PREDICTING AGRICULTURAL PRODUCE PRICES WITH CONVOLUTION NEURAL NETWORKS: IMPROVING THE LIVES OF INDEBTED FARMERS WITH DEEP LEARNING

J. Mahalakshmi[1], Y. Manisa[2], S. Hemanth Kumar[3], N.Keerthana[4], P.Jhansi Lakshmi[5]

[1] Associate Professor, Krishna Chaitanya Institute of Technology & Sciences , Markapur, A.P, India
[2,3,4,5] Scholar, Krishna Chaitanya Institute of Technology &Sciences , Markapur, India

ABSTRACT:

Farmer suicides have emerged as a pressing societal issue that governments all around the world are working diligently to address. Most farmers commit themselves because they can't sell theirThe extensive uncertainty/fluctuation in the market makes it difficult to produce at targeted profit levels, which produce costs as a result of changing market circumstances. This aims to stop farmer suicides, In order to address the issue of product price unpredictability, this study makes a first step. Introducing PECAD, a deep learning method for precise produce price forecasting according to historical price and volume trends. Despite the fact that earlier research has introduced machine learning algorithms for produce price prediction, these algorithms have two drawbacks: I they do not explicitly take into account the spatio-temporal dependence of future prices on past data; as a result, (ii) they rely on classical ML prediction models, which frequently exhibit poor performance when applied to spatio-temporal datasets. Through three key contributions, PECAD tackles these limitations: We collect actual daily prices and (produced) volume data for various crops over a period of 11 years from a website run by the Indian government; (ii) pre-process this raw information using cutting-edge imputation techniques to adjust for missing data entries; and(iii) PECAD suggests a brand-new broad and deep neural network architecture made up of two distinct convolutional neural network models that were trained on price and volume data, respectively. Our simulation findings demonstrate that PECAD surpasses current state-of-the-art baseline approaches by obtaining noticeably lower root mean squared error (RMSE) – PECAD produces a coefficient of variance that is around 25% lower than state-of-the-art baselines. To reduce farmer suicides in the Indian state of Jharkhand, we collaborate with a non-profit organization, and PECAD is now being evaluated by for possible implementation.

Keywords : Neural Networks, PECAD, Deep Learnng, RMSE, Machine learning

[1] INTRODUCTION

Small-scale farmer suicide rates have significantly increased over the past 20 years, particularly in emerging nations like India, Pakistan, etc., due to the problem of agrarian distress (and other related socio-economic issues like debt, loss of agricultural revenue, etc.). There have been around 300,000 farmer suicides in India. 2020 Copyright, Association for the Advancement of Artificial Intelligence (www.aaai.org). from 1995. All rights reserved. In the Indian state of Maharashtra alone, suicide rates among farmers reached 60,000 as of 2014, with an average of 10 each day (NCRB 2019).

Crop failures, lower agricultural productivity, an inability to turn a profit, ineffective cold chain management leading to agricultural produce waste, a lack of irrigation infrastructure, and unmanageable debt are just a few of the many variables that contribute to farmer suicides. The uncertainty surrounding agricultural pricing and markets, on the other hand, is a major factor in farmer suicides. For example, changes in global market conditions can cause large swings in the price of agricultural output at a local level (Barik 2018). Small-scale farmers with debt who frequently lack modern technology tools and awareness of global market circumstances are unable to make informed judgments about whether (and where) to sell their goods as a result of this pricing uncertainty.Due to their inability to get the expected earnings from their goods and pay back their agricultural debts, many of these formers commit suicide (see Figures 1a and 1b) (Panagariya 2008).

(a)Farmers Protesting by Throwing their Unsold Produce (b) Huge Demand for Loan Waiver at Farmer Rally

Figure 1: Agrarian Distress in India

Therefore, fast action is required to address these farmers' problems. Recent developments in machine learning (ML) methods have made it feasible to successfully apply learning algorithms to many societal concerns (Tambe and Rice 2018). This article suggests an AI/ML strategy to address the following query as a first step in resolving the issues facing farmers described above: Can data-driven systems forecast future prices of agricultural produce at various marketplaces using previous pricing and volume trends at other markets? Farmers may then utilise these AI/ML techniques to choose smart selling tactics for their product, for example, by using future price forecasts to choose when (in the future) to sell their food in order to make the most money.

To get the solution to this issue, there are a number of problems that must be resolved. First, the training process is hampered by the fact that current datasets on price patterns1 are highly sparse (i.e., they include a large number of missing values). It is crucial to create prediction models that can explicitly capture this spatio-temporal dependence because future produce prices have a long-term temporal dependence on past prices (for example, the cost of tomatoes in August 2019 may depend on their price in August 2018) as well as a spatial dependence on prices at nearby markets (for example, prices at nearby markets may be similar to those at geographically distant markets).

While prior research has provided algorithms to forecast produce prices, they I do not explicitly take into account the spatiotemporal dependence of future prices on historical data; as a result, (ii) they rely on traditional ML prediction models (such as decision trees), which frequently perform poorly when applied to spatiotemporal datasets (we validate this in our experiments). These flaws restrict the applicability of these strategies in the actual world and consequently their accuracy.

In this study, we propose PECAD (Price Estimation for Crops Using the Application of Deep Learning), an unique neural network design to forecast future agricultural product prices, in order to overcome these inadequacies. The following fresh contributions are made by PECAD to overcome the deficiencies in earlier work. First, over the course of 11 years (from 2008 to 2018), it gathers actual prices and (produced) volume of various commodities at around 1,350 agricultural marketplaces in India (an official Indian government administered website). Second, to account for missing data items, PECAD preprocesses this raw dataset using cutting-edge imputation (and other) algorithms.Third, PECAD suggests an unique neural network architecture based on the wide and deep learning paradigm that trains broad linear models and deep neural networks simultaneously utilising this data as input (Cheng et al., 2016). But PECAD employs a new combination of two distinct convolutional neural network (CNN) models for price and volume data, respectively (for the crop under consideration), and uses these CNN models as input to the broad linear model instead of cross-product feature transformations. Our simulation results demonstrate that PECAD significantly outperforms current state-of-the-art baseline methods, achieving a 25% lower coefficient of deviation than baseline methods, emphasising the significance of explicitly modelling the spatiotemporal dependence of future prices on historical data inside PECAD our ML algorithm. PECAD is now being evaluated for future deployment by a non-profit organisation that collaborates with us on reducing farmer suicides in the Indian state of Jharkhand (name omitted to protect confidentiality).

Relevant Work , We talk about previous AI/ML research that helps to ease rural suffering. Deep Gaussian processes were suggested by (You et al. 2017) to forecast crop yields us1. ing information via remote sensing http://agmarknet.gov.in. However, because their method relies on collecting satellite photos of fields, it can be costly to do so in low-resource settings in underdeveloped nations. In order to forecast future prices, we use readily accessible pricing and volume data in our work. Then, in 2017, Chen, Nowocin, and Marathe suggested a hardware and software method to lessen crop waste. The study that is most similar to ours is (Ma et al. 2019), since they also created a crop price prediction model utilising information from the same source1.Sadly, they do not take use of the spatio-temporal characteristics of price and (produced) volume data for various crops, which results in subpar performance accuracy (as we show in our experiments). In our study, we identify spatio-temporal connections in price and volume data using certain types of convolutional neural networks.

[2] LITERATURE SURVEY

For the majority of deep learning experts, recurrent networks and sequence modelling are interchangeable terms. Convolutional architectures can, nevertheless, outperform recurrent networks on tasks like audio synthesis and machine translation, according to recent research. Which architecture should one choose when faced with a fresh dataset or job for sequence modelling? For the purpose of sequence modelling, we systematically assess general convolutional and recurrent architectures. The models are assessed using a wide range of industry-standard tasks that are frequently used as recurrent network benchmarks. According to our findings, a straightforward convolutional architecture outperforms LSTMs across a wide range of tasks and datasets while exhibiting a longer effective memory. We come to the conclusion that convolutional networks should be viewed as a natural starting point for challenges involving sequence modelling rather than the traditional relationship between recurrent networks and sequence modelling. We have provided code to help with related work.

Crop waste has been a major issue for India's agricultural sector. The fact that too many small farmers lack accessible to dependable cold storage, which increases agricultural shelf-life, is a major contributing element to this issue. These farmers sometimes sell their harvests at disadvantageous low rates in order to prevent having excess produce that spoils. Unavoidably, some harvests are not sold before going bad. Farmers must plan ahead for when and where to sell their harvest since even if they have access to cold storage, they may not know how long to keep various crops there. This note details the development of a hardware and software solution that will assist farmers in lowering crop waste and raising earnings.The hardware consists of a control unit and refrigerator that run efficiently on solar power. The software is a system for forecasting produce prices, and we have experimented with various machine learning techniques for it. Note that the produce price data from primarily rural Indian markets have a significant number of missing values, in contrast to standard price forecasting tasks such as for stock market data. We are actively collaborating with farmers at two pilot sites in Karnataka and Odisha as we develop our two-pronged solution.

In the study of organisational demography, demographic heterogeneity is a key theoretical concept. The coefficient of variation is the most popular metric for assessing demographic heterogeneity. The use of the coefficient of variation raises a number of theoretical and interpretive issues, as the author demonstrates in her critical assessment of the justification for using this measure. Using the coefficient of variation may result in incorrect conclusions about the impacts of demographic heterogeneity, according to empirical analyses of turnover.

In order to calculate the post-mortem interval (PMI), which tells us about the natural occurrences and environmental factors that could have had an impact on the remains after death, forensic anthropologists use the degree of decomposition of a body. The pace of decomposition is known to be influenced by a number of variables, including temperature, precipitation, and body exposure. On the impact of body size on the pace of decay, there are, nevertheless, contradicting accounts in the literature. In order to determine how body size affects decomposition rates, this experiment compared the decomposition rates of big pigs (Sus scrofa; 60–90 kg) with tiny pigs (35 kg).In order to determine how quickly tiny pigs decompose, 15 piglets were monitored three times a week for three months in the spring and early summer. Up until the point of full skeletonization, data was collected. The total body score (TBS), which indicates the overall stage of decomposition for each pig, was calculated by adding the point values for each anatomical location. Decomposition stages were scored in accordance with distinct categories for each anatomical region. The information from 15 big pigs was used. In order to evaluate the pattern of decomposition and compare the rates of decomposition between small and big pigs, scatter plots demonstrating the correlations between TBS and PMI as well as TBS and accumulated degree days (ADD) were employed.According to the findings, both samples' early phases of decomposition involve fast breakdown. In the course of advanced phases of decomposition, large pigs had a plateau period during which little breakdown occurred. Small pigs weighing more than 20 kg attained a comparable but significantly shorter plateau at a PMI of 20–25 days, following which decomposition started right away. In contrast, the 20 kg or less little pigs did not have a plateau period and disintegration rates were rapid throughout the course of the trial. Overall, the rate of decomposition was 2.82 times faster for little pigs than for giant pigs, showing that body size does affect the pace of decomposition.

Monitoring of agriculture, especially in developing nations, can boost humanitarian operations and help stop starvation. The ability to forecast crop yields prior to harvest is a major difficulty. We provide a scalable, precise, and affordable technique to forecast agricultural yields using openly accessible remote sensing data. Our strategy enhances current methods in three different ways. First, we suggest a method based on contemporary representation learning concepts rather than the hand-crafted features that have typically been employed in the remote sensing sector. We provide an unique dimensionality reduction method that enables us to automatically acquire valuable features when training a Convolutional Neural Network or Long Short Term Memory even in the absence of sufficient labelled training data.In order to explicitly represent the spatio-temporal structure of the data and further increase the accuracy, we include a Gaussian Process component. We test our method on U.S. counties' soybean output, and the results demonstrate how much better it works than alternatives.

[3] SYSTEM ARCHITECTURE

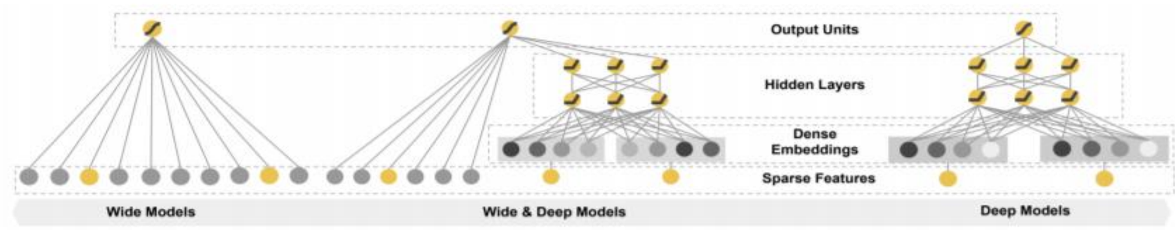

Fig. 2 System Architecture

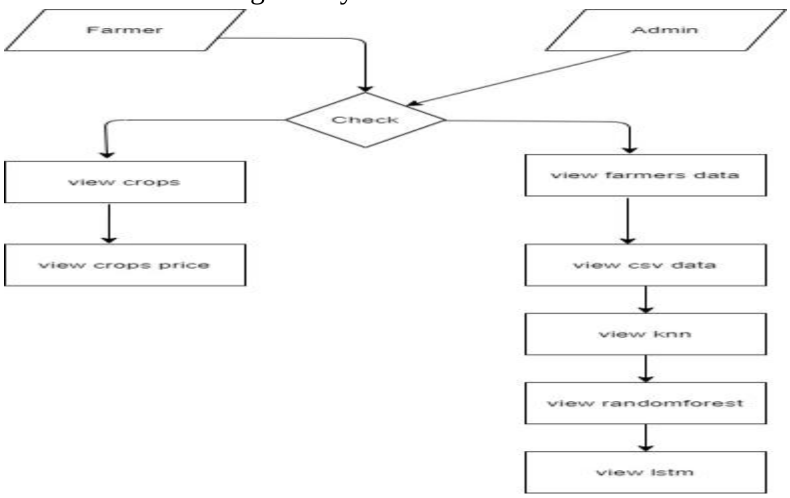

Fig. 3 Flow Chart Diagram

[4] IMPLEMENTATION

4.1 Modules Description

i) Farmer: The farmer may sign up first. He needed a working Farmer email and cellphone during registration for future conversations. Admin can activate the Farmer once they register. The Farmer can log into our system once the Admin has activated them. He may search the crop details after logging in. The farmer will receive all the crop information necessary for searching. A farmer who clicks on the crop price will see the state-specific pricing as well as information on the crop price from the previous year.

ii) Admin: Farmer activation is limited to our applications. The data set may be set by the admin. This report's data has taken state- and year-specific crop pricing information into account. After that, admin will use the dataset to apply algorithms. He will use the knn algorithm first, followed by the random forest then the cnn algorithm. We will obtain accuracy using the algorithm.

iii) Data Preprocess: The sqlite database has been used to hold the data that the admin gave. Our technique requires that we clean the data before processing it. We may fill in the missing numbers with the mean type using a pandas data frame. The histogram will be shown once the data has been cleansed.

4.2 Screenshots

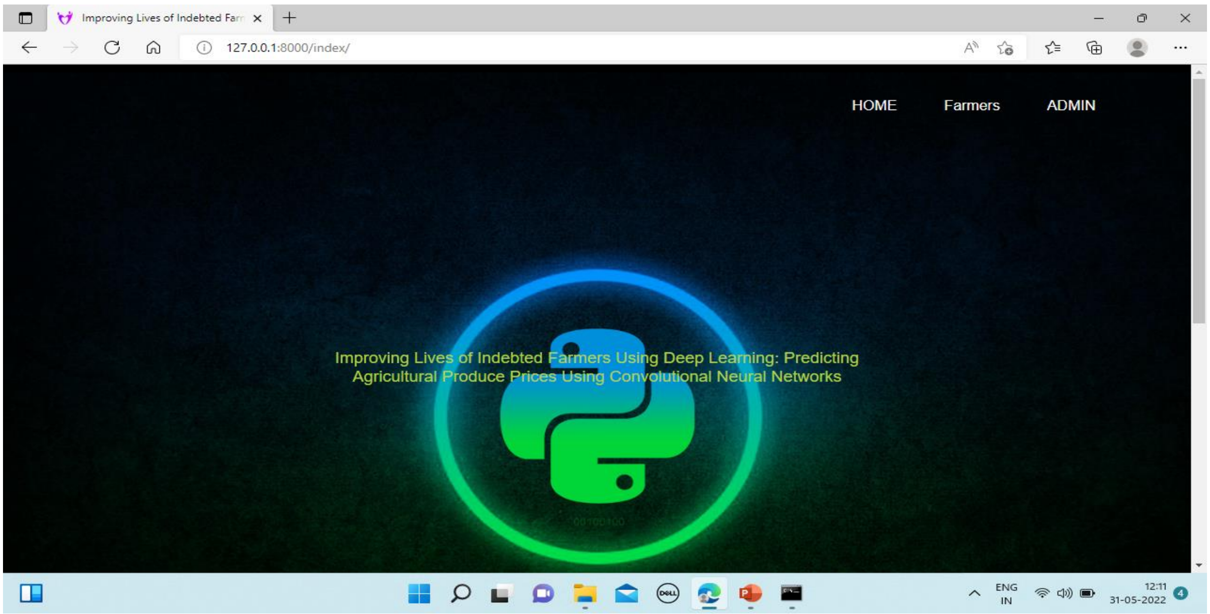

Fig. 4 Home Page

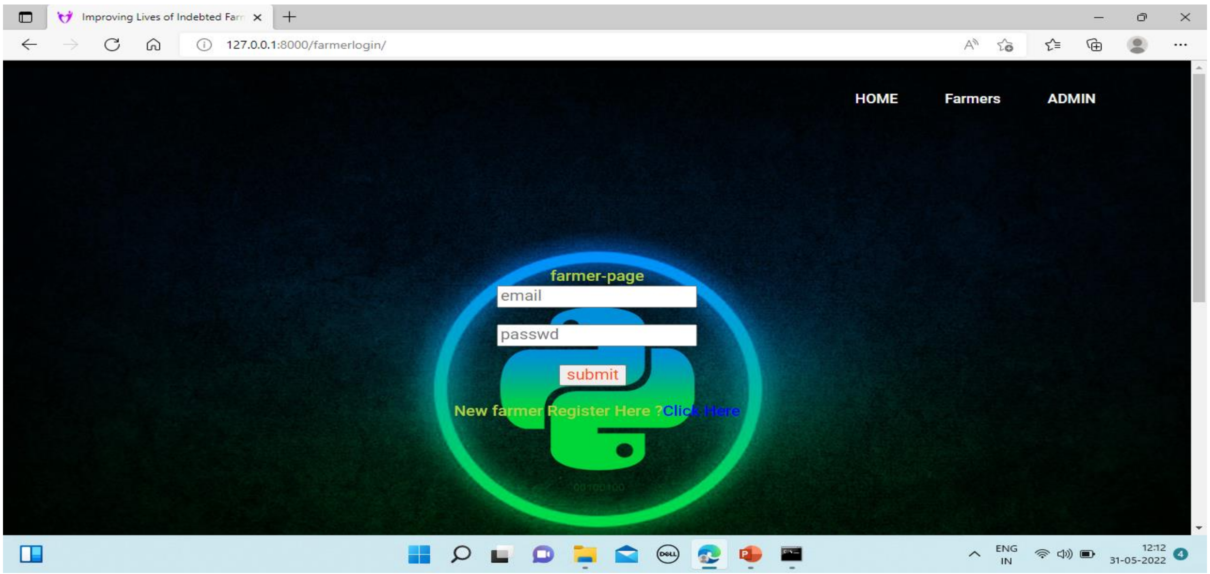

Fig. 5 Login Page

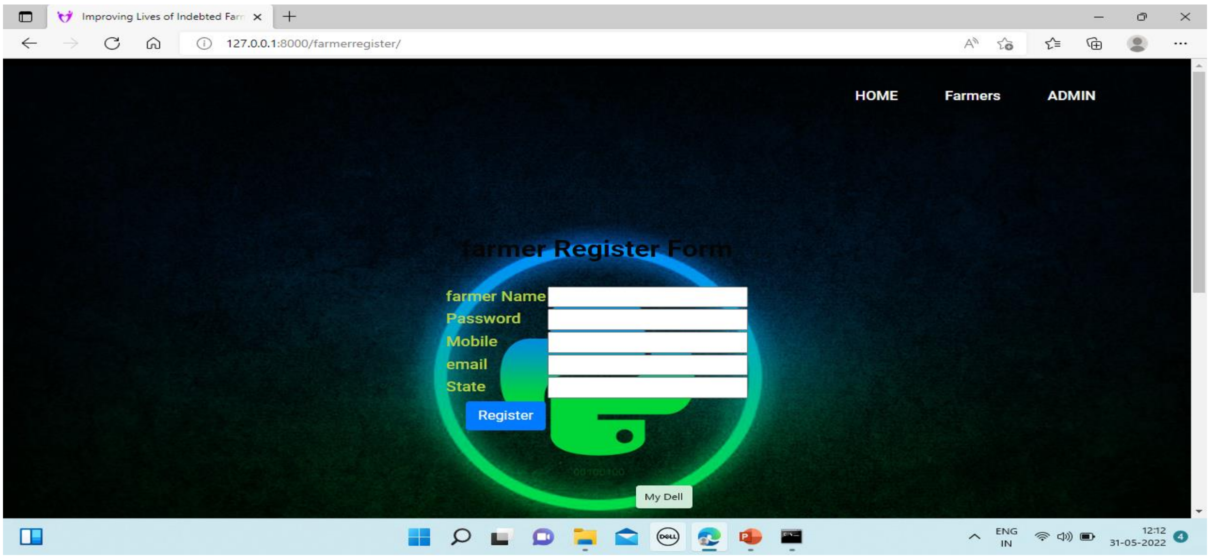

Fig. 6 Registration Page

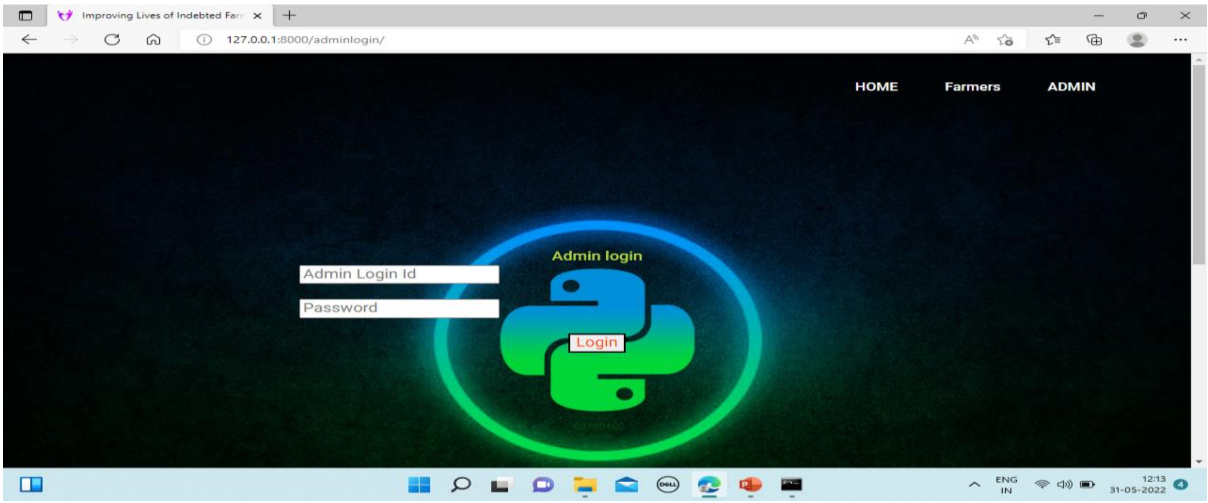

Fig. 7 Registration Page

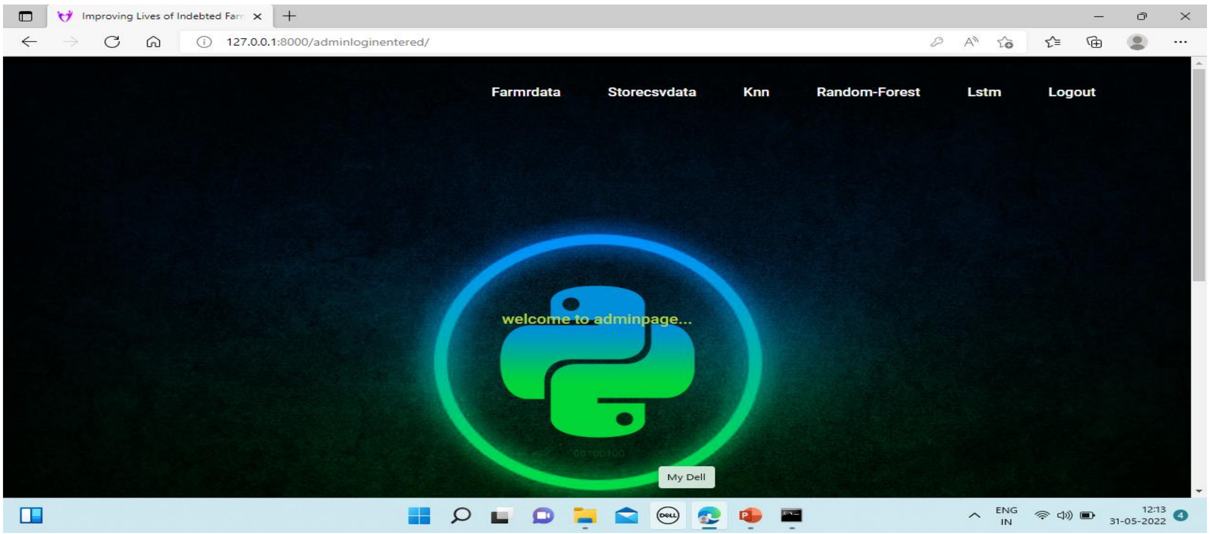

Fig. 8 Admin Page

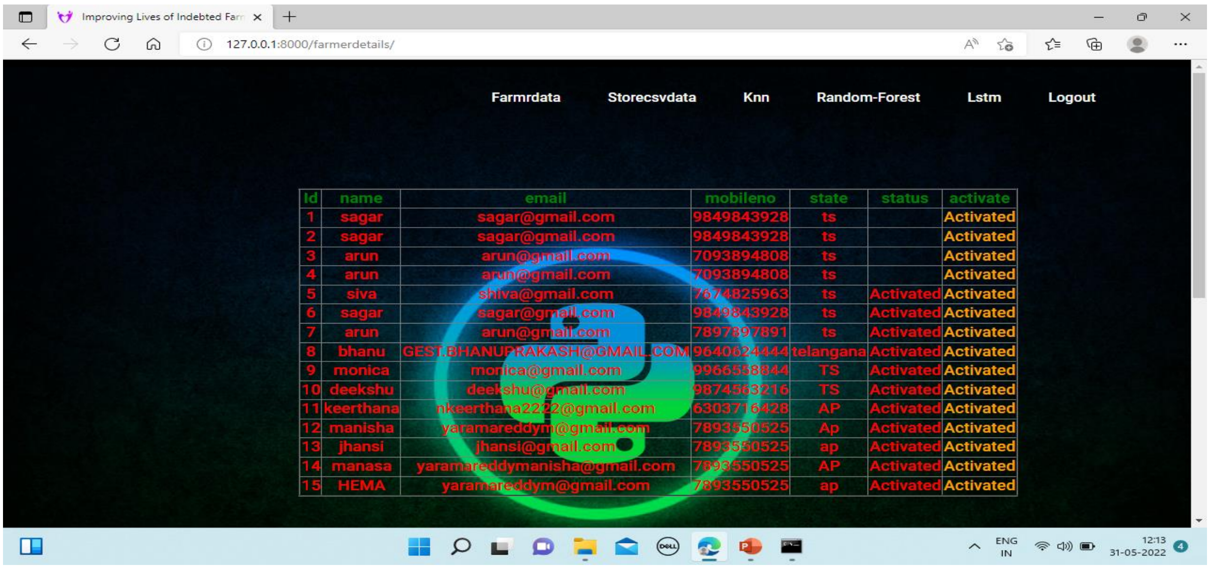

Fig. 9 Activation Page

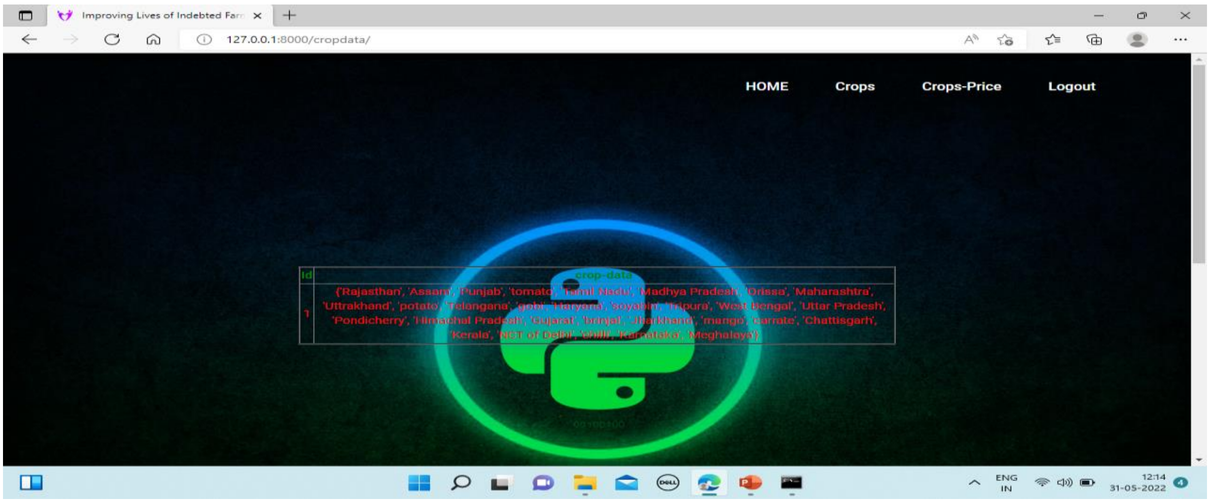

Fig. 10 Display Page

Fig. 11 Input Data Entry Page

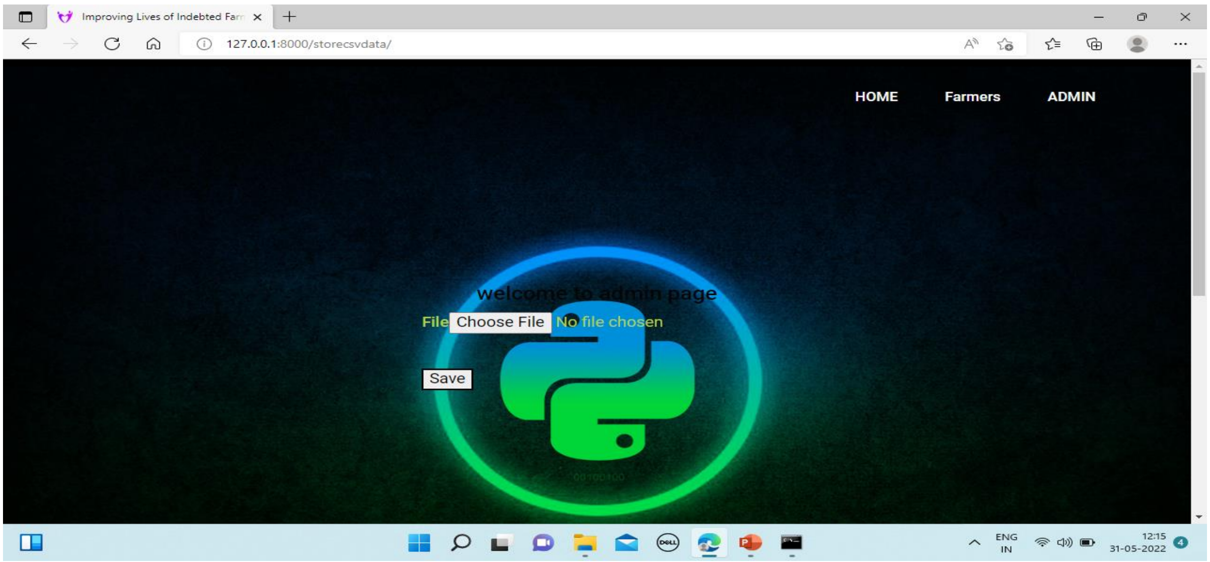

Fig. 12 Data Upload Page

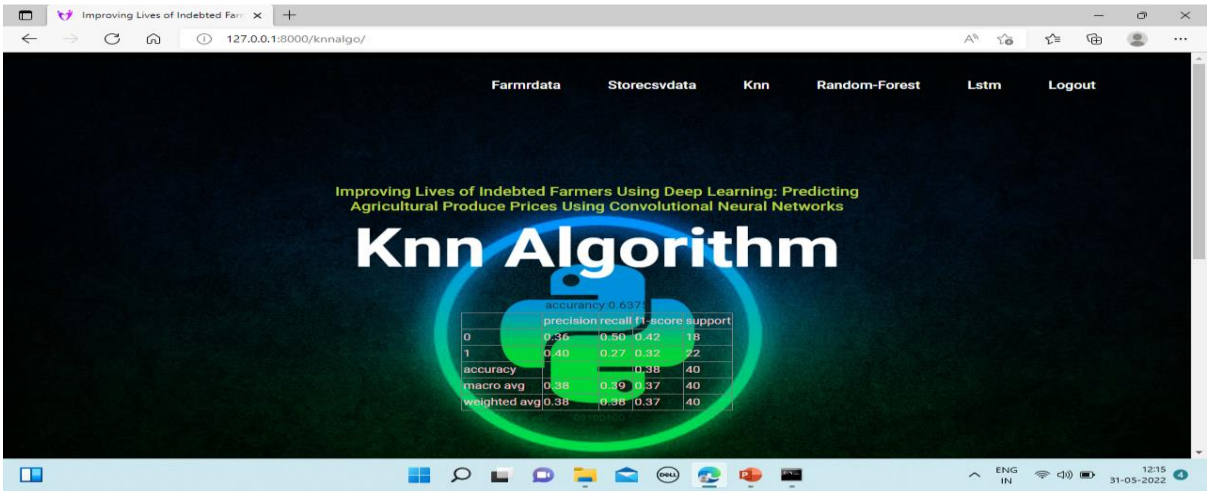

Fig. 13 Applying KNN Alogithm Page

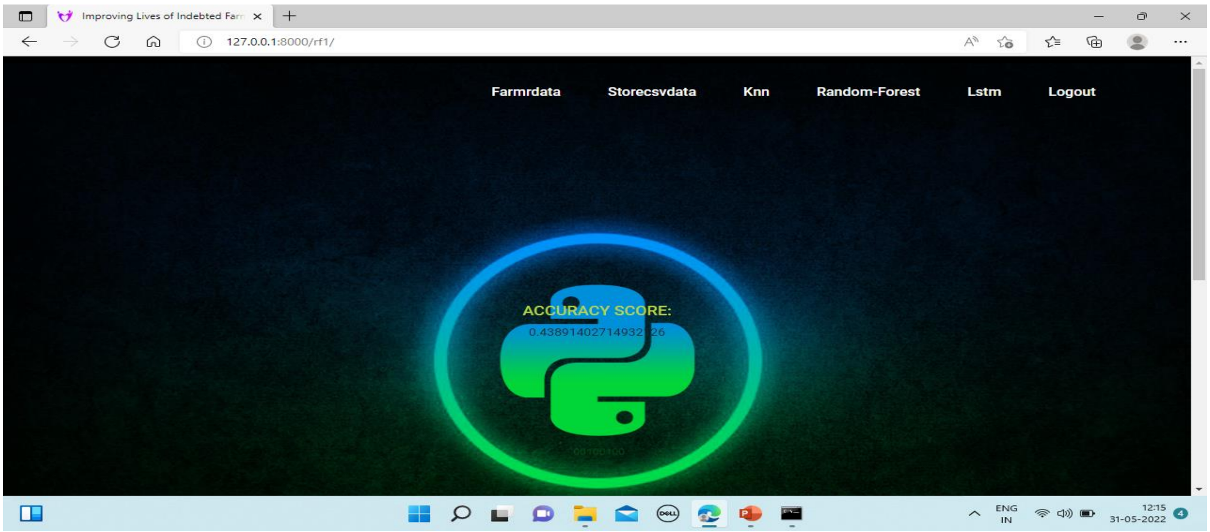

Fig. 14 Accuracy Display Page

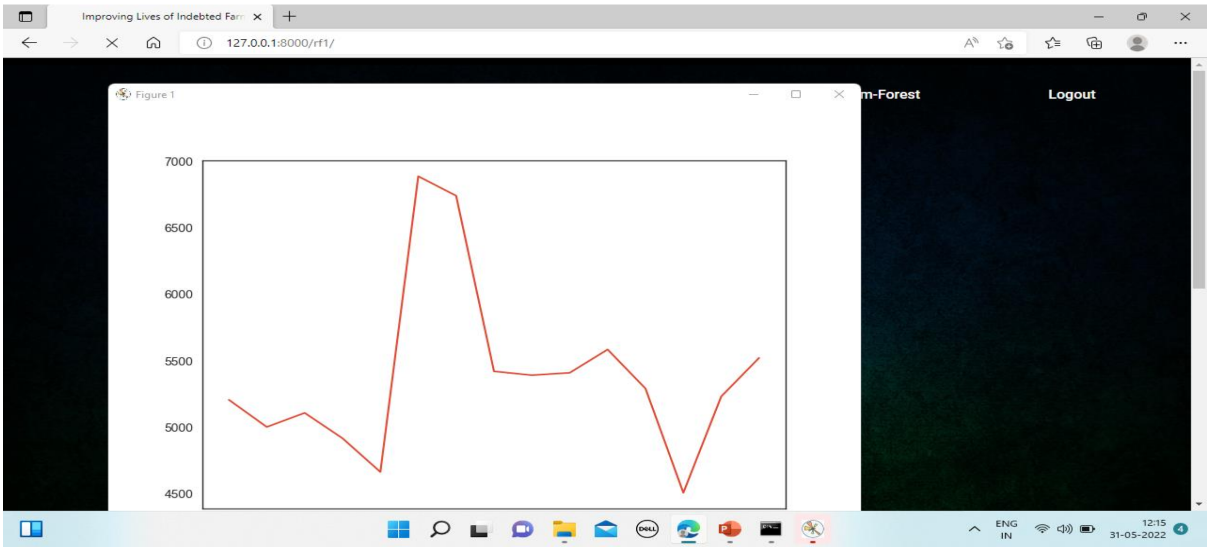

Fig. 15 Results- Accuracy Score Grapgh Page

[5] CONCLUSION

When non-profit organisations working with indebted farmers use PECAD, there are a few practical issues that need to be resolved. First, by including past weather patterns, which may influence crop availability in the future, PECAD's prediction performance may be enhanced (and hence, the future crop price). However, physical models are far better at forecasting the weather in the future, hence deep learning techniques are rarely employed to simulate weather in the actual world. Consequently, PECAD and physical weather forecast models must be combined (as part of future work).Furthermore, very advanced deep learning methods (like PECAD) for forecasting future crop prices may raise warning flags for farmers with little literacy. Public education campaigns at the organisations implementing this programme would aid in overcoming these anxieties and promoting participation. Additionally, non-profit organisations frequently do not place a high priority on using their limited funds to purchase expensive computer technology (to train and run PECAD).

Deploying PECAD as a standalone online service would allow the agencies to utilise it without our help, which is why we suggest it. Last but not least, PECAD is just one component of a much larger problem that has to be addressed in order to avoid farmer suicides. For instance, the existence of long-term crop pricing and volume trends is vital to the adoption of PECAD. While Agmarknet.gov.in provides access to this data for Indian markets, there are no comparable data sources for other developing nations. In this study, PECAD, a deep learning method for accurate produce price prediction based on historical pricing and volume trends, is introduced.Previous machine learning (ML) methods for forecasting produce prices had substantial drawbacks since they do not explicitly take into account the spatio-temporal dependency of future prices on historical data. PECAD proposes a revolutionary broad and deep learning architecture that addresses these concerns by training two independent convolutional neural network models for price and volume data, respectively (for the crop under consideration). Our simulation findings demonstrate that, by obtaining a 25% lower coefficient of variance, PECAD beats current state-of-the-art baseline approaches. PECAD is now being evaluated by an Indian non-profit organisation in the Indian state of Jharkhand that strives to prevent farmer suicides for possible deployment. Our work is done in partnership with this organisation.

In this research, PECAD, a deep learning system for accurate produce price prediction based on historical pricing and volume trends, is presented. Previous machine learning (ML) methods for forecasting product prices had substantial drawbacks since they do not explicitly take into account the spatial-temporal dependency of future prices on historical data. PECAD proposes a revolutionary broad and deep learning architecture that addresses these concerns by training two independent convolutional neural network models for price and volume data, respectively (for the crop under consideration).

REFERENCES

[1] Bai, S.; Kolter, J. Z.; and Koltun, V. 2018. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv preprint arXiv:1803.01271.
[2] Barik, N. 2018. Analysis of interventions addressing farmer distress in rajasthan. https://www.copenhagenconsensus.com/sites/ default/files/raj farmer distress sm.pdf.
[3] Chen, G. H.; Nowocin, K.; and Marathe, N. 2017. Toward reducing crop spoilage and increasing small farmer profits in india: a simultaneous hardware and software solution. In Proceedings of the Ninth International Conference on Information and Communication Technologies and Development.
[4] Cheng, H.-T.; Koc, L.; Harmsen, J.; Shaked, T.; Chandra, T.; Aradhye, H.; Anderson, G.; Corrado, G.; Chai, W.; Ispir, M.; et al. 2016. Wide & deep learning for recommender systems. In Proceedings of the 1st workshop on deep learning for recommender systems, 7–10. ACM. DARD. 2019. Vegetable Production in Kwazulu-Natal: Length of Growing Period. https://www.kzndard.gov.za/images/ Documents/Horticulture/Veg prod/length of growing period.pdf.
[5] Hastie, T.; Mazumder, R.; Lee, J. D.; and Zadeh, R. 2015. Matrix completion and low-rank svd via fast alternating least squares. The Journal of Machine Learning Research 16(1):3367–3402.
[6] Ma, W.; Nowocin, K.; Marathe, N.; and Chen, G. H. 2019. An interpretable produce price forecasting system for small and marginal farmers in india using collaborative filtering and adaptive nearest neighbors. In Proceedings of the Tenth International Conference on Information and Communication Technologies and Development,
[7] ACM. NCRB. 2019. National Crime Records Bureau. http://ncrb.gov. in/. Panagariya, A. 2008. India: The emerging giant. Oxford University Press.
[8] Sorensen, J. B. 2002. The use and misuse of the coefficient of variation in organizational demography research. Sociological methods & research 30(4):475–491.
[9] Sutskever, I.; Vinyals, O.; and Le, Q. V. 2014. Sequence to sequence learning with neural networks. In Advances in neural information processing systems, 3104–3112.
[10] Tambe, M., and Rice, E. 2018. Artificial Intelligence and Social Work. Artificial Intelligence for Social Good. Cambridge University Press.
[11] J.; Li, X.; Low, M.; Lobell, D.; and Ermon, S. 2017. Deep gaussian process for crop yield prediction based on remote sensing data. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, 4559–4565. AAAI Press.
[12] Y Sri Lalitha, SK Althaf Hussain Basha, Ayesha Mariyam, S Viswanadha Raju, "A Brief Research On Deep Learning Models", International Journal of Computer Engineering and Applications, Volume 13, Issue 6, November 2020, ISSN2321-3469.
[13] Sk. Althaf Hussain Basha, Sreedhar Jinka,Baijnath Kaushik, D.Praveen Kumar, A Jagan, "NLP: Context Free Grammars and Parse Trees for Disambiguiting Telugu Language Sentences", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT) Volume 2 , Issue 7 , pp.332-337, 2017, ISSN : 2456-3307
[14] Ayesha Mariyam, SK Althaf Hussain Basha,, S Viswanadha Raju, "A Literature Survey On Recurrent Attention Learning For Text Classification", 2nd International Conference on Machine Learning, Security and Cloud Computing (ICMLSC2020), Springer Conference 18th & 19th December2020.
[15] SK Althaf Hussain Basha, Ayesha Mariyam, and S Vishwanadha Raju "Applications of Multi- Label Classification", International Journal of Innovative Technology and Exploring Engineering(IJITEE),pp.86-89,ISSN:2278-3075,Volume-9, Issue-4S2, March 2020.
[16] Ayesha Mariyam, SK Althaf Hussain Basha Sk, and Viswanadha Raju S , "A Brief Literature Survey on Text Classification Applications", International Conference on Devices, Intelligent Systems & Communications (DISC)2020.
[17] Venkata Pavan Kumar Savala, Sk Althaf Hussain Basha, Ranganath P, P V Ravi Kumar, "Information Inclusion: The Modern Rank and The Approach Forward", International Journal of Computer Engineering and Applications(IJCEA), Volume 13, Issue 6 , December. 20, ISSN2321-3469.
[18] Sreedhar Jinka, Sk. Althaf Hussain Basha, Suresh Dara, Baijnath Kaushik, "Sequence Labelling for Three Word Disambiguation in Telugu Language Sentences", International Journal of Scientific

Research in Computer Science, Engineering and Information Technology (IJSRCSEIT) Volume 2 , Issue 7 , pp.311-315, 2017, ISSN :2456-3307.

[19] Baijnath Kaushik, Sk. Althaf Hussain Basha, Sreedhar Jinka, D Praveen Kumar, " Sequence Labelling for Two Word Disambiguation in Telugu Language Sentences", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT) Volume 2 , Issue 7 , pp.321-327, 2017, ISSN :2456-3307.

[20] Sreedhar Jinka, Sk. Althaf Hussain Basha, Baijnath Kaushik, D. Praveen Kumar, A. Jagan, " Empirical Analysis of Context Sensitive Grammars and Parse Trees for Disambiguiting Telugu Language Sentences", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT) Volume 2 , Issue 7 , pp.328-331, 2017, ISSN :2456-3307.

[21] GNR Prasad , SK Althaf Hussain Basha, Mallikharjuna Rao K M GnanaVardhan "A Review of Predictive And Descriptive Data Mining Techniques in Higher Education Domain, International Journal of Computer Engineering and Applications(IJCEA),Volume 13, Issue 6, January. 21, ISSN2321-3469.

[22] G.V.S.Raju, Sk Altaf Hussain Basha, K.Venkata Subbaih, A.V.N. Sharma, "Facilities Layout Optimization using Genetic Algorithm", International Conference on Advanced Computing Technologies (ICACT 2008), GRIET,Hyderabad,32-36,2008, ISSN: 9788178003