Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 4s (2025)

Enhancing Crop Yield Prediction using Linear Models and Deep Learning Techniques in Precision Agriculture

B. Sunitha Devi[1], K. Shahu Chatrapati[2], N. Sandhya[3]

[1] Department of CSE, CMR Institute of Technology, Hyderabad, India. sunithadevi.b2022@gmail.com

[2] Department of CSE, JNTUH, Hyderabad, India. shahujntu@gmail.com

[3] Department of CSE, VNR Vignan Jyothi Institute of Engineering &Technology, Hyderabad, India. sandhyanadela@gmail.com

Article History:

Received: *23-09-2024*

Revised: *25-11-2024*

Accepted: *08-12-2024*

Abstract:

The substantial advancements in computer science and engineering have sparked interest in precision agriculture and led to the development of more advanced instruments and methods for enhancing farming practices. This study focuses on the use of machine learning and mathematical models for fertilisation optimisation and yield prediction as part of a Precision Agriculture strategy. In particular, provides the outcomes of forecasting winter wheat production and protein content across four farms using the amounts of nitrogen fertiliser sprayed on the fields. To maximise net yields on the next crop, fertiliser treatments have to be prescribed based on these projections. In particular, contrast approaches based on neural networks (deep and shallow) and multiple regression (linear and non-linear). The greatest results are obtained by a deep neural network that incorporates spatial sampling and is based on the stacked autoencoder, according to the findings.

Keywords: Precision Agriculture, agriculture practices, mathematical frameworks, yield, winter grains.

1. Introduction

The quantity of fertiliser derived from petroleum that is wasted has greatly grown in recent years. The recent rise in algal blooms, which cause soil to become more toxic, may have this increase as one of its primary reasons, according to ecologists. The pace of carbon storage in farmland soil is also thought to be affected by climate change. In light of changes in the climate and shorter growing seasons, which are more relevant to the present investigation, farmers are trying to find innovative methods to increase their earnings. One of the main reasons people work is for the second reason. Precision Agriculture draws on cutting-edge technology, generally from the fields of engineering and computer science to solve these issues and facilitate well-informed decision-making in the agriculture sector.

Two features of Precision Agriculture have the potential to solve these problems. The first is to determine the best fertiliser rates and then use that information to make predictions that will help farmers save money and use less fertiliser. However, this can only be achieved if the yield and protein concentrations are anticipated in light of weather conditions and the field's present and past characteristics. That brings the second point and the main point of this study.

In this publication, the researchers lay out their strategy, which involves using machine learning methods to forecast yield and protein content in particular areas of interest. A multitude of machine learning fields, including medical diagnosis, forecasting the stock market, and natural language

processing, have shown an increasing interest in the use of neural network algorithms, particularly within the framework of deep learning. With ANNs, the network's "neurones" analyse the input signals independently and concurrently, which allows them to learn and recognise patterns from various input signals. This is one of the most intriguing benefits of ANNs.

A more recent innovation is the use of ANNs in precision agriculture. There are two main ways to use ANNs in this context: one models field attributes at the local level, while the other expands the inputs to include spatial context. Examine the differences and similarities between training basic feedforward neural networks and stacked autoencoders, a kind of deep learning model. Next, review the outcomes of both neural nets about multiple linear and multiple nonlinear regression models. Researchers found that deep and shallow neural netS models outperformed the more conventional regression techniques. Additionally, discovered that including the geographical context results in a statistically significant performance improvement.

The remaining sections of the article are organised as follows. The second section discusses related work. Section 3 delves more into each of the models. Furthermore, this part gives some history of Precision Agriculture. Part 4 describes in-depth experimental strategy, experimental findings, and analyses of these findings. Lastly, Section 5 presents potential areas for further investigation, and Section 6 provides a conclusion.

2. Related Work

An area known as precision agriculture (PA) has recently attracted a lot of interest. A growing number of studies are turning to artificial neural networks (ANNs) to forecast agricultural production and other related outcomes, as opposed to the more traditional linear and nonlinear regression models. Maximise Fertilisation. An outline of the most important PA-related studies using the models mentioned above is provided in the following sections.

2.1 Models: Linear and Non-Linear

For empirical model development using massive data sets, one popular approach is stepwise multiple linear regression (SMLR). When growing maize under conditions of nitrogen and water stress, this is what SMLR uses to predict grain production. In particular, they anticipated chlorophyll meter values, grain nitrogen concentration, grain yield, plant nitrogen, and biomass by applying the REG process in SAS to the stored data. To what extent ANNs outperform SMLR is something they investigate. The authors identify three significant disadvantages associated with the use of SMLR. Initially, the foundation of support vector load regression (SMLR) is the assumption that the response and input variables are linearly related. Secondly, it is assumed that the noise included in the samples has a normal distribution. Lastly, there are instances when the model overfits, which diminishes its capacity to apply to unobserved data. Nonetheless, their findings indicate that ANNs and SMLR exhibit comparable performance. Due to their ambiguous findings, the authors stress the need for more studies utilising both methods.

2.2 Machine Learning

Machine Learning has facilitated comprehensive study across several disciplines. Training artificial neural networks is a prevalent method in machine learning and has been used for many biological and

agricultural challenges. An intriguing and beneficial characteristic of artificial neural networks (ANNs) is their capacity to identify complex relationships among input and response variables, without the need of pre-establishing limitations for the distribution of samples of information. This facilitates the characterisation of intricate non-linear interactions often seen in areas like PA, arising from diverse crop conditions and several impacting variables. Research on the application of ANNs in agriculture, particularly in the areas of yield prediction and fertilisation optimisation, is the primary emphasis of this section. Summarise studies that have effectively used ANNs in the field of biology and agriculture to get a broader picture of their potential uses. To forecast yield, they train artificial neural networks in various ways using the same dataset. To find the best approach, they prioritise features such as outlier rejection, test set generalisability, and training accuracy. For their training strategies, the authors considered backpropagation as well as without weight decay, in addition to quick prop and r prop. The results showed that r prop, as a learning strategy, was marginally more effective than backpropagation, and that any kind of training and learning beat the linear method. In fertilisation models, they added a group of feedforward ANNs. They assert that while utilising a single feed-forward network in this scenario, there are two typical problems. Initially, they discovered that if a maximum goal yield is assigned before to computation, there would be a significant mistake in the fertilisation rate calculations. Secondly, they think that their model's predicting accuracy and generalisation ability are constrained when it uses a single ANN.

The researchers suggested employing an ANN with yield as the output and nutrient concentration and fertilisation rate as the input. Using a bagging approach, they trained multiple neural networks using backpropagation. Then, they clustered the networks using k-means. The next step in assembling a group of ANNs was to choose a single network that exemplified each cluster. Using the Lagrange multiplier approach, which is connected with constraints that require the weights to add to one, they calculated the ensemble's combining weights. Then, using the ensemble's output as a basis, they defined a nonlinear objective function and used nonlinear programming to optimise the rates of fertiliser delivery. Improvements in using ANNs for PA have been substantial since the aforementioned studies. To determine the optimal crop for a given field given a variety of soil and atmospheric conditions, they used a feed-forward neural network trained using backpropagation. Temperature, precipitation, pH, nitrogen, and potassium levels were among the input characteristics they used. In addition, they foretold the optimal fertiliser rate that would benefit the harvest. For farmers without access to costly soil testing equipment, their projections were a lifesaver. The findings revealed quite accurate forecasts, which might be of great assistance to these farmers. The NDVI, water stress index, canopy surface temperature, and absorbed photosynthetically active radiation are the inputs used by their Support Vector Regressor and DNN.

There has been the application of a deep neural net. Their DNN and SVR are fed data on water stress index, canopy surface temperature, absorption photosynthetically active radiation, and normalised difference vegetation index (NDVI). By using a CNN and a Long-Short Term Memory net for the sorting of histograms formed from rarely sensed pictures, the DNN can attain RMSE values of about for maize yield prediction, while the SVR does not. In terms of total RMSE values, the CNN performs the best.

To evaluate measurement data, examine sensor information, and update basic knowledge, data mining—a distinct area of study—often employs methods based on Supervised Self-organising Maps. The model's input nodes, according to their proposal, would represent the primary components of grain crop production. Using neural networks to forecast wheat yield using publicly accessible, but often of poor quality, in-season data, the model categorised the data to forecast productivity and wheat output. To maximise the effectiveness of fertiliser, the writers also used data mining techniques including neural networks. Various networks were constructed and tested; the results showed that, as more data became accessible, the networks' prediction accuracy improved.

3. Background

An overview of Precision Agriculture will be provided before delving into the specifics of the methodology suggested in this article.

3.1 Agricultural Expertise

The goal of precision agriculture (PA) is to enhance crop yields via the strategic use of different technological tools. The area began with the invention of the Global Positioning System, which allowed robots to manage remedies with localised needs for each field by determining coordinates anywhere on Earth. Several fields of engineering and computer science have recently seen significant growth The next sections will discuss the subfields of PA that have been studied, namely yield mapping and fertilisation optimisation.

3.1.1 Mapping Yields

The process of yield mapping entails visualising agricultural output for a certain region about a given geographic location. Physical observations or yield estimates derived from computer models may form the basis of these maps. The four main types of yield maps are inference, prediction, interpolation, and aggregation. When establishing a target yield, for instance, this is helpful when working with a soil map. On the prediction map, any value where the yielding element is expected to be substantially quantifiable is filled up. Produced by measuring yield at discrete points within a predefined region, interpolation maps are more coarsely detailed than prediction maps. Afterwards, a local estimation technique is used to estimate the yield values among data points. Ultimately, by measuring or predicting the source data, aggregation maps provide aggregated statistics. In PA, data aggregation and prediction are the most often used yield mapping techniques out of the four. The three measurements needed for yield mapping—the yield measurement, an area of measurement, as well as the precise location of measurements within this area—allow for the quantification of grain, harvested crop volume or mass according to position within a specified field. Yield mapping is intrinsically linked to our goal of optimising fertilisation since it often serves as a foundation for determining the optimal fertilisation rate for fields.

3.1.2 Optimisation of Fertilisation

Over the last 50 years, the rise in agricultural productivity has led to a substantial increase in fertiliser application. The increase in fertilisation leads to heightened output; nevertheless, similarly escalates agricultural emissions, particularly nitrogen emissions, in both groundwater and surface water. The problem, along with enhancing nitrogen application to boost yield and therefore profit, prompted

a study in fertilisation optimisation. The predominant technique for enhancing fertilisation application is the use of variable rate technology. Their summary provides a synopsis of the many precision agriculture fertilisation experiments. Nutritional management study examining the effects of fertilisation on crop yields is the subject of their first literature review. When it comes to site-specific management, the authors discovered that the majority of studies support the idea that applying at different rates yields better results than employing a uniform rate. To find out how much nitrogen to apply to specific areas of a field, variable rate technology uses several optimisation models. Instead of applying a uniform rate to the whole field, it determines the optimal nitrogen rates for every plot and utilises them as a foundation for return prediction. The mechanics of sprayers as well as spreaders that can apply these various rates, however, suffer more when rates are altered often or if variations entail very big changes. These devices are also more costly. The fact that the authors provide two trials with uniform rate application yielding larger profits is thus not surprising, especially considering that these researchers were directed when the device was less comfortable.

3.2 Regression: Linear and Non-linear

One way to define potential correlations between variables in statistical analysis is multiple regression, which employs mathematical procedures. In linear regression, the response values are best predicted by fitting a straight line across the data. In addition to helping explain or evaluate a scientific theory, this approach is often used for making predictions about the values of one variable given another. Similar to linear regression, nonlinear regression uses non-linear surface forms instead of a linear one. The form might vary substantially since it is governed by the general dispersion of the data. This study's non-linear model estimates yield as well as protein values using a hyperbolic function.

3.3 Networks of Neutrals

An architecture for distributed and parallel computing called an artificial neural network mimics the brain's natural operation. The building's processing components are linked via bidirectional or unidirectional signal stations. All of the processing parts have their local memory, so they can only process data that is local to them. The present status of the input signal and the stored values in this reminiscence are the only inputs that an element takes into consideration during processing. All the components have one output, which may be any mathematical signal, and as many collateral connections as are needed. Here, go into more depth about the feed-forward method and present the idea of a stacked autoencoder, a deep approach that is not used for PA.

3.3.1 The feed-forward neural network

Every node in a feed-forward ANN contributes data to the final output layer by passing it forward via each following layer. To find a node's output in a network, utilise the following approach:

$$y = f\left(w_0 + \sum_{i=1}^{n} w_i\, x_i\right)$$

Where f(.) is often referred to as an “activation function”. Given z= $w_0 + \sum_{i=1}^{n} w_i\, x_i$, these activation functions may be linear (f(z) = z), rectified linear (f(z) = max, z), logistic (f(z) = 1/(1+exp (-z)), hyperbolic tangent (f(z) = tanh z), or radial (f(x) = exp(||x - c||/σ)).

The modifications to the weights inside the network's interior are determined by the backpropagation (BP) method. Figure 1 is an example of a feed-forward artificial neural network demonstrating the backpropagation process. Backpropagation is a widely used method for training neural networks. The Parallel Distributed Processing (PDP) grouping at Stanford presented it to a broad audience. Although other researchers found BP earlier, the PDP lab is often credited with developing the first practical way to use backpropagation, which is why they are seen as the ones responsible for the present approach to BP. In BP, the goal is to minimise the network's mean squared error as a loss function, and the updates are related to:

$$\Delta w_{ij} = \frac{\partial Err(x)}{w_{ij}} = \eta \delta_i x_{ji}$$

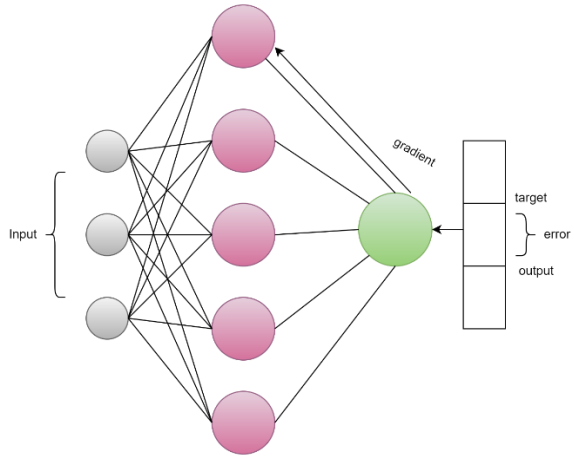

Figure. 1 Backpropagation is used to train a feed-forward neural network.

where at the output layer,

$$\delta_j = (r_j - y_j) y_j (1 - y_j)$$

where an input from the i^{th} nodes to the j^{th} node is x_{ij}, The intended response value is r_j. The inner layers

$$\delta_j = o_j (1 - o_j) \qquad \sum_{k \in downstream(j)} \delta_k W_{kj}$$

and o_j The output produced by the j^{th}Nodes are represented. This update rule is predicated on a logistic activation function, so keep that in mind. The updating rule will be somewhat modified by other activation functions due to the varied derivatives being computed.

3.3.2 Automatic Stacked Encoder

This paper's study makes use of a stacked autoencoder (SAE), a distinct kind of ANN that employs deep learning for yield prediction. The idea of deep learning originated from the belief that improved generalisations on complicated recognition tasks may be achieved by increasing the modelling of very nonlinear interactions among the variables. This would need multiple layers of abstractions in feature space. One artificial neural network (ANN) that aims to replicate input-to-output mapping using a

hidden layer to encode input is the autoencoder. The first step of an autoencoder is to encode the input with the hidden layer; the second is to decode the encoded values and use them to rebuild the input. The autoencoder model learns valuable data characteristics since it is designed to prioritise specific input qualities to be duplicated. This is the first step in creating a stacked autoencoder. Firstly, train the only autoencoder using backpropagation, using the inputs as both the goal values and the training set. In the second phase, an additional hidden-layer net is superimposed on top of the current autoencoder. This network strives to recover the hidden values that were created by the previous autoencoder's hidden layer. The input layer of the new network is mapped to those values. A "stacked" autoencoder is the product of this procedure repeated while linking the learnt autoencoders (Figure 2).

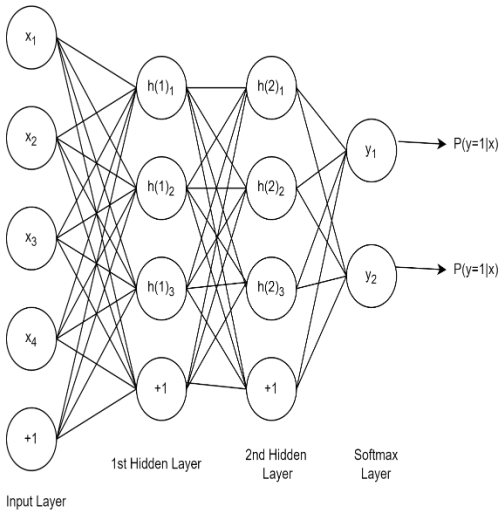

Figure. 2 A softmax classifier combined with a stacked autoencoder

The idea behind stacking autoencoders is that it creates a hierarchical abstraction of characteristics for the data under analysis, as each hidden layer corresponds to a set of detectors for features that target the signals of the preceding layer. The last stage in stacking autoencoders for training is to put another learning-focused network on top. After this is complete, by training the weights collectively, the ideal mapping between the initial set of inputs and the desired yield/protein levels might be discovered.

3.4 Geographic Sample

The technique known as "spatial sampling" involves gathering data in a two-dimensional area. As many geographical variations of the study variables as possible are often accounted for in the sampling strategy. Additional measurements may be taken after the first data has been collected and recorded, depending on the data's fluctuations. Criteria to optimise the data often inform these further measures. Because certain factors are more important in some regions than others, or because of differences in topography, the final data obtained may be skewed. In comparison to data that is evenly spaced, this data is far more difficult to analyse due to its uneven spacing. Thus, it is crucial to handle these anomalies to conduct the dataset analyses more effectively. Because the fields researched & the prescription maps generated for them are grid-based, a kind of spatial sampling is inherent to this study because of the grid-based layout of the fields. Since each grid cell contains data points, it is reasonable to sample from both the target cell and its nearby neighbours. Both the von Neumann and Moore

neighbourhoods are typical for this kind of sample (Figure 3 left and 3 right, respectively). Our spatial sampling method relies on the Moore neighbourhood because it gives a more complete picture of the surrounding samples. For each given cell, we averaged the values of all the points within. The location information for every point of data is derived from the mean values of the eight cells immediately around it. Two data adjustments were used to reduce noise and smooth the information for more consistent spatial information. One was to use the average method already discussed, and the other was to make the grid cells larger so that more points could fit into each cell.

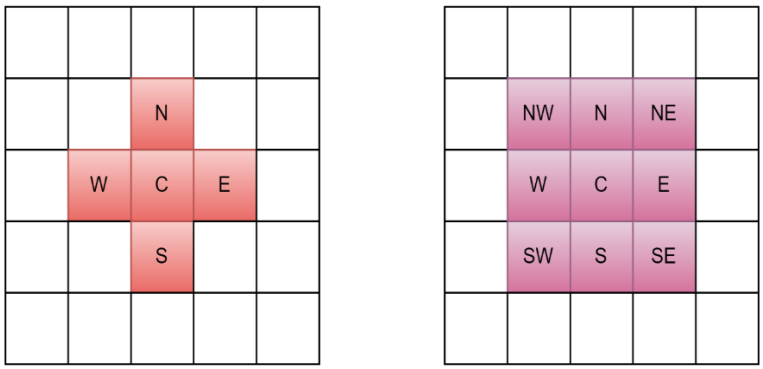

Figure. 3 Sample settings for neighbourhoods. "C" is the symbol for the centre cell. The neighbourhoods of von Neuman and Moore appear on the left and right, respectively.

4. Experiments

Section 4.1, covers the specifics of evaluating several ML and SS methods. Section 4.2 displays and compares the results of the four well-established methods: artificial neural networks (ANN), support vector machines (SSE), not linear regression, however, and line regression. The next stage is to compare the k-nearest neighbour algorithm-sampled data with non-sampled data about its geographic distribution. These results are discussed and explained in Section 4.3, which is the last section.

4.1 The approach

Utilising the fertiliser rate as a variable, a basic linear regression across the protein & yield points is utilised. A hyperbolic curve was fitted across the data using the nonlinear model. Because there is a saturation threshold for fertiliser rates beyond which yield & protein values stop increasing, a hyperbolic fit is used. Create a total of 6 models by implementing and evaluating both spatial and non-spatial versions of the ANN as well as the SAE. It was determined which design yielded the greatest results by experimenting with various settings. A few examples of these parameters include the total number of ANN and SAE epochs, the total number of layers that are hidden, and the total number of nodes that are hidden within each layer. Depending on the field's accessible features, the ANN's optimum performance was attained with just one hidden layer containing 15 – 100 hidden nodes; however, these values were changed for each dataset. For non-spatial data, the SAE uses two hidden layers; for spatial data, it uses three, with the number of hidden nodes decreasing with each successive layer. Starting with the top layer and working down the way, utilise 500, 250, and 125 concealed nodes for the spatial information, respectively. The consistency of the number of concealed nodes for non-spatial information varied with the available data points. Two distinct metrics, the Root Mean Squared

Error and the coefficient of determination (R^2), were utilised to assess each model. Utilise the 10-fold cross-validation to determine each of these, and then average the findings over all of the folds. Researchers performed a paired Student t-test with Bonferroni alteration to look at the difference in means.

Utilising the information gathered from four separate fields, all of which belonged to separate farmers. Produce and protein levels for a given spot on the field, together with details about that spot's location and slope, nitrogen application rates, observed precipitation, and Normalised Difference Vegetation Index (NDVI) from prior years make up the data. Based on the additional data that is given, the models that have been put into place are attempting to forecast yields and protein levels. Both the yield and the protein values were from already-harvested fields as the recommended map was made using information from previous harvest bins.The cell densities in prescription charts might vary according to the field and grid size. Table 1 shows that there is a lot of variation in the data about protein and yield points for each location. Since only locations with eight neighbouring cells containing data are taken into account throughout the spatial sampling method, the amount of information gradually decreases.

Table.1 Data quantity for every field.

	Sec35mid	**Sre1314**	**Davidson MW**	**Carlin W**
Protein points	1020	2999	561	656
Yield points	17874	24647	11803	15622

4.2 Result

Table 2 displays the root-mean-squared error (RMSE) for yield predictions, whereas Table 2displays the results for protein predictions. Table 3 shows that the non-linear system did not offer any outcomes for protein prediction. The boldface indicates statistically significant differences in RMSE values.

Field	**Measure**	**Linear**	**Non-Lin**	**ANN**	**ANN Sp**	**SAE**	**SAE Sp**
Sec35mid	RMSE	10.17	10.02	10.65	**8.98**	10.08	**8.96**
	R^2	30.04	52.67	24.08	45.16	53.96	65.79
Sre1314	RMSE	11.24	11.26	11.52	11.06	10.88	10.16
	R^2	1.38	14.79	9.17	20.68	17.68	32.68
david	RMSE	15.26	15.93	16.37	**12.48**	16.04	**12.07**
	R^2	38.95	36.94	30.47	50.65	35.17	59.36
carlin	RMSE	12.08	10.21	10.24	**8.07**	9.94	**9.16**
	R^2	1.46	6.78	-4.31	16.18	11.48	9.69

Table. 2 Forecasted yields per field. Emphasised are the notable variations. For geographic data, the sign "Sp" is equivalent.

Pitch	**Measure**	**Linear**	**Non-Lin**	**ANN**	**ANN Sp**	**SAE**	**SAE Sp**
Sec35mid	RMSE	1.41	N-A	1.45	1.23	1.98	1.97
	R^2	24.61	N-A	23.14	49.73	0.79	17.42
Sre1314	RMSE	1.23	N-A	1.17	0.91	1.16	1.14
	R^2	-56.97	N-A	-41.38	11.65	0.42	5.71

david	RMSE	1.69	N-A	1.61	1.48	1.68	1.62
	R^2	-32.21	N-A	-15.69	-9.61	0.06	0.06
carlin	RMSE	1.18	N-A	1.19	1.14	1.38	1.28
	R^2	11.16	N-A	11.13	16.46	4.83	0.49

Table. 3 Protein prediction for all areas. Bold text indicates statistically significant differences. "Sp" represents spatial data.

4.3 Discussion

The results of this study on yield prediction indicate that geographical data often enhances a model's precision. Aside from one area, the spatial ANN & spatial SAE exhibit much superior performance compared to the other models. The absence of advancement in this area may result from the distribution of yield sites; these spots are densely clustered along the field's length but are much more dispersed throughout its breadth compared to other fields. This successfully augments the point density inside a single grid cell while reducing the overall number of cells considered for the spatial information. In general, the application of machine learning methods to spatial information surpasses conventional methods in generating predictions. Further investigation into both regression algorithms and geographical sampling methods may give valuable insights into yield prediction as well as improve precision. The findings for protein prediction provide a contrasting narrative. While enhancements are consistently seen in the outcomes for the same models using geographical data, the findings do not exhibit substantial differences from one another. Nonetheless, the ANNs exhibited markedly superior performance compared to the other models in two of the domains. Notably, the two fields have the largest and lowest quantities of protein points. A thorough examination of the information could give more understanding of the source of this phenomenon. Also, the SAE isn't very good at this, particularly when contrasted with how well it does with yield prediction. Since SAEs do better with a large amount of data, the most reasonable explanation is that there aren't enough data points to train the protein models, which means they can't learn very well. On the other hand, when we examine the RMSE values, we can see that the predictions made by each model are quite accurate. It seems from the data that protein values are not significantly different from yield values, which might make it simpler to anticipate the right protein values and ultimately lead to decreased RMSE values.

5. Conclusion

Precision Agriculture may significantly enhance agricultural yields by using several mathematical methods & experimental methodologies. This field of study has realized a rise in computational investigation, particularly in machine learning. Utilising machine learning to develop predictive models may enhance the accuracy of fertilisation prescription maps and yield/protein forecasts, hence improving production. This notwithstanding the extensive prior study on machine learning, there has been less investigation into the use of deep learning methodologies in precision agriculture. To create productivity and protein prediction techniques, this study examines multiple regression, a shallow feedback net, and stacking automatic encoders in geographic and non-spatial situations. Compared to the other methods tested, our results show that feed-forward net and spatial stacked autoencoders provide much better accuracy in the studied domains.

6. Future work

Equal Distance & Random Selection are two more techniques of sample spatial framework that are not covered in this book but are certainly worthy of investigation. They are interested in determining how well various forms of spatial context perform when included. Given that the ANN and SAE performed similarly, more research into their inner workings is necessary. It could be helpful to examine whether an autoencoder can be adjusted to better fit this particular challenge as the presently used stacked autoencoder was first developed for another study. The data being analysed determines how the ANN and SAE structures are modified. Stated differently, distinct models are established for every distinct area. "Transfer learning," or the application of information acquired in one discipline to another, is a topic of study that is now getting attention. Investigating how this could reduce training complexity or perhaps raise overall exactness would be intriguing. The data utilised to forecast the yield & protein levels is another factor to take into account for further research. It may be useful to look at which values have the most effects on exact forecast whereas others have little to no effect. Further understanding of the systems and how they forecast the yield, as well as protein values, may be gained by examining how particular values affect the various system types and contrasting the variations in the effect of those values for individual types. This research only focuses on predicting protein and yield, particularly in connection to fertiliser application. As before said, forecasting the ideal fertilisation rate is a crucial component of our research, and examine the degree to which may utilise these models to provide optimised fertilisation prescriptions.

References

[1] Awad, Mohamad M. "Toward precision in crop yield estimation using remote sensing and optimization techniques." *Agriculture* 9.3 (2019): 54.

[2] Rama Devi, O., et al. "Optimizing Crop Yield Prediction in Precision Agriculture with Hyperspectral Imaging-Unmixing and Deep Learning." *International Journal of Advanced Computer Science & Applications* 14.12 (2023).

[3] Aworka, Rubby, et al. "Agricultural decision system based on advanced machine learning models for yield prediction: Case of East African countries." *Smart Agricultural Technology* 2 (2022): 100048.

[4] Kuradusenge, Martin, et al. "Crop yield prediction using machine learning models: Case of Irish potato and maize." *Agriculture* 13.1 (2023): 225.

[5] Rao, M. Venkateswara, et al. "Brinjal Crop yield prediction using Shuffled shepherd optimization algorithm based ACNN-OBDLSTM model in Smart Agriculture." *Journal of Integrated Science and Technology* 12.1 (2024): 710-710.

[6] Abbas, Farhat, et al. "Crop yield prediction through proximal sensing and machine learning algorithms." *Agronomy* 10.7 (2020): 1046.

[7] Wei, Marcelo Chan Fu, et al. "Carrot yield mapping: A precision agriculture approach based on machine learning." *Ai* 1.2 (2020): 229-241.

[8] Sishodia, Rajendra P., Ram L. Ray, and Sudhir K. Singh. "Applications of remote sensing in precision agriculture: A review." *Remote sensing* 12.19 (2020): 3136.

[9] Akhter, Ravesa, and Shabir Ahmad Sofi. "Precision agriculture using IoT data analytics and machine learning." *Journal of King Saud University-Computer and Information Sciences* 34.8 (2022): 5602-5618.

[10] Agarwal, Sonal, and Sandhya Tarar. "A hybrid approach for crop yield prediction using machine learning and deep learning algorithms." *Journal of Physics: Conference Series*. Vol. 1714. No. 1. IOP Publishing, 2021.

[11] Gopal, PS Maya, and R. Bhargavi. "A novel approach for efficient crop yield prediction." *Computers and Electronics in Agriculture* 165 (2019): 104968.

[12] Gómez, Diego, et al. "Potato yield prediction using machine learning techniques and sentinel 2 data." *Remote Sensing* 11.15 (2019): 1745.

[13] Medar, Ramesh, Vijay S. Rajpurohit, and Shweta Shweta. "Crop yield prediction using machine learning techniques." *2019 IEEE 5th international conference for convergence in technology (I2CT)*. IEEE, 2019.

[14] Ju, Sungha, et al. "Optimal county-level crop yield prediction using MODIS-based variables and weather data: A comparative study on machine learning models." *Agricultural and Forest Meteorology* 307 (2021): 108530.
[15] Coulibaly, Solemane, et al. "Deep learning for precision agriculture: A bibliometric analysis." *Intelligent Systems with Applications* 16 (2022): 200102.
[16] Sunitha, Gurram, et al. "Modeling of chaotic political optimizer for crop yield prediction." *Intelligent Automation and Soft Computing* 34.1 (2022): 423-437.
[17] Kumar, Y. Jeevan Nagendra, et al. "Supervised machine learning approach for crop yield prediction in the agriculture sector." *2020 5th International Conference on Communication and Electronics Systems (ICCES)*. IEEE, 2020.
[18] Radočaj, Dorijan, Mladen Jurišić, and Mateo Gašparović. "The role of remote sensing data and methods in a modern approach to fertilization in precision agriculture." *Remote Sensing* 14.3 (2022): 778.
[19] Abdel-salam, Mahmoud, Neeraj Kumar, and Shubham Mahajan. "A proposed framework for crop yield prediction using hybrid feature selection approach and optimized machine learning." *Neural Computing and Applications* (2024): 1-28.
[20] Elavarasan, Dhivya, and P. M. Durai Raj Vincent. "Fuzzy deep learning-based crop yield prediction model for sustainable agronomical frameworks." Neural Computing and Applications 33.20 (2021): 13205-13224.
[21] Wang, Xinlei, et al. "Winter wheat yield prediction at county level and uncertainty analysis in main wheat-producing regions of China with deep learning approaches." *Remote Sensing* 12.11 (2020): 1744.
[22] Patro, Pramoda, et al. "A hybrid approach estimates the real-time health state of a bearing by accelerated degradation tests, Machine learning." 2021 Second International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE). IEEE, 2021.
[23] Gopi, P. S. S., and M. Karthikeyan. "Multimodal machine learning based crop recommendation and yield prediction model." *Intell Autom Soft Comput* 36.1 (2023): 313-326.
[24] Richetti, Jonathan, et al. "A methods guideline for deep learning for tabular data in agriculture with a case study to forecast cereal yield." *Computers and Electronics in Agriculture* 205 (2023): 107642.
[25] Kok, Zhi Hong, et al. "Support vector machine in precision agriculture: a review." *Computers and Electronics in Agriculture* 191 (2021): 106546.