of FinTech patents: The case of Taiwan's financial industry. *Journal of Computational and Applied Mathematics*, *390*, 113215.
10. Agarwal, S., Alok, S., Ghosh, P., & Gupta, S. (2020). Financial inclusion and alternate credit scoring for the millennials: role of big data and machine learning in fintech. *Business School, National University of Singapore Working Paper, SSRN, 3507827*.
11. Raj, B., & Upadhyay, V. (2020). Role of FinTech in Accelerating Financial Inclusion in India. In *3rd International Conference on Economics and Finance organised by the Nepal Rastra Bank at Kathmandu, Nepal during February* (pp. 28-29).
12. Chen, T. H., & Chang, R. C. (2021). Using machine learning to evaluate the influence of FinTech patents: The case of Taiwan's financial industry. *Journal of Computational and Applied Mathematics*, *390*, 113215.
13. Narayanan, V. L., Pandi, G. R., Kaleeswari, K., & Veni, S. (2023). MACHINE LEARNING ALGORITHM FOR FINTECH INNOVATION IN BLOCKCHAIN APPLICATIONS. *ICTACT Journal on Soft Computing*, *14*(1).
14. Maheshwari, S., & Chatnani, N. N. (2023). Applications of artificial intelligence and Machine learning-based supervisory technology in financial Markets surveillance: A review of literature. *FIIB Business Review*, 23197145231189990.
15. Singh, A. K., Sharma, P. M., Bhatt, M., Choudhary, A., Sharma, S., & Sadhukhan, S. (2022, November). Comparative Analysis on Artificial Intelligence Technologies and its Application in FinTech. In *2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)* (pp. 570-574). IEEE.
16. Stojanović, B., Božić, J., Hofer-Schmitz, K., Nahrgang, K., Weber, A., Badii, A., ... & Runevic, J. (2021). Follow the trail: Machine learning for fraud detection in Fintech applications. *Sensors*, *21*(5), 1594.
17. Pattnaik, D., Ray, S., & Raman, R. (2023). Applications of artificial intelligence and machine learning in the financial services industry: A bibliometric review. *Heliyon*.
18. Anil, K., & Misra, A. (2022). Artificial intelligence in Peer-to-peer lending in India: a cross-case analysis. *International Journal of Emerging Markets*, *17*(4), 1085-1106.
19. Biju, A. K. V. N., Thomas, A. S., & Thasneem, J. (2023). Examining the research taxonomy of artificial intelligence, deep learning & machine learning in the financial sphere—a bibliometric analysis. *Quality & Quantity*, 1-30.
20. Malhotra, P., Patel, P., Shah, N., Veera, R., & Sanghavi, R. (2019). Customer Churn prediction for FinTech Companies using Artificial Neural Networks. *i-Manager's Journal on Computer Science*, *7*(4), 46.
21. Goodell, J. W., Kumar, S., Lim, W. M., & Pattnaik, D. (2021). Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. *Journal of Behavioral and Experimental Finance*, *32*, 100577.