International Journal of Advanced Science and Technology
Vol. 29, No. 9s, (2020), pp. 3193-3203

Deep Learning based Recommendation System for Profitable Agricultural Plantation

V.Ramachandran, R.Ramalakshmi*

Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education
Krishnankoil-626126, India
Email: ramachandran.veer@gmail.com, rama@klu.ac.in

Abstract

Crop planting patterns are influenced by factors like climatic conditions, yield prediction, availability of resources, profitability, and farm area. Unless rising production costs are matched by the market value of an agrarian commodity, farmers incur mounting losses or debt. Predicting the market value of a crop to be gleaned at harvest time is a key decision for a farmer hoping to profit from his hard work. A farmer's decision to plant crops is based on the current market price of a product, rather than its anticipated future price, which can, at times, result in losses rather than profits. This paper proposes a recommendation system that will enable farmers to predict the market value of a planted crop. The recommendation system uses Recurrent Neural Networks with Long Short-Term Memory (LSTM) cells to predict the price of the crop to be gathered in during the harvest. The proposed system is analyzed with data from the data site (www.data.gov.in)of the Government of India.The system was able to predict crop prices during the harvest accurately, and will help farmers choose the right crop, intended for high profitability, to be planted.

Keywords: *Deep Learning, Recurrent Neural Networks, Crop Price Prediction and Precision Agriculture.*

1. Introduction

India, with 179.8 MHa of net cropland, stands first in terms of global net cropland area (9.6 percent of the global area) according to the United States Geological Survey 2017. Agriculture in India contributes nearly 17% of India's gross domestic product (GDP), and around 50% of India's workforce is directly or indirectly connected to agriculture. Increased production costs over the years, coupled with the price instability for commodities in the agrarian market, are major problems for small-scale farmers in developing countries like India.A sustainable yield is fundamental to cropping systems [Fowler,(2016)]. Harvesting and market demand are unpredictable in certain circumstances. A low demand causes the market price of a particular agricultural commodity to crash, making survival impossible for farmers. As per the NCRB report of 2015, bankruptcy or indebtedness and farming-related issues have been reported as major causes of suicides among farmers/cultivators [Merriott,(2016)]. All the more reason, then, for a crop suitability recommendation system that plays a significant role in making agriculture profitable. In India, a crop recommendation system based on available resources and farmers, who are able to select the right crop to be planted for a good yield, have long practiced soil type. Farmers evinced interest in planting water-intensive crops with high profit margins [Ji, (2018)]. A seemingly unrelated regression model was used to study factors driving crop cultivation in the Hong he Hani Rice Terraces.In the current market, rural family units tended to plant crops yielding high returns. This was because they judged profits based on crop yields per unit, rather than market trading trends [Liverpool-Taise, (2017)].

Farmers need to be made aware of optimum crop suitability, or crop design. Given that the volatility/unpredictability of the agricultural commodity market is a major concern, increased production from the field does nothing to generate the revenue farmers expect, culminating in acute agrarian distress.Price fluctuations and future price predictions are unfamiliar terrain for Indian farmers. For the most part, farmers decide on a suitable crop for future planting based on its current market price, which could likely fluctuate at harvest time, generating less revenue than expected and having farmers stare mounting losses in the face [Hu,(2016)].The demand for raised Minimum Support Prices and future contracts with farmers clearly depict the uncertainty in the agricultural commodity market. The agricultural market, with 17% GDP of the country, it affects the economic sustainability of the farmer hugely. Predicting the price of a crop planted for harvesting is crucial for farmers, especially the small-scale ones. The proposed system in this paper predicts future prices of commodities in the agrarian market, and helps farmers choose the right crop/s for a profitable agricultural venture.

2. Related Work

Neural networks have played a key role in the success of recommendation or prediction applications. Developing a prediction model for crop-planting recommendations, considering the price of the crop at harvest time, will increase farmers' returns. The demand for the increase in the food production increases the demand to bridge the productivity gaps and profit for the farmers. Effective utilization of available resources, and predicting crop suitability, has attracted the attention of researchers worldwide. Wright et al. investigated the selection of corn, instead of cotton and rice, in the Southern Mississippi River Basin, considering the economic feasibility of the product. The results demonstrated that corn was far more profitable, 7 times out of 24, than rice and cotton [Wright,(2018)]. Cover crops contribute to a farm's annual profits. Generally perceived positively by farmers, net annual returns from cover crops were, however, largely negative for most farmers [Plastina,(2018)]. A crop recommendation system centered on profitability will increase farmers' net annual returns. The optimal use of resources might be inappropriate in circumstances where the rigidness of a high price model is a priority [Tan,(2017)].

Profitable farming is an important aspect of precision agriculture that has drawn the attention of researchers and research organizations from all over.The Qatar National Research Fund (QNRF) sponsored a decision support tool project that recommends suitable crops and techniques for cost-effective in-country farming [Huda,(2018)]. A combination of the Global Circulation Model (GCM) with a crop simulation model produces optimal crop designs, resulting in increased profitability and reduced risks for farmers [Rodriguez,(2018)]. Several mathematical tools have been developed that predict crop suitability for a specific place, based on the crops grown there earlier and available natural resources [Ramachandran (2018)]. Most, however, failed to process the uncertainties that surfaced in the data gathered. Anithaand Acharjya developed a system that processed uncertainties and predicted crop suitability using hybrid fuzzy approximation and neural networks [Anitha,(2018)].To understand analyses, transparency in data processing is essential.

Developing tools with additional details that promise to deliver high accuracy may appear attractive, but the probability of their being expensive, time-consuming and lacking in transparency is entirely likely [Renton,(2011)].Developing a simple and transparent system is necessary from a farmer's perspective so it can be used effectively and with no complications. [Kuehne,(2017)] designed the ADOPT model to predict diffusion in agricultural practices and an increased understanding of the adoption process, which includes variables such as economies of scale and associated risks.A Convolutional Neural Networks (CNN) model based on image classification was built for crop yield prediction using NDBI and RGB data acquired from UAVs. It was concluded that CNN performed better with RGB data than the NDVI data [Nevavuori, (2019)]. [Nguyen, (2019)] developed a smart system for short-term price prediction using time series models. The proposed Autoregressive Integrated Moving Average (ARIMA) model was tested for two seasonal products, considering different brands of each product. ARIMA was not appropriate for predicting long-term trends. A contingent valuation method was used to elicit farmer's willingness-to-accept (WTA) price to sell their crop straws in northeast China [Zuo, (2020)]. [Alidoost, (2019)] have used copula based analysis to estimate the crop related variables: yield, production and price of a crop for the given extremes of air temperature and precipitation. They used potatoes in Netherlands as a case study.

India, with its multitudes of small-scale farmers, needs a prediction system for suitable crop selection using the recurrent neural network, which delivers increased profitability. The contributions of the paper are:

1. To propose a model to predict crop price for agriculture plantation.
2. To develop a Recurrent Neural Network (RNN) model with Long Short-Term Memory (LSTM).
3. To validate the accuracy of the proposed model with Mean Square Error (MSE) as loss function.
4. To analyze the proposed model for the dataset provided by the Government of India.

The rest of the paper is organized as follows: section 3 discusses Recurrent Neural Network model, section 4 explains the architecture of proposed model with RNN, the dataset exploration is discussed in section 5 and paper concludes in section 6.

3. The proposed prediction model

We propose a recommendation and prediction system for farmers that predicts the market price, at harvest time, of the crop/s planted. It also recommends crops that can be profitably cultivated. As a general rule, farmers' predictions of the returns expected from a particular crop are based on current or previous market prices, and decisions made can occasionally backfire, resulting in significant losses. A loss for small-scale farmers impacts their life significantly, sometimes destroying their livelihood and getting them to quit farming for good. India, a country with a high ratio of small-scale farmers, is greatly affected by inaccurate profitability predictions, and the sheer numbers of farmers committing suicide every year testify to the same What a farmer decides to plant depends on his ability to predict, in the near future, the price of the planned crop at harvest time.Assessing the future market price of a crop to be planted differs, based on individual farmers' perceptions.Since the price prediction of agrarian products plays a strategic role in planting crops, this study proposes a Recurrent Neural Network (RNN)-based model to predict agrarian product prices at harvest time.

3.1 The Recurrent Neural Network (RNN)

Human beings have persistent thoughts, and can recall the past without too much difficulty. Traditional neural networks, however, have no persistence and remember no past data sequences.Recurrent Neural Networks (RNNs) are far removed from traditional neural networks. RNNs have loops, and can be thought of as multiple copies of the same neural network. Fig 1 depicts the structure of a typical RNN. Fig 2 shows an unrolled RNN loop.

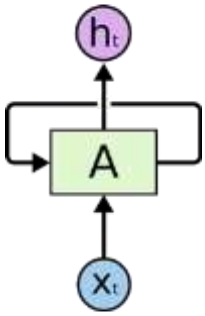


Fig 1: Structure of a typical RNN

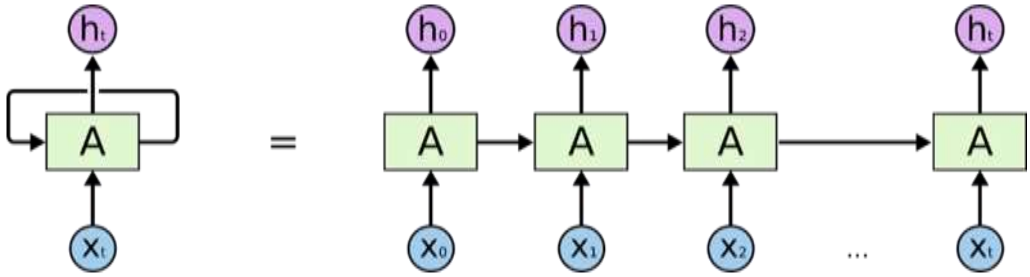
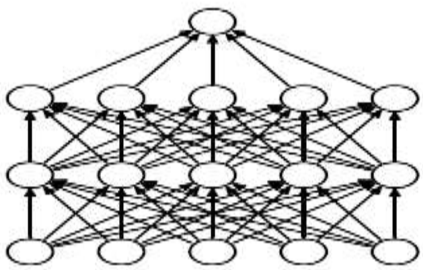


Fig 2: An unrolled RNN loop

3.2 Dropouts

A dropout is a way to prevent neural networks from over fitting. The idea of the dropout is to randomly drop units, along with their connections, in a neural network during training [Srivastava,(2014)].Fig 4 below shows how the dropout significantly enhances neural network performance in supervised learning.



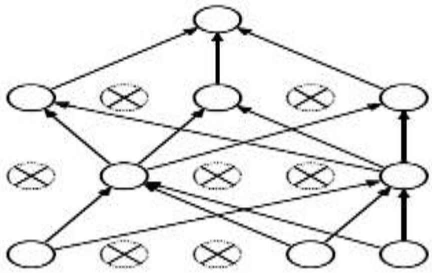
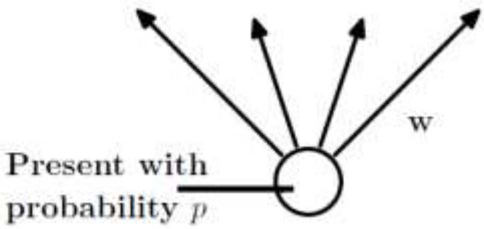


Fig 3: A standardized neural set **Fig 4: The neural set after a drop out application**



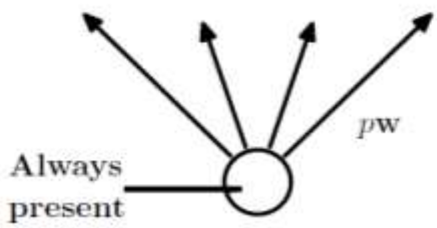


Fig 5: Probability at training time **Fig 6: Probability at testing time**

Fig 3 above shows a fully connected neural network, while Fig 4 displays the result of applying a dropout to the network.In a dropout, a unit as in Fig 5 may be present during training with the probability, *p*, and connected to the next layer with *w* weights. The unit is always present during testing as in Fig 6.

ISSN: 2005-4238 IJAST

4. Building our neural network model

Our neural network Architecture consists of a sequential stack of layers with four Long Short-Term Memory (LSTM) cells, each with a dropout of 20% with 50 units, the architecture is depicted in Figure 7. These LSTM cells are connected to a dense layer with an output neuron. The cost or loss function is a function that compares the model's output with the intended values. Since ours is a regression model, we use the Mean Squared Error (MSE) as our loss function, and the optimizer used is the ADAM [Kingma,(2014)]. Our model is trained with a batch size of 4 in 100 epochs. In a neural network, one epoch is equal to an instance of a forward/backward pass of all the training examples. A batch size is equal to the numerical value of the training examples in an instance of a forward/backward pass.

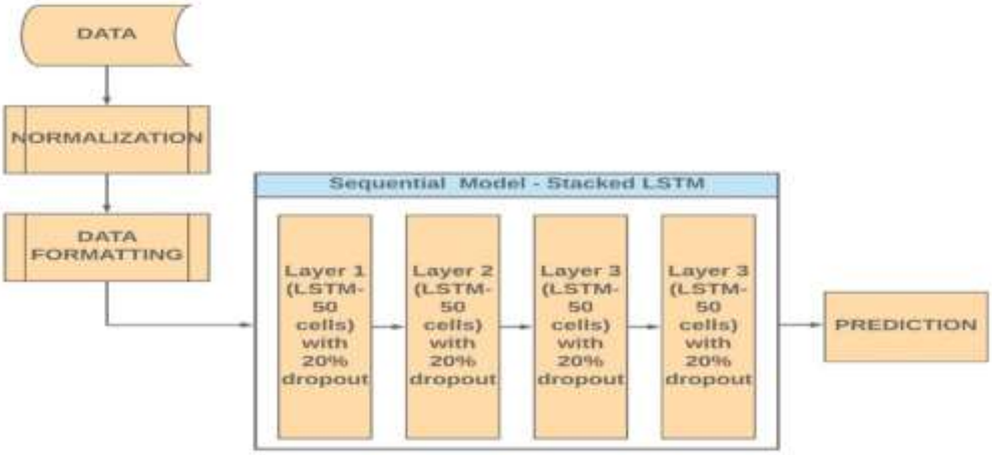
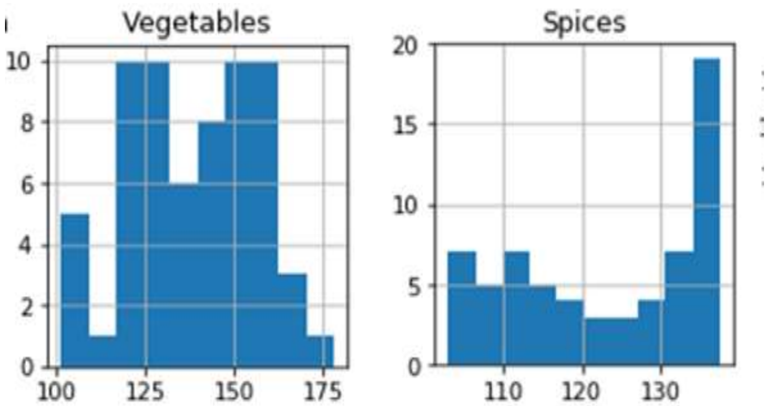


Fig. 7: The architecture of the proposed prediction system

4.2Dataset exploration

Our dataset is downloaded from the website (www.data.gov.in) of the Government of India, the site provides various data.The dataset consists of about 30 columns. Column 1 represents the sector (urban, rural, or both), Column 2 the year (ranging from 2013 to 2018), and Column 3 the month of the year. Column 4 provides the prices of products ranging from cereal and milk to vegetables. Before the raw data was processed, histograms are generated to help understand the price history of the commodities studied. The histograms shown in Fig 8 gives information on the minimum and maximum prices that can be expected during the harvest.Given that histograms are simple representations of data, no complex prediction algorithm is required.



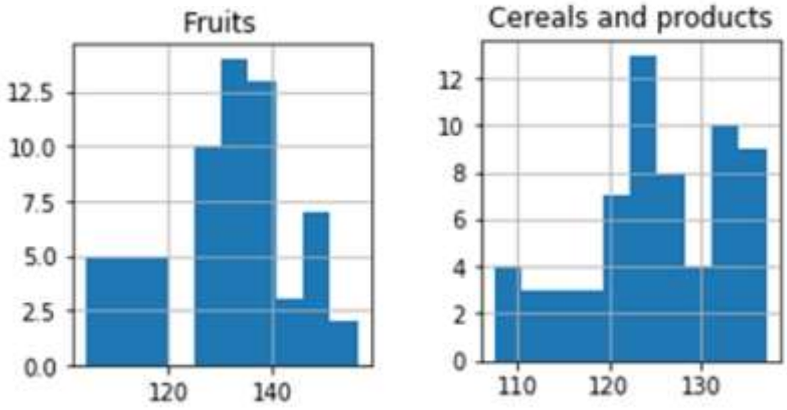


Fig.8: Histograms illustrating the minimum and maximum prices of products

4.3Visualizing the increase in prices over months

After generating histograms for individual commodities, a consolidated graph is generated as shown in Fig 9 that depicts how prices increase over time, with specific reference to the number of months. This helps farmers have an overview of the market price of commodities, and the graph clearly shows huge fluctuations therein. Only a few commodities show a steady increase.

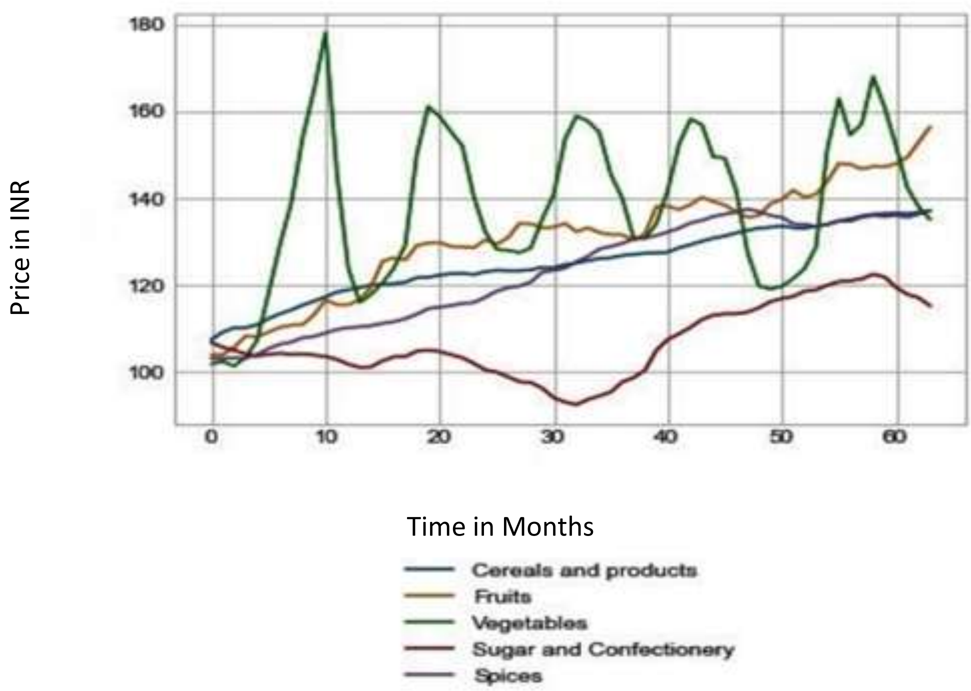


Fig.9: Graph depicting the prices of commodities over time.

4.4Preparing our dataset:

Since product prices are not scaled, feature scaling is needed to normalize the data. All prices are normalized to a scale in the range of (0, 1).Recurrent Neural Networks are supervised learning models that need both independent and dependent variables for training. Given the product prices at hand, we take every three months' price as a row value for X (independent variable), and the price in the 4th month for Y (dependent variable). Therefore, our neural network will predict prices for the next six months, if data for the last three years is provided. The proposed model is impleted in python with TensorFlow package.

5. The prediction model:

The dataset studied has multiple products, and we considered cereal for sample prediction. We found that the model was able to make highly accurate predictions. The dataset consisted of 6 years' data, ranging from 2013 to 2018. We used the data in the 2013 to 2017 range, and predicted the price of cereal for 2018. The results showed that the prediction was similar to the original price of cereal in 2018.Accuracy from the test results was 98% and it is shown in Fig 10.

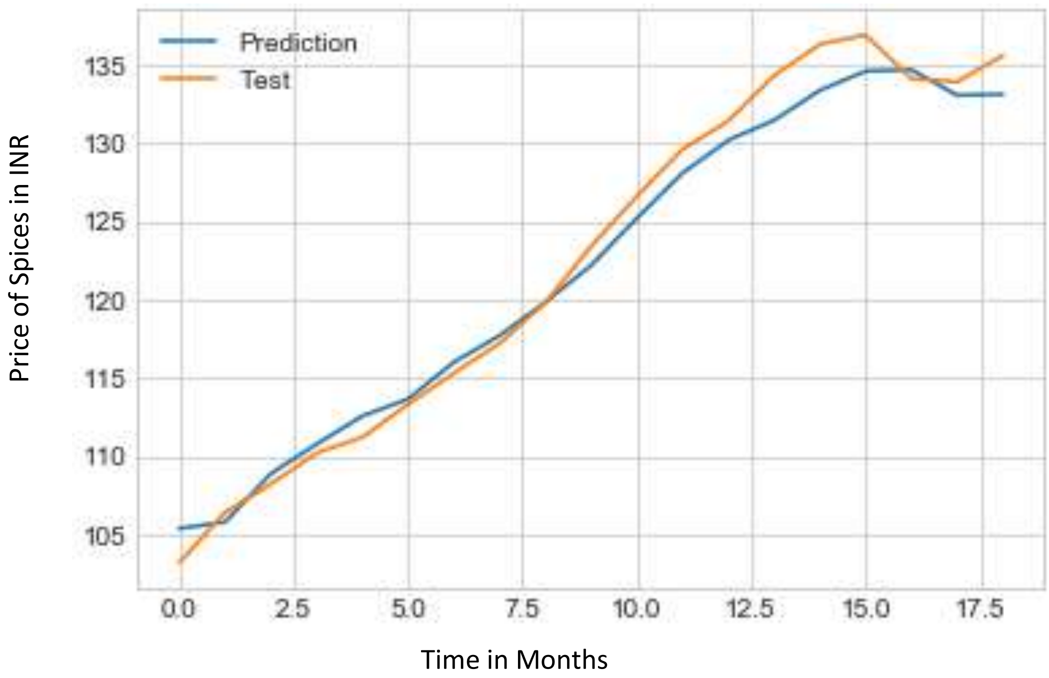


Fig.10: Accuracy of the prediction model

The performance of the prediction model for the year 2018 is shown in table-1. The price of products for the month of January, February and March were the input for the model and the price of the products for the month of April is predicted. The predicted price and the original price for the month of April were compared in the table.

Table 1: Predicted vs Original price of the products

Product Name	January	February	March	Original April Price	Predicted April Price
Cereals and products.	₹136.6	₹136.4	₹136.8	₹137.1	₹137.94054
Fruits.	₹147.9	₹149.4	₹152.9	₹156.4	₹146.45842
Vegetables.	₹152.1	₹142.4	₹138.0	₹135.1	₹145.20792
Sugar and Confectionery.	₹119.5	₹117.5	₹117.1	₹115.2	₹122.09604
Spices	₹136.0	₹135.6	₹136.3	₹137.2	₹136.40706

The performance of our model depends upon the loss function used, which is the Mean Squared Error (MSE), and the results show that the MSE is minimal. Our model has significantly reduced the cost function. Fig 11 below shows the decreased cost function over epochs, with the X-axis representing epochs and the Y-axis the corresponding loss value. The neural network performs well.

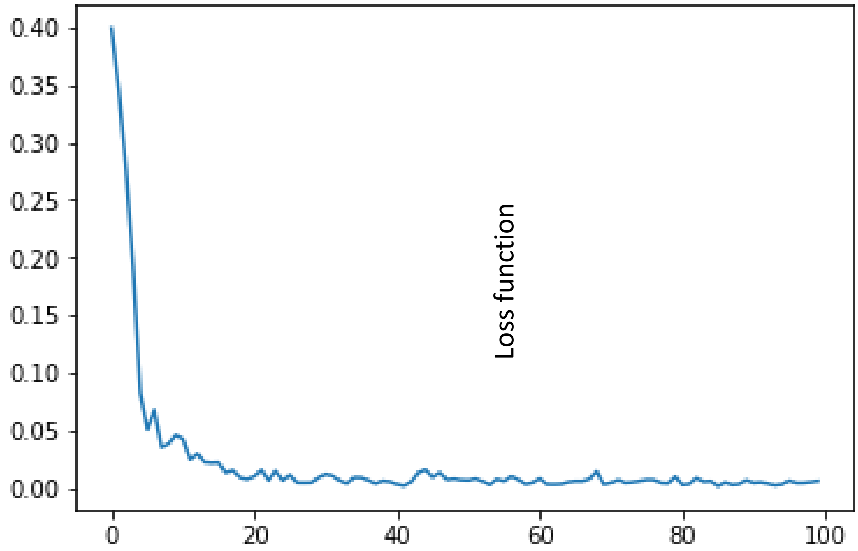


Fig.11: The Mean Squared Error (MSE)

No of Epochs

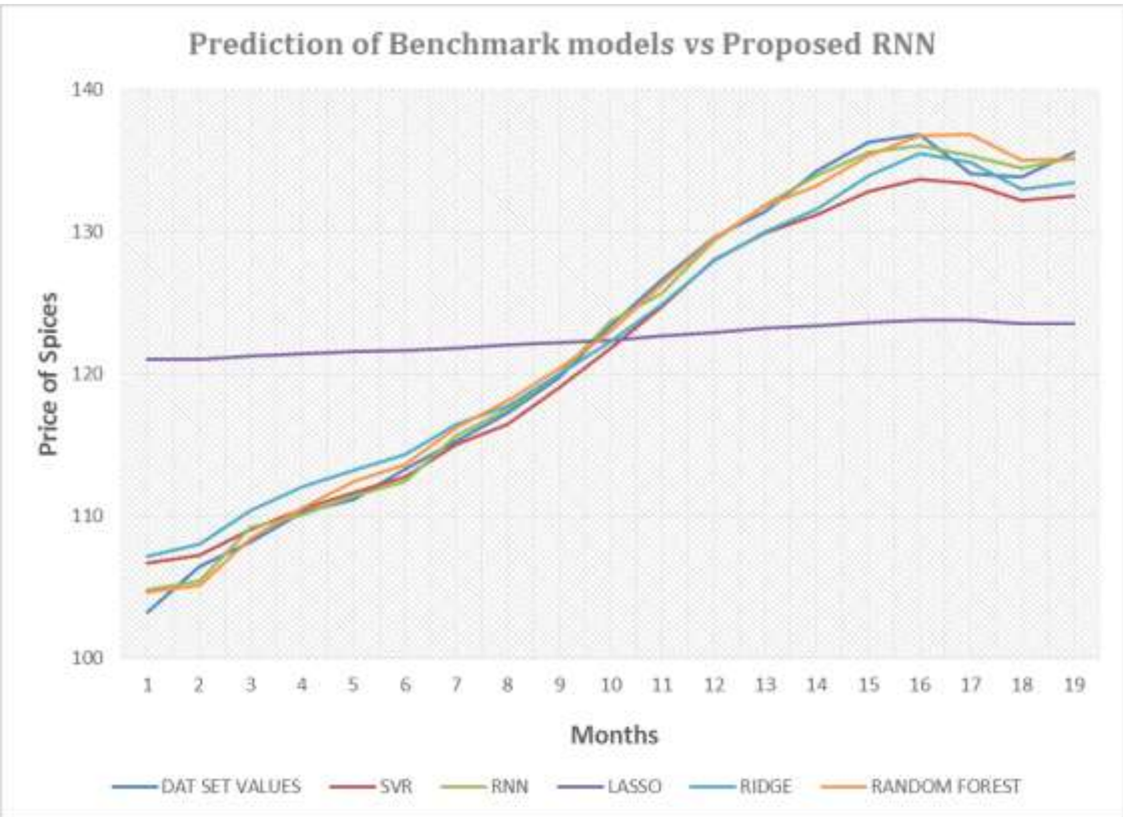


Fig.12:Comparison of Benchmark models with Proposed RNN

The results of the proposed model has to be compared with benchmark models, it is essential to state that for time series data, prediction is done by splitting the historical data and validating the data by time period for the training data. We have considered the SVR, Lasso, Random Forest and Ridge as the Benchmark Models since it's been used in time series data prediction by several research works [Pavlyshenko, (2019)] [Rabiei, (2019)] [Li,(2019)][Liu,(2019)]. Fig 12 depicts proposed work using RNN is performing well in comparison with the SVR, Lasso, Random forest and Ridge models. The Random Forest model performs well when compared to the other benchmark models SVR, Ridge and Lasso models.

6. Conclusion:

Predicting the prices of crops that are to be planted, and recommending alternate crops for planting, is a complex task if agriculture is to be made viable**.** Farmers make planting decisions based on the availability of resources and current or previous market price**.** Poor decisions impact farmers hugely**.** The RNN**-**based prediction model resolves this problem by predicting, at harvest time, the price of a crop planted in season much earlier**.** The model predicts with high accuracy, validated using the MSE**.** With this system, farmers will be able to predict profits from the crop/s planted, and determine alternate profitable planting options**.** The dataset used does not carry individual crop or product prices **-** they are, instead, categorized into fruits, vegetables, cereals and spices**.** The prediction model was tested for categorized products, rather than individual ones**.** The resources available, climatic conditions, and the farm**'**s planting history are to be considered for a future extension of this work**.**

References

1. Anitha, A., &Acharjya, D. P. (2018). Crop suitability prediction in Vellore District using rough set on fuzzy approximation space and neural network. Neural Computing and Applications, 30(12), 3633-3650.
2. Bhargava, A. K., Vagen, T., &Gassner, A. (2018). Breaking Ground: Unearthing the Potential of High-resolution, Remote-sensing Soil Data in Understanding Agricultural Profits and Technology Use in Sub-Saharan Africa. World Development, 105, 352-366.
3. Dewi, C., & Chen, R. C. (2019, April). Decision Making Based on IoT Data Collection for Precision Agriculture. In Asian Conference on Intelligent Information and Database Systems (pp. 31-42). Springer, Cham.
4. Fowler, K. R., Jenkins, E. W., Ostrove, C., Chrispell, J. C., Farthing, M. W., &Parno, M. (2015). A decision making framework with MODFLOW-FMP2 via optimization: Determining trade-offs in crop selection. *Environmental Modelling & Software*, *69*, 280-291.
5. Fulton, J. P., & Port, K. (2018). Precision agriculture data management. Precision agriculture basics, (precisionagbasics), 169-188.
6. Hassan, F. M., Zakir, S. B., Binta, A. S., & Hossain, S. M. (2018). Agricultural yield and profit prediction using data analysis techniques (Doctoral dissertation).
7. He, F., Borisova, T., Athearn, K., Hochmuth, R., & Barrett, C. (2019). Economic Analysis to Support the Development of Agricultural Best Management Practices and Protect Florida's springs.
8. https://data.gov.in/sectors/Agriculture-9212
9. https://medium.com/hub-konam-foundation/predicting-crop-yield-and-profit-with-machine-learning-1d9c3216faf7
10. https://www.indiaspend.com/farmer-suicides-down-21-in-2016-farming-sector-suicides-at-decade-low-56122/
11. Hu, M., Liu, Y., & Wang, W. (2016). Altruistic rationality: The value of strategic farmers, social entrepreneurs and for-profit firms in crop planting decisions. Available at SSRN.
12. Huda, A. K. S., Issaka, A., Kaitibie, S., Haq, M., Goktepe, I., Moustafa, A. & Huda, N. (2018, March). Improving Vegetable Crop Production in Qatar: Strategies to determine optimum planting time minimise production risk and maximise water and nutrient use efficiency. In Qatar Foundation Annual Research Conference Proceedings (Vol. 2018, No. 1, p. EEPP126). Qatar: HBKU Press.
13. Image credits in proposed: http://colah.github.io/posts/2015-08-Understanding-LSTMs/.

14. Ji, X., &Cobourn, K. M. (2018). The Economic Benefits of Irrigation Districts under Prior Appropriation Doctrine: An Econometric Analysis of Agricultural Land Allocation Decisions. Canadian Journal of Agricultural Economics/Revue canadienned'agroeconomie, 66(3), 441-467.
15. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
16. Kuehne, G., Llewellyn, R., Pannell, D. J., Wilkinson, R., Dolling, P., Ouzman, J., & Ewing, M. (2017). Predicting farmer uptake of new agricultural practices: A tool for research, extension and policy. Agricultural Systems, 156, 115-125.
17. Liverpool-Tasie, L. S. O., Omonona, B. T., Sanou, A., &Ogunleye, W. O. (2017). Is increasing inorganic fertilizer use for maize production in SSA a profitable proposition? Evidence from Nigeria. Food policy, 67, 41-51.
18. Merriott, D. (2016). Factors associated with the farmer suicide crisis in India. Journal of epidemiology and global health, 6(4), 217-227.
19. Michler, J., Tjernstrom, E., Verkaart, S., &Mausch, K. (2018). Money matters: The role of yields and profits in agricultural technology adoption.
20. Plastina, A., Liu, F., Miguez, F., & Carlson, S. (2018). Cover crops use in Midwestern US agriculture: perceived benefits and net returns. Renewable Agriculture and Food Systems, 1-11.
21. Renton, M. (2011). How much detail and accuracy is required in plant growth sub-models to address questions about optimal management strategies in agricultural systems? AoB Plants, 2011.
22. Rodriguez, D., De Voil, P., Hudson, D., Brown, J. N., Hayman, P., Marrou, H., &Meinke, H. (2018). Predicting optimum crop designs using crop models and seasonal climate forecasts. Scientific reports, 8(1), 2231.
23. Rodriguez, D., De Voil, P., Hudson, D., Brown, J. N., Hayman, P., Marrou, H., &Meinke, H. (2018). Predicting optimum crop designs using crop models and seasonal climate forecasts. Scientific reports, 8(1), 2231.
24. Srivastava, Nitish, et al. "Dropout: a simple way to prevent neural networks from overfitting." The Journal of Machine Learning Research 15.1 (2014): 1929-1958.
25. Tan, Q., Cai, Y., & Chen, B. (2017). An enhanced radial interval programming approach for supporting agricultural production decisions under dual uncertainties and differential aspirations. Journal of cleaner production, 168, 189-204.
26. Wang, A. X., Tran, C., Desai, N., Lobell, D., &Ermon, S. (2018, June). Deep transfer learning for crop yield prediction with remote sensing data. In Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies (p. 50). ACM.
27. Wright, R., Griffin, T. W., Guha, G., &Bouldin, J. (2018). Economic Investigation of Crop Selection in the Southern Mississippi River Basin. The Journal of the ASFMRA, 173-191.
28. Yang, L., Liu, M., Lun, F., Yuan, Z., Zhang, Y., & Min, Q. (2017). An analysis on crops choice and its driving factors in agricultural heritage systems—A case of Honghe Hani Rice Terraces System. Sustainability, 9(7), 1162.
29. Nevavuori, P., Narra, N., & Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. *Computers and Electronics in Agriculture*, *163*, 104859.
30. Nguyen, H. V., Naeem, M. A., Wichitaksorn, N., & Pears, R. (2019). A smart system for short-term price prediction using time series models. Computers & Electrical Engineering, 76, 339-352.
31. Zuo, A., Hou, L., & Huang, Z. (2020). How does farmers' current usage of crop straws influence the willingness-to-accept price to sell?. Energy Economics, 86, 104639.
32. Alidoost, F., Su, Z., & Stein, A. (2019). Evaluating the effects of climate extremes on crop yield, production and price using multivariate distributions: A new copula application. Weather and climate extremes, 26, 100227.
33. Pavlyshenko, B. M. (2019). Machine-learning models for sales time series forecasting. Data, 4(1), 15.

34. Rabiei, M. R., Arashi, M., &Farrokhi, M. (2019). Fuzzy ridge regression with fuzzy input and output. Soft Computing, 23(23), 12189-12198.
35. Li, T., Zhou, Y., Li, X., Wu, J., & He, T. (2019). Forecasting daily crude oil prices using improved CEEMDAN and ridge regression-based predictors. Energies, 12(19), 3603.
36. Liu, J., Xiao, A., Lei, G., Dong, G., & Wu, M. (2019). Intelligent predicting of salt pond's ion concentration based on support vector regression and neural network. Neural Computing and Applications, 1-15.
37. Ramachandran, V., Ramalakshmi, R., & Srinivasan, S. (2018, November). An automated irrigation system for smart agriculture using the Internet of Things. In 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV) (pp. 210-215). IEE