ARTIFICIAL NEURAL NETWORKS AND THEIR USE IN AGRICULTURAL GENETIC ENGINEERING and PROTEIN ESTIMATION

ABSTRACT

A mathematical model called an artificial neural network uses certain algorithms to anticipate and predict various events. This model has multiple layers, including input, hidden, and output layers. By altering its algorithms, various outputs can be produced based on the input utilized. Biological neuron mechanization has an impact on artificial neural networks. As biological neurons have a propensity to learn and train sets of data for producing biased outputs by spotting and removing variances in them, ANN also works on these principles. Although this model has many uses, it has historically been employed in biological experiments using the supervised learning method, one of which is to predict protein secondary structure. This allows one to identify the positions of different amino acids that are ordered complicated protein structures, which are very appealing in genetic engineering. By this model scientists can map out and isolate a desirable gene in genetic material without going through intense laborious experiments. This report summarizes all the objectives linked to artificial neural networks as well as their applications in bioengineering by examining many related studies.

Keywords: Bioengineering, Artificial Intelligence, Protein analysis, Bioinformatics, Neural Network

1. Introduction

Artificial neural networks (ANN) have been developed as expansions of statistical models of human nervous systems (Hill *et al.*, 1994). Following the development of simpler neurons, neural networks have seen a surge in attention (Osenblatt,1958). Artificial neurons, also referred to as nodes or neurons, are the basic computational nodes in neural networks. Dynamic, self-organizing, and effective learning are some qualities that benefits artificial neural networks over conventional logic-based artificial intelligence in processing intuition and unstructured data. When compared to conventional artificial intelligence and information processing technologies, artificial neural networks use an altogether different mechanism by employing parallel processing system (Kasbov *et al.*, 2016).

The essence of an ANN model is the network transformation and intermittent nature of parallel distributed information architectures, and it can imitate people to varying degrees and levels. Artificial neural networks are information processors inspired by the brain that don't require programming (He *et al.*, 2004). AANs, or machine learning algorithms, are used significantly in a variety of fields, including agriculture and bioinformatics. Due to their capacity to represent complicated interactions between inputs and outputs in a non-linear manner, artificial neural networks (ANNs) have become more and more popular in the field of agriculture. ANNs have been employed in agriculture for several tasks, including predicting crop production, spotting diseases, evaluating soil quality, and forecasting the weather (Khairunniza-Bejo *et al.*, 2014; Gandhi and Armstrong, 2016). The task of protein estimation, which entails predicting different aspects of proteins, including their function, structure, and interactions, has made extensive use of ANNs (Shimon and Gad, 1994). ANNs are active in a number of interdisciplinary domains, including computer science, cognitive science, neurology, life science and artificial intelligence (Wu *et al.*, 2018). Being universal non-linear function approximators, these networks can be used without any preliminary knowledge of the nature of the input being represented. These two justifications encourage their modeling of the extremely complicated mechanisms connected to cell development and biocatalysis (Montague and Morris, 1994).

Artificial neural networks (ANNs) play a crucial role in climate and environmental research by enabling the analysis, modeling, and prediction of complex systems. A deeper comprehension of processes and phenomena is made possible by ANNs' exceptional ability to capture non-linear correlations and patterns in climate and environmental data (Fathian *et al.*, 2019; Castellano-Navarro *et al.*, 2021). They assist academics simulate and anticipate the future, evaluate the effects of climate change, and support decision-making processes. They also enable data-driven modeling and prediction (Imran *et al.*, 2020). ANNs also help with pattern detection and categorization of various environmental parameters, such as forms of land cover and extreme weather conditions. Additionally, ANNs can be used to quantify the uncertainty related to forecasts for the climate and the environment, giving decision-makers critical information (Lapides *et al.*, 2021).

2. What is Artificial Neural Network?

"Artificial neural networks (ANN) derive from biological neural networks (BNN), with interconnected neurons arranged in different layers, akin to the human brain." These neurons are called nodes. Inputs in the artificial neural network are represented by dendrites of biological nervous system, while nuclei represent nodes, weights are represented by synapses, and outputs are represented by axons." Artificial neural networks mimic the human brain to enable computers to make human-like decisions. They are created by programming computers to act like interconnected brain cells. Please cut down the length of this passage.

BNN	ANN
Dendrites	Inputs

Cell nucleus	Nodes
Synapse	Weights
Axon	Output

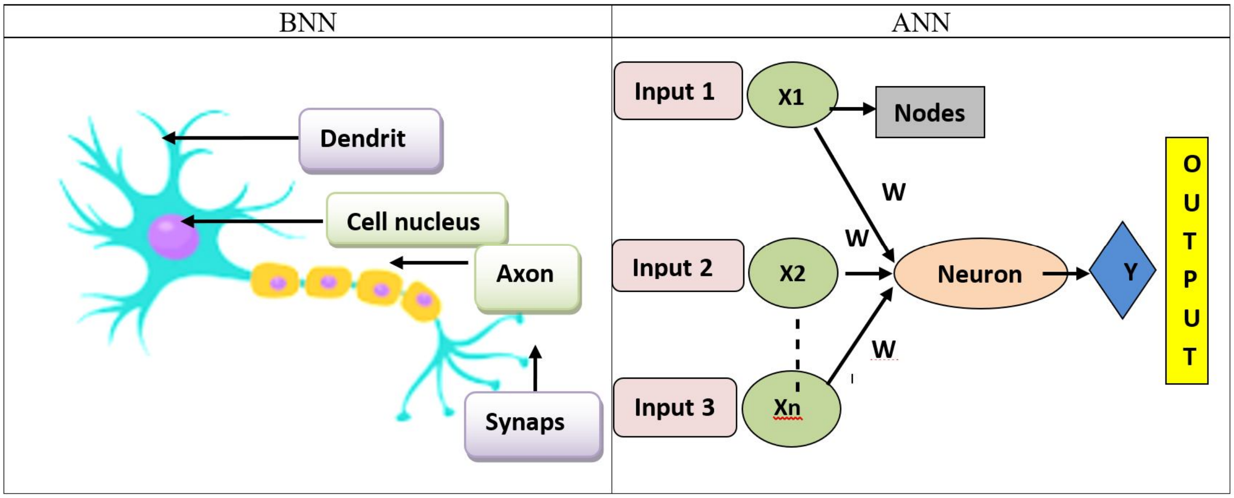

Fig. 1. Similarities between biological neural network and artificial neural network:

3. Principles of ANN

The structure and operation of the nervous system serve as the foundation for the principles underlying Artificial Neural Networks (ANNs). Artificial neurons (ANNs) are composed of interconnected processing nodes that are arranged in layers (McCulloch and Pitts, 1943). The ability of ANNs to learn complex relationships between inputs and outputs makes them effective for a number of applications, including computational linguistics, voice recognition, and image classification. The main advantage of neural networks is their capacity to uncover some information hidden in data that was previously undiscovered (but they are not able to extract it). The process of "training" or "learning" of neural networks is the way to "capture" the unknown information (Daniel *et al*., 1997). The principle model of ANN consist of different parts which is interlinked or interconnected to perform in respective manner of biological neurons to give accuracy outputs, the similarity between both is shown in Fig. 1. The different part includes: *Artificial Neurons* which are designed to operate like a biological neuron, accepting a variety of signals from numerous nearby neurons and processing them in a straightforward manner that is predetermined (Zupan, 1994). Each artificial neuron gets information from other neurons and uses a mathematical function to process that information to create an output. The network can analyze and transport data from one layer to another since the output of one neuron serves as the input for other neurons. Artificial neuron networks (ANNs) are composed of several layers, including an input layer, one or more hidden layers, and an output layer (Miguel *et al*.,2019). Raw input data, such as visuals, sounds, or text, are received by the input layer, while the output layer generates the network's final prediction or judgment. The information is processed by the hidden levels before being sent to the output layer for further processing. Weights are used to represent the strength of connections between neurons in neural networks. During the training phase, the weights are modified to enhance the network's performance. A collection of weights contains the information that is transmitted from one processing element to another (Agatonovic-Kustrin and Beresford, 2000). *Activation Function*, the activation function of each neuron determines how the input is processed. It is the functions that act upon weighted inputs at a neuron to produce the neuron output (Montague and Morris, 1994). Sigmoid, ReLU (rectified linear unit), and Tanh (hyperbolic tangent) are some examples of activation function. The network is given non-linearity through the activation function, enabling it to learn intricate correlations between inputs and outputs (Sharma *et al*., 2020). Based on the predictions that were incorrect, the learning algorithm is responsible for changing the network's weights. To reduce the error between the network's predictions and the actual goal values, this is accomplished by utilizing optimization techniques like as gradient descent (Donaldson *et al*., 1993). And lastly *Back propagation*. By updating the network's weights based on the gradient of the error, a process known as back propagation, thus allowing network to learn from its errors and make predictions more accurate over time (Hornik *et al*., 1989).

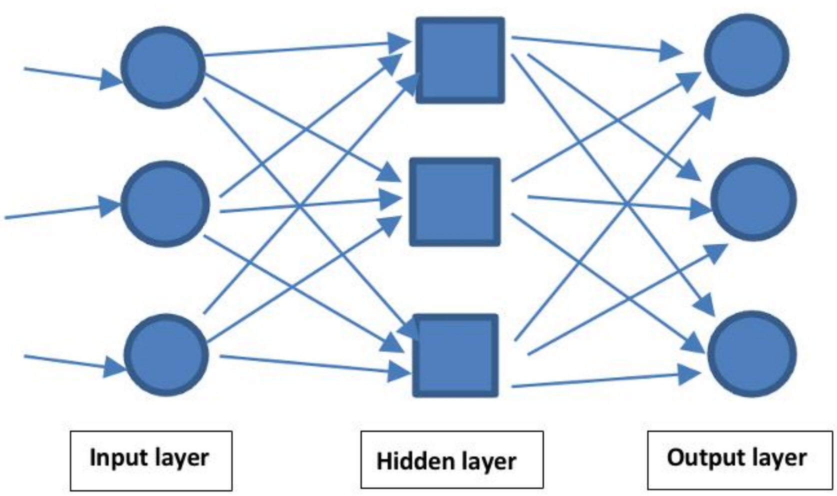

Fig. 2. Three-layered ANN model: where input, hidden and output layer are interlinked to each other to give valid output.

3.1 Working of ANN

A neuron processing unit in a synthetic neural network can represent a wide range of things, such as features, letters, concepts, or practical abstraction patterns. The three categories into which the processing units of the network are divided are input unit, output unit, and hidden unit, a diagrammatic representation is mimic in Fig. 2. The input unit will accept signals and data from outside sources (Balcazar, 1997). The output unit creates the system's final output. Unnoticeable from the outside of the system, the hidden unit is positioned between the input and output units (Agatonovic-Kustrin and Beresford, 2000). The connection weights contribute to the effectiveness of the connections between the neurons. The network processing unit's connectivity structure is an expression of how information is represented and processed. The Fig. 3 shows the complete procedure for predicting unknown output by ANN.

A transfer function is used to represent the neurons' non-linear activity in a condensed statistical model of the neuron, while connection weights are used to explain the synaptic effects, which regulate the influence of the linked input signals. The neuron impulse is then calculated using the weighted sum of the digital signal after they have been interpreted by the transfer function. Artificial neurons are capable of learning by changing the weights to match the chosen learning strategy (Peter and Richard, 2005).

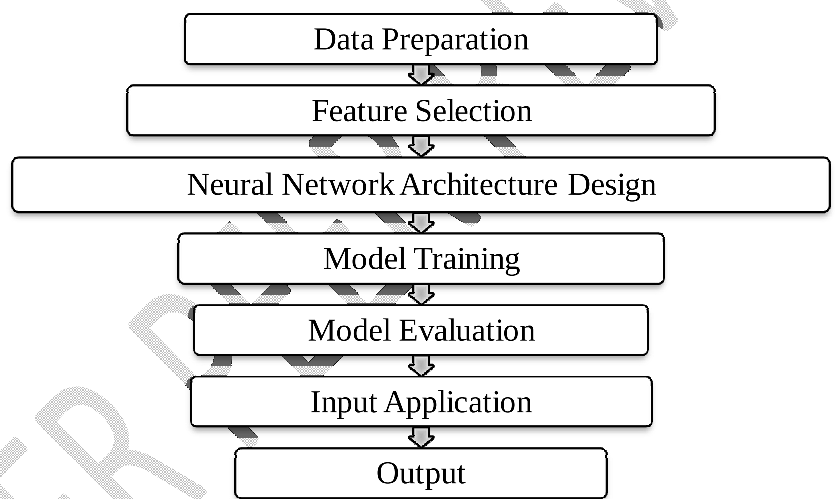

Fig. 3. Flowchart consists of different steps involved in output prediction using ANN Model.

4. Application of ANN

ANNs have numerous applications in a variety of fields. Its use has grown in recent decades due to its high level of accuracy in prediction and forecasting. Some of the applications in which ANNs are frequently used include information processing and pattern recognition, risk analysis, information storage, speech recognition, and many medical and research assessments (Hill *et al.*, 1994). It is also used to forecast and predict many outputs mentioned in Fig. 4. Climate forecasting is an important application of ANNs in agriculture, as it can help farmers make informed decisions about crop management in the face of uncertain environmental conditions. Sahi *et al*. (2020) used ANNs to predict rainfall in India based on historical climate data, achieving high accuracy in predicting rainfall for different regions in India. It is employed in the forecasting of market trends. ANN and discriminate analysis have been used to predict stock price performance (Yoon and Swales, 1990). ANNs have been used for weed and pest management in crops, which is important for reducing crop losses and improving yields. Xia *et al*. (2021) used ANNs to detect weeds in soybean crops based on hyper spectral imaging data, also Bhardwaj *et al*. (2020) used ANNs to predict the infestation of whiteflies in cotton crops, outperforming other models in terms of accuracy and efficiency.

ANNs are a good substitute for sophisticated non-linear models used inu economic. In a well-done study, Donaldson *et al*. (1993) evaluated the applicability of several well-known conditional volatility models for capturing the fat-tailed and alternate method aspects of investment returns using stock data that was taken from the London, New York, Tokyo, and Toronto exchanges. ANN was also employed as a backup technique for forecasting business venture outcomes and national economic trends. Some examples for the same are; projection of banking crises in Texas by Tam (1991) and Tam and Kiang (1992). Using financial and political factors, Roy and Cosset (1990) also employed logistic regression and artificial neural network models to forecast country risk assessments. One of the most significant uses of ANNs in agriculture is the prediction of crop production. Crop yield can be predicted by ANNs using a variety of input variables, including meteorological, soil characteristics, and management techniques. According to a study by Kachwala *et al*. (2021), ANNs fared better than other models at forecasting the yield of wheat crops in India using an assortment of meteorological and non-climatic characteristics. Informed decisions about crop productivity and higher yields can be made by farmers with the aid of accurate crop yield prediction.
However, ANN has multiple challenges. The architecture and programming of artificial neural networks are constantly evolving, in contrast to many statistical modeling techniques that are stable and fixed. Second, commercialized software for artificial neural networks usually lags developments in the field, whereas software for statistical methods is freely accessible and of good quality. Finally, artificial neural network models are harder to understand and give a physical interpretation to than many other forecasting models. Fourth, ANN can experience overfitting concerns because they have more variables to examine than most conventional forecasting models. Finally, ANN take up more computational time than statistical models (Hill *et al*., 1994). But other than its few flaws it has great potential in future agriculture, biotechnology and other fields in predicting and forecasting outputs.

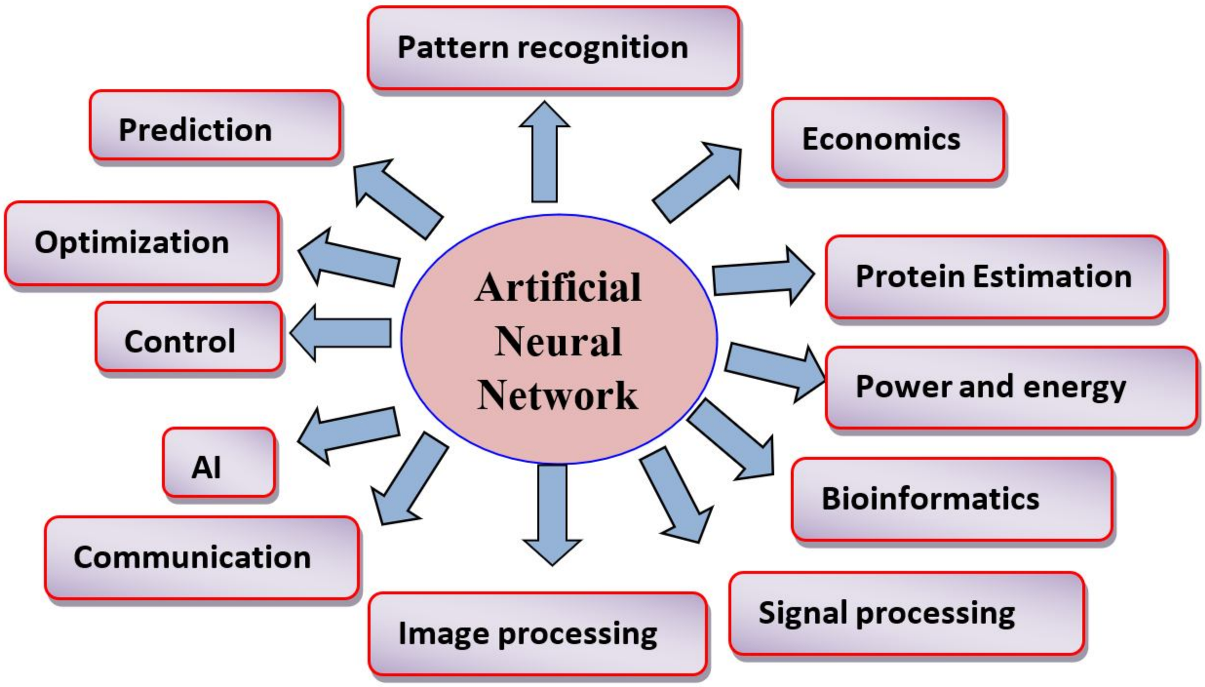

Fig. 4. **Different applications of artificial neural network (ANN).**

4.1. Application in bioengineering

There are several uses for artificial neural networks (ANNs) in bioengineering. ANNs may mix and use both experimental and theoretical data to solve issues. Analysis, estimation, and classification or pattern recognition are some of the many applications of ANNs. In the pharmaceutical sector, supervised associating networks can be utilized as an alternative to the conventional response surface methodology (Agatonovic-Kustrin and Beresford, 2000). Some of the application is mentioned in table 1 along with it’s importance.

Bioengineering is an emerging field that involves the application of engineering to biological systems. By increasing crop yields, increasing crop resilience to pests and diseases, and lowering the environmental effect of agricultural methods, bioengineering has the potential to greatly improve crop production in agriculture. According to Klümper and Qaim (2020), GM crops have contributed to higher yields, reduced pesticide use, and increased farmer profits. The production of bio agent such as *Tricoderma*, *Pseudomonas*, BGA, PSB, *Rhizobium*, *Xanthomanas* and many more helps in improving soil physical and chemical properties and maintaining healthy Rhizosphere (Kour *et al.*, 2020). The other applications including:

Drug discovery

To help find new drug candidates, ANNs can be used to forecast the affinity of medicines for their intended protein targets. Moreover, ANNs can be used to predict drug toxicity, assisting in the identification of chemicals that appear to be safe for administration in humans (Carrico *et al.*, 2018) as well as the manufacture of pesticides for plants.

Gene expression analysis
ANNs can be used to analyze large gene expression data sets and identify patterns in the data that are indicative of disease or other biological processes. The orientation and synthesis of RNA and DNA molecule sequences, as well as the bending and molecular structure of RNA strands, have all been accomplished using ANNs trained with GA algorithms (Reidys *et al.*, 1997; Micheal, 1998).

Disease Diagnosis
ANNs have been used for disease diagnosis in crops, which is crucial for effective disease management (Pérez-Bueno *et al.*, 2016). A study by Jamshidi *et al.* (2020) used ANNs to diagnose apple tree diseases based on image analysis, achieving high accuracy in detecting and classifying different diseases in apple trees. Accurate disease diagnosis can help farmers prevent and manage diseases in their crops, leading to higher yields and quality produce.

Protein folding
ANNs can be used to predict the 3-D proteins structure, which is crucial for understanding protein function. Gribskov *et al.* (1987) developed the position-specific scoring matrix (PSSM) to identify widely related proteins. It has accomplished considerable success in predicting protein binding sites, protein secondary structure, and unstructured areas (Chen and Jeong, 2009; Jones, 1999).

Image analysis
ANNs can be used to analyze medical images, such as X-rays, PET, Histopathology imaging and MRI scans, to help diagnose illness and aid in the assessment of potential treatments for better recovery (Najarian and Splinter, 2005).

Plant Phenotyping
Plant phenotyping is another area where ANNs have been applied, which is important for understanding plant growth and development. Cruz *et al.* (2021) used ANNs to predict the biomass of maize plants based on hyperspectral imaging data, achieving high accuracy in predicting the biomass of maize plants. Accurate plant phenotyping can help farmers optimize their crop management practices and improve their yields (Costa *et al.*, 2019).

Personalized medicine
ANNs can be used to develop personalized medicine strategies by predicting patient response to treatments based on genetic and other patient data. A multi-layers feed-forward neural network was designed to predict the human intestinal absorption (HIA%) of therapeutic drugs based on their molecular structure (Wessel *et al.*, 1998).

Predictive maintenance in bioprocessing
ANNs can be used to predict equipment failure and optimize maintenance schedules in bioprocessing, reducing downtime and improving efficiency. Overall, ANNs have the potential

to revolutionize bioengineering by providing new insights into biological systems and improving the efficiency and accuracy of various bioengineering applications.

Table 1. The application of ANN and their utility in different biotechnical sciences.

Application	**Utility**	**Reference**
Protein structure prediction	Improves drug discovery process by providing insights into protein-ligand interactions and help in production of GM crops or plant.	Wu *et al*., 2021; Carrico *et al*., 2018
Biomedical signal processing	Improves diagnosis and monitoring of various diseases by accurately analyzing and interpreting biomedical signals.	Acharya *et al*., 2018
Tissue engineering	Enhances the design of artificial tissue, artificial seeds and organs by predicting their mechanical properties.	Abbasi *et al*., 2020
Bioinformatics	Facilitates the analysis of large biological datasets, such as genomics and proteomics, by predicting relationships and interactions between genes, proteins, and other biomolecules.	Angamuthu and Subramanian, 2021; Micheal, 1998
Biomaterials design	Enables the design of novel biomaterials with specific properties such as biocompatibility and degradation rate, for various biomedical applications.	Raza *et al*., 2021

5. Protein estimation

A vital component of agriculture is protein estimation because proteins are essential for plant growth and development. In addition to being the foundation to produce enzymes, hormones, and other compounds, proteins are engaged in several metabolic pathways. Artificial neural networks (ANNs) are used to estimate protein content by analyzing vast volumes of data pertaining to protein content and structure (Seonwoo *et al*., 2017). ANNs are computer models that imitate how the human brain functions (Fig. 1) and can be used to assess and predict outcomes based on large amounts of unstructured information (He *et al*., 2004). Protein estimation is crucial for evaluating the nutritional quality and economic value of agricultural products, as well as for identifying genetic factors that contribute to protein content and facilitating the development of high-protein crop varieties. Muneer *et al*. (2021) concluded protein estimation is essential for evaluating the nutritional quality of wheat flour, similarly Bhaduri *et al*. (2020) highlighted the importance of protein estimation in soybean breeding programs. The authors employed quantitative trait locus (QTL) mapping to identify genomic regions linked with protein content in soybean.

In protein analysis, ANNs are trained on a huge dataset of protein sequences as well as their associated protein concentrations. The ANN processes this training data to produce a model that forecasts a protein's concentration based on its pattern. Protein sequence is transmitted to the ANN, which analyses the data and produces a predicted secondary structure for the protein, to make the conclusion (Seonwoo *et al*., 2017), diagrammatic representation of protein estimation and the different steps involved is explained in Fig. 5. The prediction can take the form of a continuous value, which represents the likelihood that each protein residue will form a specific secondary structure, or a discrete value, which represents the expected secondary structure for each protein residue.

In several instances, it has been demonstrated that ANNs are effective at estimating proteins with great accuracy and precision. Spencer *et al*. (2015) coupled PSSM and Atchley factors with DBN utilised for amino acid sequencing in order to predict protein structures. When doing high-throughput analysis, which requires swift and precise analysis of numerous protein samples, the use of ANNs in protein estimation is extremely helpful. In general, protein estimation by ANN is an approach that shows great potential for forecasting protein concentrations and has the significance to be useful in several sectors, including biochemistry, biotechnology, plant breeding, and clinical research.

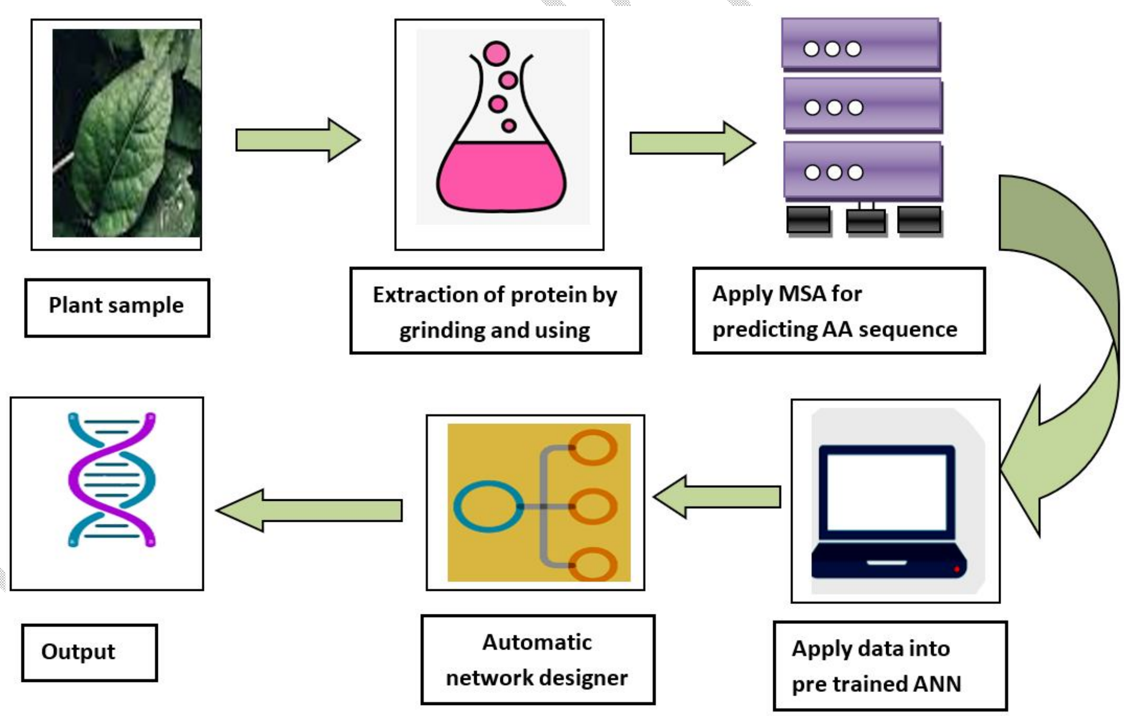

Fig. 5. Diagrammatic representation of protein estimation by plant cell using ANN.

5.1 Different ANN models for protein analysis

Convolutional neural networks, feedforward neural networks, and recurrent neural networks are some of the ANN types that can be used to predict protein secondary structure. Depending on the unique needs of the analysis, such as the complexity of the protein sequences and the required level of accuracy in secondary structure prediction, a specific ANN is employed (Montague and Morris, 1994). See table 2 for more information. The various kinds of artificial neural networks that can be applied to protein analysis include:

Feed forward Neural Networks

These ANNs, which are the most basic kind, have an input layer, one or more hidden layers, and an output layer. Protein sequence information is ingested by the input layer, and the anticipated protein concentration is delivered by the output layer. The information is processed by the hidden layers before being sent to the output layer. DEEPred is a multi-task feed-forward neural network that has been shown to be particularly useful for analysing Gene Ontology and offers a considerable improvement over existing techniques when the amount of training instances is high (Rifaioglu *et al*., 2019).

Convolutional Neural Networks (ConvNets)

ConvNets are particularly well suited for protein analysis as they can handle the two-dimensional structure of proteins. ConvNets, a popular tool for image analysis, can be modified for protein analysis by employing filters to find patterns in the protein sequence (Krizhevsky *et al*., 2012). Detecting protein binding sites and determining the absolute binding affinity of proteins with ligands are few bioinformatics applications that employ’s 3DCNN (Doerr *et al*., 2017; Jime'nez *et al*., 2017).

Recurrent Neural Networks (RNNs)

RNNs are made to handle long data sequences, like protein sequences, by storing and retrieving information over time. This makes RNNs well suited for protein analysis, as they can capture the relationships between amino acids in a protein sequence. RNNs' abilities to map a variable-length input sequence of some other sequence or make fixed-size predictions hold promise for bioinformatics study (Seonwoo *et al*. 2017).

Long Short-Term Memory (LSTM) Networks

By using memory blocks to store data, LSTMs are a sort of RNN that can handle long sequences of data, such as protein sequences (Hochreiter and Schmidhuber, 1997). Because they can capture the correlations between amino acids over a large length of data, LSTMs are especially helpful for protein analysis (Zhang *et al*., 2017). Long-range dependencies are modelled by LSTM, which employ several gate vectors at each place to govern the passage of information along a sequence (Zhou *et al*., 2020).

Deep BeliefNetworks (DBNs)

DBNs are a form of ANN that use numerous layers of hidden units to simulate complex relationships in the data. DNNs have been used for protein analysis and have been shown to achieve high levels of accuracy in protein concentration prediction. With each consecutive layer, DNNs construct more complex features using low level features as their input. DBN exploit unlabeled data and can help avoid over fitting, researchers are able to obtain normalized results, even when labelled data are inadequate as is common in the real world (Erhan *et al*., 2010).

Table 2: Comparative study of different ANN models and their application in protein estimation.

ANNs Model	Purpose	Application	Reference
Feed forward Neural Network	Predicting protein solubility	Useful in designing new biopharmaceuticals compounds and studying Gene Ontology.	(Dusad *et al*., 2020; Rifaioglu *et al*., 2019)
CNN	Predicting protein-ligand binding affinity, identify patterns in the protein sequence	Useful in drug discovery and design	(Ahmadi *et al*.,2021; Krizhevsky *et al*., 2012)
LSTM	Predicting protein secondary structure, used in positioning AAs in protein	Useful in drug design and discovery	(Ding *et al*., 2020; Zhang *et al*., 2017)
RNN	Predicting protein-protein interactions	Useful in understanding protein functions and designing new drugs	(Sankar *et al*., 2021)
DBN	Estimating protein content in rice, protein concentration prediction	Offers better accuracy compared to traditional methods, can be used under insufficient input data.	(Dong *et al*., 2021; Erhan *et al*., 2010)

6. Observation

A growing number of agricultural applications, including crop production prediction, insect detection, and disease diagnostics, utilises artificial neural networks (ANNs). As an illustration, ANNs have been used to accurately predict maize production based on climate and soil variables (Liu *et al*., 2021; Yoo *et al*., 2016). In addition, ANNs can be used to identify plant diseases using leaf images (Mohanty *et al*., 2016). Furthermore, ANNs can be used in protein estimation, such as in the determination of protein content in wheat grains (Mousa *et al*., 2019). Despite the promising applications of ANNs in agriculture and biotechnology, there are still some limitations that need to be addressed for better results.

One of the limitations of ANNs in agriculture is the lack of interpretability. The interactions between the input and output factors in ANNs are frequently referred to as "black-box" models since they are difficult to comprehend. This can make it difficult to identify the key factors that affect crop yield or pest infestations, which are important for making informed decisions in agriculture. Moreover, the generalisability of ANNs is another concern. Models trained on one dataset or location may not necessarily perform well on a different dataset or location. This is particularly relevant in agriculture, where crop growth can vary widely depending on soil type, weather conditions, and other factors.

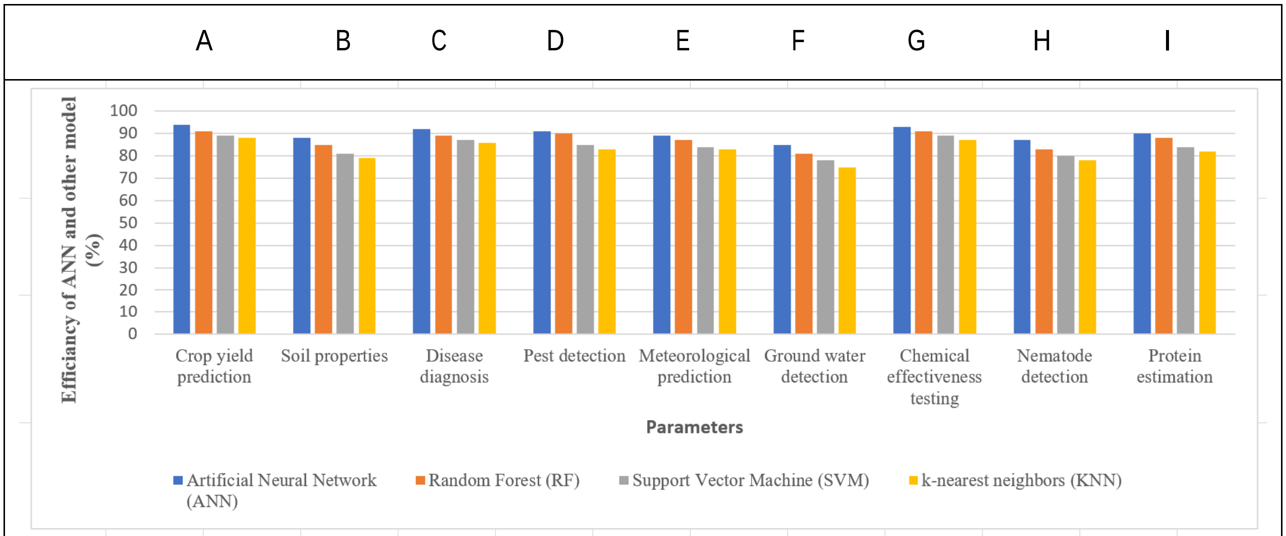

Fig. 6. Efficiency of ANN and models on different applications; where A) Suresh et al., 2021; B) Wang et al., 2020; C) Lee et al., 2021; D) Chandra et al., 2021; E) Li et al., 2020; F) Mishra et al., 2021; G) Zhang et al., 2020; H) Ali et al., 2021 and I) Rani et al., 2020.

Another limitation of ANNs in agriculture is the need for substantial volumes of data to train the model accurately. In some cases, there may not be enough data available, or the data may be of poor quality, which can lead to inaccurate predictions. Furthermore, ANNs may not perform well when faced with unexpected or extreme conditions, such as unusual weather patterns or new pest outbreaks.

ANN also have inability to handle complex and dynamic systems. Agricultural systems are complex and dynamic, involving multiple factors such as climate, soil, crop variety, and management practises. It's possible that ANN can't fully capture the complexity of these systems, which results in forecasts being inaccurate. Moreover, overfitting is a problem that affects ANN, which happens when the model is too complicated and matches the training data too tightly, resulting in subpar performance on new data. This issue can be particularly problematic in agriculture where the data is often limited, leading to an increased risk of overfitting.

To overcome these limitations, there are several things that can be done in the future. One approach is to combine ANNs with other modelling techniques, such as decision trees or fuzzy logic, to improve the issues of interpretability and overfitting, also increase model accuracy in predicting outcomes. Additionally, efforts should be made to collect high-quality information from diverse sources, such as remote sensing and meteorological data, to improve the performance of ANNs in agriculture.

A recent study by Gao *et al*. (2021) proposed a hybrid model that combines ANN with fuzzy logic for the prediction of soil moisture content in agriculture. The hybrid model was found to outperform both ANN and fuzzy logic alone, highlighting the potential benefits of combining different methods in agriculture. Similarly, Kumar *et al.* (2021) demonstrates the potential of using ANNs for crop yield prediction in rice farming. The authors used a multi-layer perceptron (MLP) ANN model to predict the yield of rice crops based on various input variables, such as weather conditions, soil properties, and fertiliser application rates. The results showed that the ANN model outperformed traditional statistical models in predicting rice yield. However, the researcher concluded that further studies is required to improve the accuracy and interpretability of the model. While ANNs have the potential to improve agriculture, there are limitations and drawbacks that need to be addressed for better results. By combining ANNs with other modelling techniques and collecting high-quality data, we can improve the accuracy and interpretability of ANNs in agriculture, leading to better decision-making and more sustainable agricultural practices.

7. Conclusion

AANs are a computational representation of the human brain that is made up of artificial neurons. The non-linearity of ANN gives it an edge over conventional logic-based machine intelligence in processing intuition and unstructured input. AANs have proven to be effective techniques for estimating protein. They are well suited for numerous proteins estimating tasks, such as function prediction, structure prediction, and interaction prediction, due to their capacity to learn complicated correlations between inputs and outputs. Research on the use of ANNs to protein estimate is still ongoing, and new developments are anticipated to further boost the precision and reliability of protein estimation.

It also has great application in designing and discovery on new pharmaceutical compounds or drugs which are better suitable for human consumption and have higher effectiveness over subject. There is always scope for better, although ANNs have been successful in solving complex problems but they also have some drawbacks which includes, more computational requirements, complex architectures make it difficult to understand, time consuming, vulnerability to adversarial attacks and requirements of large amount of training data to learn patterns and relationships. But overall ANNs are great asset in predicting and forecasting of many outputs with high level of accuracy.

ANNs are utilized for a wide range of tasks, including image processing, protein estimation, risk and failure analysis, medicine design, capital estimation, and many others (Fig. 4). It is impossible to overestimate the significance of Artificial Neural Networks (ANNs) in agriculture given their vast range of applications for assisting farmers in making better decisions and increasing crop yields. ANNs can help in several ways, including accurate agricultural production prediction, disease detection, soil quality evaluation, climate forecasting, weed and pest control, and plant phenotyping. In Fig. 6, the graph shows clear positive response of ANN with many parameters as compared to other models. ANNs are the best choice for jobs involving high-dimensional data and ambiguous environmental variables because they can represent complicated non-linear interactions between inputs and outputs.

References

1. Abbasi, N., and Ansari, R. (2020). Application of artificial neural networks for prediction of mechanical properties of tissue engineering scaffolds: A review. Journal of Materials Science: Materials in Medicine, 31(2), 14.
2. Acharya, U. R., Hagiwara, Y., and Koh, J. E. W. (2018). Diagnosis of cardiovascular diseases via deep learning techniques: A review. Journal of Healthcare Engineering, 2018, 1-16.
3. Ahmadi, M., Sharifi, M., and Eslahchi, C. (2021). Deep learning-based prediction of protein-ligand binding affinity using convolutional neural networks. Journal of Molecular Graphics and Modelling, 108:107885.
4. Rifaioglu, A. S., Doğan, T., Martin, M. J., RengulCetin-Atalay and Atalay, V. (2019). DEEPred: Automated Protein Function Prediction with Multi-task Feed-forward Deep Neural Networks. 9:7344.
5. Ali, S., Anjum, S., Khan, A. R., and Nawaz, M. A. (2021). A review on the application of machine learning techniques for plant nematode detection. Computers and Electronics in Agriculture, 184, 106015.
6. Angamuthu, K., and Subramanian, K. (2021). A comprehensive review of artificial neural network applications in bioinformatics. Journal of Molecular Graphics and Modelling, 107, 107848.
7. Balcazar, J. (1997). Computational power of neural networks: A characterization in terms of Kolmogorov complexity. IEEE Transactions on Information Theory, 43(4): 1175-1183.
8. Bhaduri, K., Bhattacharya, R., Singh, A. K., Singh, R. K., Sarkar, A., Pal, A. K., ... and Singh, N. K. (2020). Genome-wide association study for protein content in soybean [*Glycine max* (L.) Merrill]. Plant Breeding, 139(5), 902-910.
9. Bhardwaj, R., Mahajan, J. and Singh, N. (2020). Artificial Neural Network-based prediction of whitefly (*Bemisia tabaci*) infestation in cotton crop. Journal of Agrometeorology, 22(2): 251-256.
10. Carrico, I., *et al*. (2018). Artificial neural networks in drug discovery. Journal of Computer-Aided Molecular Design, 32(2), 149-164

11. Castellano-Navarro, A., Nieto-García, M., & Sánchez-García, E. (2021). Detecting and classifying satellite images of atmospheric patterns with convolutional neural networks. Remote Sensing, 13(9), 1669.
12. Chandra, P., Bhatt, D. K., and Singh, P. K. (2021). Pests detection using machine learning algorithms: A review. Computers and Electronics in Agriculture, 182, 106035.
13. Chen, X.-W., and Jeong, J.C. 2009. Sequence-based prediction of protein interaction sites with an integrative method. Bioinformatics 25, 585–591.
14. Costa, C., Schurr, U., Loreto, F., Menesatti, P., and Carpentier, S. (2019). Plant phenotyping research trends, a science mapping approach. *Frontiers in plant science*, *9*, 1933.
15. Cruz, M. S., G. W. A. Menezes, J. A. R. Silva, F. A. C. Oliveira, and M. G. da Silva. (2021). Biomass Prediction of Maize Plants Using Artificial Neural Networks and Hyperspectral Imaging. Computers and Electronics in Agriculture 183 (2021): 106042.
16. Daniel S., Kvasnicka, V., Pospichal, J. (1997). Introduction to multi-layer feed-forward neural networks, Chemometrics and Intelligent Laboratory Systems. 39(1), 43-62, ISSN 0169-7439.
17. Ding, H., Li, J., Zeng, X., and Li, Y. (2020). Prediction of protein secondary structure using long short-term memory neural networks. Computational Biology and Chemistry, 84, 107187.
18. Donaldson, R.G., M. Kamstra and H.Y. Kim. (1993). Evaluating alternative models for conditional stock volatility: Evidence from international data, Working Paper (University of British Columbia).
19. Dong, Y., Zhang, J., Ma, Y., and Sun, L. (2021). Estimation of rice protein content using deep belief network models. Computers and Electronics in Agriculture, 180, 105921.
20. Dusad, K., andDwivedi, V. D. (2020). Prediction of protein solubility using feedforward neural networks. Journal of Molecular Liquids, 304, 112698.
21. Erhan D., Bengio Y., Courville A, *et al*. (2010). Why does unsupervised pre-training help deep learning? *J Mach Learn Res.* 11:625–60.
22. Fathian, A., Amiri, B., & Marofi, S. (2019). Estimation of daily global solar radiation using artificial neural network model. International Journal of Environmental Science and Technology, 16(7), 3909-3922.

23. Gandhi, N., and Armstrong, L. J. (2016). A review of the application of data mining techniques for decision making in agriculture. *2nd International Conference on Contemporary Computing and Informatics (IC3I)* (pp. 1-6). IEEE.
24. Gao, L., Zhang, J., Cui, X., Liu, X., Chen, J., and Wang, S. (2021). A hybrid model based on artificial neural networks and fuzzy logic for prediction of soil moisture content in agriculture. Journal of Hydrology, 592, 125821.
25. Gribskov, M., McLachlan, A.D., and Eisenberg, D (1987). Profile analysis: Detection of distantly related proteins. Proc. Natl Acad. Sci. U. S. A. 84, 4355–4358.
26. Peter H. Sydenham and Richard Thorn (2005). Handbook of Measuring System Design.. John Wiley and Sons, Ltd. ISBN: 0-470-02143-8.
27. He, G., Zhu, P., Cao, Z., *et al.* (2004). Lyapunov exponents and chaotic regions of chaotic neural networks. Journal of Zhejiang University, 31(7), 387–390.
28. Hill, T., Marquez, L., O'Connor, M., and Remus, W. (1994). *Artificial neural network models for forecasting and decision making. International Journal of Forecasting, 10(1), 5–15.*doi:10.1016/0169-2070(94)90045-0.
29. Hornik, K., M. Stinchcombe and H. White (1989). Multilayer feedforward networks arc universal approximators. Neural Networks, 2(S). 3599366.
30. Imran, M., Zhang, Y., & Malik, R. N. (2020). Future climate change scenarios over Pakistan using the CMIP5 multi-model ensemble. Climate Dynamics, 54(1-2), 109-129.
31. Zhou, J. Q. Lu, R. Xu, L. Gui and H. Wang (2020). EL_LSTM: Prediction of DNA-Binding Residue from Protein Sequence by Combining Long Short-Term Memory and Ensemble Learning. IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 17, no. 1, pp. 124-135, 1.
32. Jamshidi, M., Sadeghi, T., Abdollahi-Arpanahi, R., and Mohtasebi, R., "Deep Learning Approach for Apple Tree Diseases Recognition: Transfer Learning Based Convolutional Neural Networks." Computers and Electronics in Agriculture 178 (2020): 105764.
33. Jime´nez J, Sˇ kalič M, Martı´nez-Rosell G, De Fabritiis G. (2018). KDEEP: protein-ligand absolute binding affinity prediction via 3D-convolutional neural networks. J Chem Inf Model. 58(2): 287–296. PMID: 29309725.
34. Jones, D.T. 1999. Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. 292, 195–202.

35. Zupan, J., Introduction to Artificial Neural Network (ANN) Methods: What They Are and How to Use Them, Acta ChimicaSlovenica 41/3/1994, pp. 327-352.
36. Kachwala, T., Bhardwaj, R., and Yadav, A. (2021). Yield Prediction of Wheat Crop Using Artificial Neural Network: A Study of Rajasthan, India. Communications in Computer and Information Science, 1431, 399-408
37. Khairunniza-Bejo, S., Mustaffha, S., and Ismail, W. I. W. (2014). Application of artificial neural network in predicting crop yield: A review. *Journal of Food Science and Engineering*, *4*(1), 1.
38. Klümper, W., and Qaim, M. (2020). A meta-analysis of the impacts of genetically modified crops. PloS one, 15(11), e0241543.
39. Kour, D., Kaur, T., Devi, R., Rana, K. L., Yadav, N., Rastegari, A. A., and Yadav, A. N. (2020). Biotechnological applications of beneficial microbiomes for evergreen agriculture and human health. In *New and Future Developments in Microbial Biotechnology and Bioengineering* (pp. 255-279). Elsevier.
40. Krizhevsky A, Sutskever I, Hinton GE. (2012). Imagenet classification with deep convolutional neural networks In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. Advances in Neural Information Processing Systems. p. 1097
41. Kumar, A., Singh, A. K., Kaur, S., and Kumar, V. (2021). Crop yield prediction in rice farming using multi-layer perceptron artificial neural network model. Agricultural Water Management, 243, 106513.
42. Lapides, D., Gupta, H., & Ravela, S. (2021). Quantifying uncertainty in continental scale wildfire predictions using Bayesian deep learning. Journal of Advances in Modeling Earth Systems, 13(7), e2020MS002400.
43. Lee, K. H., Yoo, J. H., and Park, H. J. (2021). Disease detection using machine learning: A review. Computers and Electronics in Agriculture, 188, 106303.
44. Li, J., Cai, J., and Liu, Y. (2020). Meteorological forecasting with machine learning algorithms: A review. Journal of Hydrology, 590, 125456.
45. Wessel, M.D., Jurs, P.C., Tolan, J.W., Muskal, S.M., Chem, J. (1998). Inf. Comput. Sci. 38. P. 726–735.
46. McCulloch, W. S., Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 5: 115–133.

47. Miguel A. Zuniga-Garcia, G. Santamaría-Bonfil, G. Arroyo-Figueroa, Rafael Batres.(2019). Prediction Interval Adjustment for Load-Forecasting using Machine Learning. Applied Sciences. DOI:10.3390/app9245269.
48. Mishra, S., Singh, R., and Singh, B. (2021). Groundwater level prediction using artificial neural network: A case study in India. Journal of Water Process Engineering, 39, 101727.
49. Mohanty, S. P., Hughes, D. P., and Salathé Jr, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7, 1419.
50. Montague, G., Morris, J. (1994). Neural-network contributions in biotechnology, Trends in Biotechnology. 12(8), 312-324, ISSN 0167-7799,
51. Mousa, W. A., Raheem, M. A., and Alhazmi, M. I. (2019). Prediction of wheat grains protein content using artificial neural networks. Computers and Electronics in Agriculture, 163, 104844.
52. Muneer, F., Zhang, X., Li, X., and Wang, C. (2021). Non-destructive prediction of protein content in wheat flour using near-infrared spectroscopy and partial least squares regression. Journal of Food Science and Technology, 58(2), 764-773.
53. Najarian, K., Splinter, R., (2005) *Biomedical Signal and Image Processing*. CRC Press, New York.
54. Osenblatt, F., (1958) The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev* 65: 386.
55. Pérez-Bueno, M. L., Pineda, M., Cabeza, F. M., andBarón, M. (2016). Multicolor fluorescence imaging as a candidate for disease detection in plant phenotyping. *Frontiers in plant science*, *7*, 1790.
56. Rani, M., Patel, C. B., and Kumar, S. (2020). A comparative study on artificial neural network and partial least square regression for the estimation of protein content in wheat. Journal of Food Science and Technology, 57(2), 406-414.
57. Raza, M. A., Iqbal, N., Zaheer, A., and Abbas, G. (2021). Machine learning-based biomaterials design for biomedical applications: A review. Journal of Biomaterials Science, Polymer Edition, 32(9), 1093-1123.
58. Reidys, C., Stadler, P.F. andSchuster, P. (1997). Generic properties of combinatory maps: Neutral networks of RNA secondary structures. *BltnMathcal Biology* 59, 339–397.
59. Agatonovic-Kustrin, S. neural networks. Applied Intelligence, 12(1–2), 15–25.

60. Agatonovic-Kustrin, S., Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. Journal of Pharmaceutical and Biomedical Analysis. 22, 717–727.
61. Hochreiter, S., and Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation. 9(8), 1735-1780.
62. Sahi, A.K., M. Kumar, M. Singh, and V.K. Srivastava (2020). Prediction of Rainfall in India Using Artificial Neural Networks. Neural Computing and Applications. 32(23), 17327-17338.
63. Sarkar, A., and Azuaje, F. (2021). Protein-protein interaction prediction using recurrent neural networks. BMC Bioinformatics, 22(1), 40.
64. Seonwoo, M., Byunghan, L., Sungroh, Y. (2017). Deep learning in bioinformatics, *Briefings in Bioinformatics*. 18(5), 851–869.
65. Shimon, S., Gad, A. (1994). Neural networks for decision support: Problems and opportunities; Decision Support Systems. 11 (5), 393-414, ISSN 0167-9236
66. Sharma, S., Sharma S. (2020). ACTIVATION FUNCTIONS IN NEURAL NETWORKS, International Journal of Engineering Applied Sciences and Technology. 4(12) 310-316. ISSN No. 2455-2143.
67. Spencer, M., Eickholt, J., Cheng, J. (2015). A deep learning network approach to ab initio protein secondary structure prediction. IEEE/ACM Trans compatible Bioinformat.12 (1):103–12.
68. Suresh, R., Prasad, V. K., Babu, B. K., and Prasad, G. (2021). Crop yield prediction using artificial neural network models. Communications in Computer and Information Science, 1430, 290-297.
69. Tam, K.Y. and Kiang, M.Y., (1992). Managerial Applications of Neural Networks: The Case of Bank Failure Predictions, Management Science, 38, 926-947
70. Tam, K.Y. (1991). Neural network models and the prediction of bank bankruptcy, Omega, The International Journal of Management Science. 19, 429-445.
71. Wang, X., Sun, G., and Zhou, G. (2020). A comprehensive comparison of machine learning methods for predicting soil properties. Soil and Tillage Research, 197, 104512.

72. Wu, J., Wu, Z., Yan, Y., Wang, Z., and Zhang, J. (2021). Protein-ligand binding affinity prediction using machine learning: A review. Current Topics in Medicinal Chemistry, 21(16), 1515-1527.
73. Wu, Yc., Feng, Jw. (2018). Development and Application of Artificial Neural Network. *Wireless PersCommun* 102, 1645–1656.
74. Xia, G., Zhang, J., Li, W., Qiu, Z., Li, J., and Li, Y. (2021). Identification of Weeds in Soybean Crops Based on Hyperspectral Imaging and Artificial Neural Network. Computers and Electronics in Agriculture 184: 106063.
75. Yoo, S. H., Park, J. W., and Jung, K. H. (2016). A deep learning network approach to abiotic stress detection and prediction in maize. Computer and Electronics in Agriculture, 124, 236-241.
76. Yoon, Y. and Swales, G. (1990). Predicting stock price performance, Proceedings of the 24th Hawaii International Conference on System Sciences. 4, 156-162.
77. Doerr, S., Martı´nez-Rosell, G., Rose, A.S., De Fabritiis, G. (2019). DeepSite: Protein-binding site predictor using 3D-convolutional neural networks. Bioinformatics. 33(19): 3036–3042.
78. Zhang, J., *et al*. (2017). Artificial neural network-based analysis of amino acid sequences for prediction of protein stability change. Journal of Biomolecular Structure and Dynamics, 35(3), 443-451.
79. Zhang, L., Li, X., Zhang, S., and Li, H. (2020). Machine learning approaches in pesticide efficacy prediction: A review. Science of The Total Environment, 703, 135538.