AgriMine: A Deep Learning integrated Spatio-temporal analytics framework for diagnosing nationwide agricultural issues using farmers' helpline data

Article *in* Computers and Electronics in Agriculture · October 2022

DOI: 10.1016/j.compag.2022.107308

9 authors, including:


Samarth Godara
Indian Agricultural Statistics Research Institute
30 PUBLICATIONS **89** CITATIONS
SEE PROFILE


Durga Toshniwal
Indian Institute of Technology Roorkee
165 PUBLICATIONS **2,758** CITATIONS
SEE PROFILE


Rajender Parsad
Indian Agricultural Statistics Research Institute
191 PUBLICATIONS **1,660** CITATIONS
SEE PROFILE


R S Bana
Indian Agricultural Research Institute
177 PUBLICATIONS **1,510** CITATIONS
SEE PROFILE

AgriMine: A Deep Learning integrated Spatio-temporal analytics framework for diagnosing nationwide agricultural issues using farmers' helpline data

Samarth Godara[a,b], Durga Toshniwal[b], Rajender Parsad[a], Ram Swaroop Bana[c,*], Deepak Singh[a], Jatin Bedi[d], Abimanyu Jhajhria[e], Jai Prakash Singh Dabas[c], Sudeep Marwaha[a]

[a]*ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India*
[b]*Indian Institute of Technology Roorkee, Uttarakhand, India*
[c]*ICAR-Indian Agricultural Research Institute, New Delhi, India*
[d]*Thapar Institute of Engineering and Technology, Punjab, India*
[e]*ICAR-National Institute of Agricultural Economics and Policy Research, New Delhi, India*

Abstract

In the current scenario, exploring new means to gain accurate information regarding agriculture-related problems is the need of the hour. In this direction, we propose a multi-stage framework to perform spatial mapping and time series analysis on more than 26 million farmers' helpline call-log records, made available by the Ministry of Agriculture, Government of India. The proposed spatial analysis framework delivers hidden patterns regarding the crop-wise density of farmers calling for help from various regions of the country. Furthermore, the proposed step-plot concept gives insights into the time span of the problems in the agriculture sector. Additionally, the proposed framework explores the potential of high-end forecasting models, including five Deep Learning-based models to predict the topic-wise demand for help (number of query calls) by the producers of the target regions. To elaborate on the utility of the presented work, the article outlines two case studies corresponding to policy recommendations regarding agriculture extension and other related domains using AgriMine.

Keywords: Artificial intelligence in agriculture, Data analytics in agriculture, Big Data, Decision making, Deep Learning, Helpline center data, spatio-temporal analysis

---

1. Introduction

Agriculture plays a vital role in providing livelihood all over the globe. It contributes to food safety and health worldwide and helps empower nations' economies. Moreover, agriculturists have always been striving to meet the increasing demand for crop production by incorporating the latest technological advancements in the sector. To optimize the usage of the agriculture sector's resources, policymakers, agri-industrialists, market experts, and the research community need to know the on-goings of the farming communities. Nevertheless, gathering day-wise information regarding the farmers' activities on a large scale seems impossible with even today's technology. This scenario calls for a robust platform to gain up-to-date information regarding the problems in farming communities. However, the low literacy rate in the farming community makes it challenging for developing countries to gain such information. A significant fraction of the farmers' community can not be reached through the latest technological advancements, including Twitter, online QA forums, blogs, etc. In this direction, the present study explores a new way to obtain novel insights regarding the farming communities (with a low-literacy rate) on a large scale.

*Corresponding author

Email addresses: `sgodara@cs.iitr.ac.in, samarth.godara@icar.gov.in` (Samarth Godara), `durgatoshniwal@gmail.com` (Durga Toshniwal), `rajender.parsad@icar.gov.in` (Rajender Parsad), `rsbana@gmail.com` (Ram Swaroop Bana), `deepaksingh2112@gmail.com` (Deepak Singh), `jatin.bedi@thapar.edu` (Jatin Bedi), `abhimanyujhajhria@gmail.com` (Abimanyu Jhajhria), `jpsdabas_catat@iari.res.in, jaiiari5@gmail.com` (Jai Prakash Singh Dabas), `sudeep@icar.gov.in` (Sudeep Marwaha)

In the past, numerous researchers have proposed systems to gather information from the agriculture sector ([1], [2], [3]). Broadly, these systems are based on computers/smartphones with internet connectivity; thus, in rural territories of developing countries, their use is extremely limited due to diverse practical problems. Similarly, for capturing information from farmers, many researchers suggested using IoT technology ([4], [5]). A major disadvantage of such systems is the pre-requisite of installing data capturing devices under actual farming conditions, making it a problematic option for large-scale data collection.

Regular monitoring of the agriculture sector and its analysis has always been a focal point for the Indian Government. For conducting surveys in various sectors (including agriculture, healthcare, etc.), the Ministry of Statistics and Programme Implementation (MoSPI), the Government of India, is the nodal agency in the country [6]. Likewise, the National Sample Survey Office (NSSO) is India's leading institute for countrywide survey design and monitoring [7]. While the computer-administered surveys have an edge and several benefits [8], to obtain data on agriculture and allied sectors and rural livelihoods, NSSO still employs conventional paper-survey methods across the country. Furthermore, most surveys designed and carried out by NSSO remain primarily focused on agricultural production. At the same time, farmers' problems-oriented surveys are rare.

In the current global scenario, only a few systems seem to exist for collecting and analyzing information regarding farmers' problems nationwide. The existing methods used to gain information about farmers' problems (surveys, camps, etc.) are too costly and time-consuming [9]. On the other hand, the low cost of the data collection in helpline centers makes it a feasible solution to gain information and analyze mob behavior. Additionally, since, through this medium, it is easy to collect data for extended periods, the helpline data have also been used by the researchers for intervention assessment for the past few years, in other domains [10]. A few of such studies include studying the mental health of children [11], psychoanalysis of women [12], suicide prevention [13], etc.

Motivated by these works, in the present study, we propose a framework to perform data analytics to mine several types of insights from the call-log records of the "Kisan Call Center" (KCC) network. KCC is a helpline service initiated by the Government of India to provide agriculture-related help to Indian farmers over telephone calls. Since the farmers' call-log records from KCC can provide beneficial inferences and hidden insights regarding the Indian farmers' behavior, several researchers have already used data mining and machine learning tools on the KCC data. In an existing study, [14] analyzed three years (2015–17) of KCC data to categorize similar questions using NLP to determine which queries are frequently asked by farmers. Besides, the researchers used Hadoop-based MapReduce algorithms to draw insights such as the hours of maximum query calls during the day, type of crops that have been enquired about more frequently, type of query, etc. In 2017, [15] proposed a text-mining-based framework to extract information from the KCC call-log records to improve the KCC system. The system had five primary phases, i.e., Data Collection, Data Integration, Data Selection, Text Mining (including frequent term identification), and Decision Making. Using the same source of information, [9] proposed a framework for mining association rules from the call-log records using the Apriori algorithm integrated with the TOPSIS technique. Furthermore, [16] presented Deep Learning-based forecasting models for forecasting the number of query calls from the target Indian states.

Although there exist many analysis-works on the same farmers' helpline dataset, due to the unavailability of profound methodologies, researchers were not able to extract the following insights from the dataset:

- Geo-coordinates plots corresponding to the farmers' queries.
- Day-wise time-series plots of the number of farmers' queries from selected regions.

- Identification of the exact dates corresponding to the different phases (uprising, peak, and downfall trend) of the problems in the farmers' community.

In the presented study, we first extract only the unique addresses present in the dataset; subsequently, we obtain geo-coordinates to the addresses and then fill in the obtained geo-coordinates in the dataset. The maps extracted using the proposed approach provide a very informative medium for the decision-makers. These give information regarding the regions where the farmers are less aware/active and correspond to the target crop. Moreover, the proposed temporal-analysis methodology, including automated query annotation and step-plot comparison, gives precise information regarding the time the selected region's farmers request help corresponding to particular crops. We first transform the available transactional data into time series to perform the temporal-based analysis. Subsequently, we use the proposed concept of "step-series" to obtain the exact date corresponding to the several phases (uprising, peak, and downfall) of the farmers' problems. Such in-depth analyses are valuable in designing agricultural policies, farmers' training, marketing strategies, and many more. Besides, obtaining accurate forecasts regarding the topic-wise demand for help (in terms of the number of query calls) plays a vital role in the decision-making process. In this direction, we use six state-of-the-art forecasting models: Fast Fourier Transform (FFT), N-BEATS, Recurrent Neural Network (RNN), Long Short-term Memory Network (LSTM), Gated Recurrent Unit (GRU), and Temporal Convolutional Network (TCN)-based model to forecast the number of queries for the next seven days.

With approx. $4 billion of an annual budget to the agriculture sector [17], including $100 million of the annual budget for agriculture extension activities [18], the Indian government has always tried to push the sector to new levels. Therefore, optimizing capital expenditure at the national and state level is a must for the country's policymakers. Furthermore, in the current scenario, there seems to exist no robust methodology which can help in accurate Spatio-temporal planning of the agriculture-related intervention activities in the country. In this direction, to demonstrate the utility of the proposed framework, the later sections of the article brief on two agriculture-based decision-making case studies, including the Spatio-temporal analysis and recommendations corresponding to the Indian farmers' problems regarding the mushroom crop and for the rice-related issues from the producers of the Uttar Pradesh state.

The remainder of the article is organized as follows: Section 2 describes the data and the methodology used to develop the proposed AgriMine framework. A detailed description of the experiments, results of the study, and case studies associated with them are presented in section 3. A brief discussion on the policy recommendations from the extracted insights is given in section 4, and a summary of the proposed study is given in section 5.

2. Methodology

The methodology proposed in the study can be divided into four phases, i.e., data collection, pre-processing, spatial analysis, and temporal analysis. In the initial phase, the data is collected from the KCC data servers. Later, the downloaded data is processed and transformed into the desired form. The processed information is then fed to two modules (spatial and temporal analysis) independently to extract insights from the dataset (figure 1). The remainder of the section gives details of each of the modules present in the proposed AgriMine framework.

1. Data Collection: In 2004, the Government of India launched a flagship scheme entitled "Kisan Call Centres" (K.C.C.) to harness the information and communication potential in agriculture [19]. Through the project, using a toll-free helpline number "1800-180-1551", agricultural queries of the farmers are received, and solutions are provided to the farmers instantly. The helpline operators maintain records of all the questions, answers,

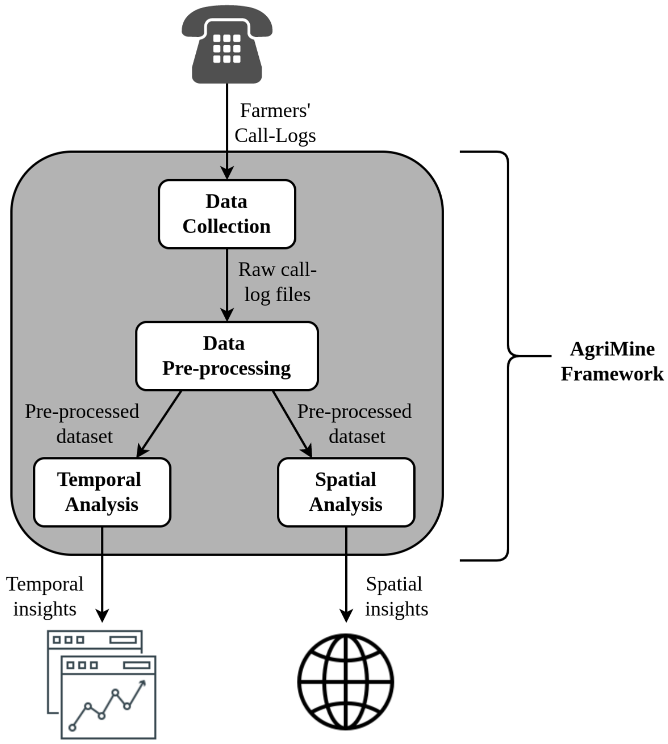


Figure 1: Block diagram of the AgriMine framework

crop category, type of query, date and time of call, and the farmer's location (table 1). Moreover, the "Kisan Knowledge Management System" (K.K.M.S.) under the Ministry of Agriculture and Farmers Welfare, Government of India, maintains the records and are made available as open access in .json format [19]. The call-log records are also maintained on the *Open Government Data Platform India* [20].

The KKMS servers maintain separate call record files corresponding to each district for each month. Furthermore, the call records used in this study are from March 2013 till February 2021. In the study, a custom web-crawler is designed to download the files automatically in an iterative manner. Overall, 55,844 raw data files are downloaded in the study, including ≈26.8 million farmers' call-log records. Each call-log record is in .json file format with a total of 11 attributes. Details of each attribute present in the record files are given in table 1.

Table 1: Information in the downloaded call-log records

S.No.	**Attribute Name**	**Description**
1.	BlockName	Name of the farmer's block
2.	Category	Category of the query, for example, "Avian", "Beekeeping", "Cereals", etc.
3.	CreatedOn	Date and time of the query-call
4.	Crop	Crop regarding which the farmer enquired
5.	DistrictName	Name of the farmer's district
6.	KccAns	The answer communicated to the farmer by the helpline center
7.	QueryText	The question asked by the farmer to the helpline center

Table 1: Information in the downloaded call-log records

S.No.	Attribute Name	Description
8.	QueryType	Type of the query, for example, “Plant Protection”, “Weed Management”, etc.
9.	Season	Season in which the query was made, including, “Jayad”, “Kharif”, and “Rabi”.
10.	Sector	The sector that the query belongs to, for example, “Animal Husbandry”, “Fisheries”, “Horticulture”, etc.
11.	StateName	Name of farmer’s state , for example, “Punjab”, “Haryana”, “Maharastra”, etc.

2. Data pre-processing - In this phase, the raw call-log data files are transformed to a usable single dataset file. The Big-data is handled using ‘pandas’ python package. Moreover, the whole pre-processing phase can be further divided into following four sub-steps (figure 2):

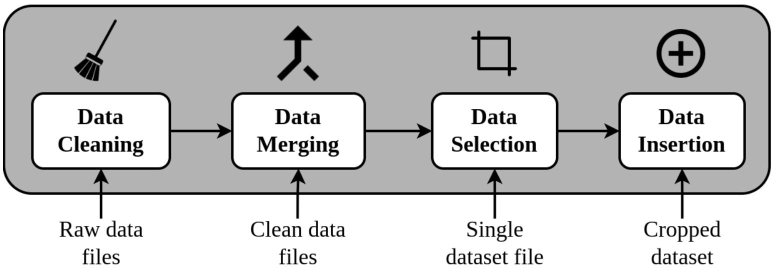


Figure 2: Block diagram of the Data Pre-processing phase

   (a) Data cleaning is generally the first step in data mining. It involves the removal of erroneous data, filling null values, and correcting the instability in the dataset. In the data cleaning step, all characters except the alphabets, numbers, and some special characters (including ‘,’ and ‘-’) are removed from the dataset.

   (b) Data merging is the process of combining datasets from multiple sources into a single location. The complete raw data used in the study consists of 55,844 separate files. To bring all the records together in a single file, we merge data, where every record present in the downloaded files is read and stored in a common file.

   (c) Data selection selects only the relevant data section from the complete dataset. Therefore, after obtaining the dataset in a single file, the attributes that are irrelevant to our study, including “Category”, “KccAns”, “QueryText”, “Season”, and “Sector” are removed from the dataset—removing these from the dataset help in memory-efficient processing of the subsequent phases.

   (d) Data insertion involves the addition of new rows/columns in the dataset. In this step, the attribute “CreatedOn” is partitioned to form three new attributes named “Year”, “Month”, and “Date”. This step helps in grouping the data based on the combinations of “Year” and “Date”. The pre-processed dataset contains eight attributes of the raw data in it, including “BlockName”, “DistrictName”, “StateName”, “QueryType”, “Crop”, “Year”, “Month”, and “Date” (table 1).

3. Spatial Analysis: In this phase, the geo-coordinates corresponding to the queries related to the target crops are plotted on the country map, the information regarding the block name, district name, and state name are used to obtain the query calls' geo-coordinates. The overall task is divided into the following six sub-steps (figure 3):

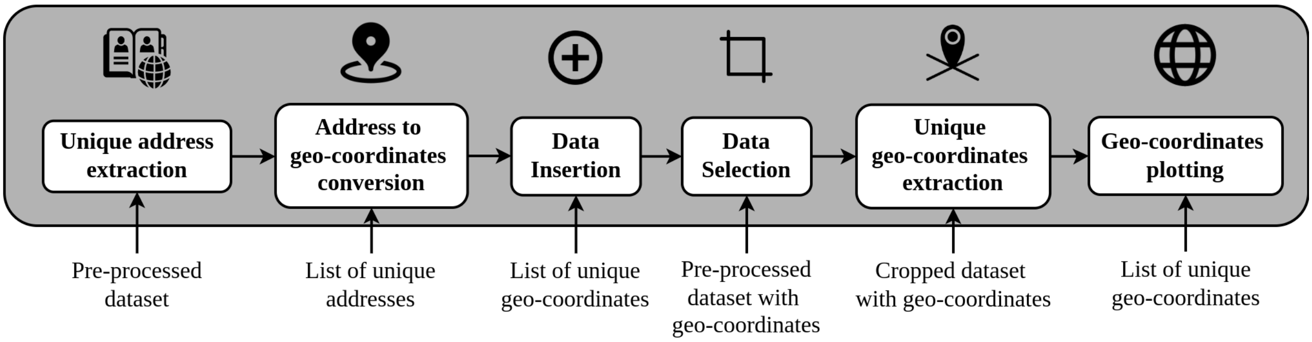


Figure 3: Block diagram of the Spatial-analysis phase

(a) Unique address extraction: In this sub-step, the addresses associated with all the call-log records are initially separated. Later, the duplicate addresses present in the address list are removed to obtain a list of unique addresses in the dataset. The extracted unique addresses are in the form of the combination of the block, district, state, and country name of the farmer (figure 4). After the processing, a total of 7925 unique addresses were obtained in this sub-step.

(b) Address to geo-coordinates conversion: To obtain the latitude and longitude corresponding to the set of unique addresses, the python packages "opencage" and "geopy" are used. These packages provide the geo-coordinates corresponding to any address string (consisting of the block, district, state, and country name) that is fed to the modules. The packages mentioned above return a list of geo-coordinates corresponding to the input address string, from which the first set of latitude and longitude is selected (figure 4).

(c) Data insertion: Once the unique latitudes and longitudes are obtained corresponding to the unique addresses of the dataset, the information is inserted in the processed dataset. First, the list of unique geo-coordinates is expanded with duplicate records corresponding to all the call records in the dataset. Later two new attributes are created named "Latitude" and "Longitude" in the dataset, and the values are filled in the attributes.

(d) Data selection: To plot the geo-coordinates corresponding to the target-crop queries, a subsection corresponding to the target crop is selected from the complete dataset. This step matches the input crop name with the values present in the "Crop" attribute of the dataset. Later, all the records with the positive-match results are reverted to the subsequent step.

(e) Unique geo-coordinates extraction: After obtaining the crop-related dataset, the latitudes and longitudes are separated from it in list form. Later, all the duplicate copies are removed from the list to get distinctive locations. This process gives the list of unique geo-coordinates (combinations of latitudes and longitudes) to be plotted on the map.

(f) Geo-coordinates plotting: The last sub-step of the spatial analysis framework first plots the Indian-states' border lines on the map. Later, the latitude and longitudes obtained in the previous sub-step are plotted.

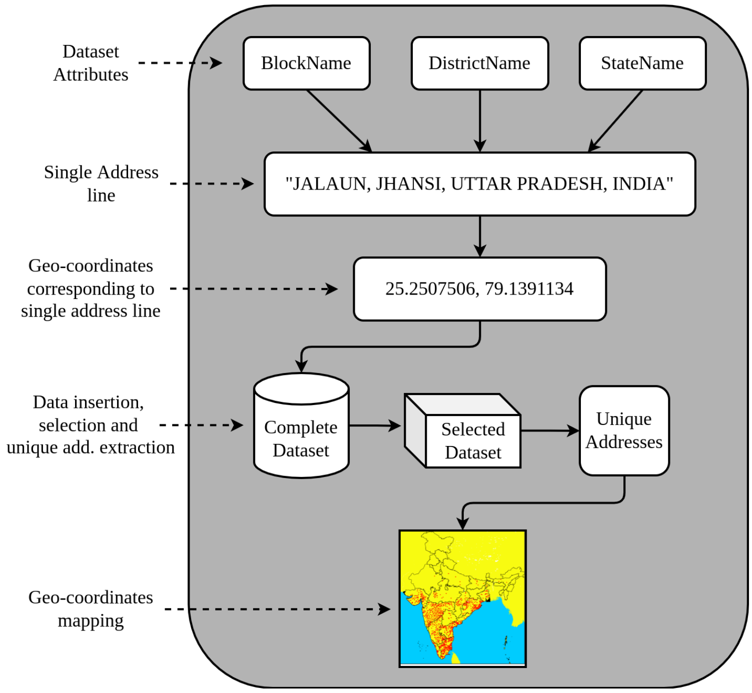


Figure 4: Output examples of the spatial analysis steps

It is to be noted that the sub-steps (a), (b), and (c) are to be executed only once. Subsequently, multiple requests of plotting geo-coordinates can be carried out by executing sub-steps (d), (e), and (f) iteratively.

4. Temporal Analysis: In this phase, the temporal insights are extracted from the dataset in terms of the per-day number of query calls made by farmers from the target region (state) of India. Through the proposed methodology, we obtain time-series specific to the selected crop, corresponding to the target region, and associated with the chosen query category. Additionally, we propose the concept of step-series, a new way of representing the different phases of agricultural problems. The complete temporal analysis module can be divided into three steps (figure 5).

   (a) Time series extraction: The first step of the temporal-analysis phase involves extracting the per-day query count corresponding to the combination of the target state, crop, and query category. This step is further divided into the following four sub-steps:

      i. Data Selection: In this sub-step, the data corresponding to the combination of the input state, crop, and query category is separated. The sub-step selects the records found to have a match for the input parameters against the values present in the "StateName", "Crop", and "Category" attributes.

      ii. Data transformation: In order to obtain a time series t (equation 1) from the selected dataset, the data is transformed from its transactional nature into time series using equation 2.

$$t = \big(t_i : i \in N \mid i \leq (365 \times 8)\big) \tag{1}$$

$$t_i = \frac{1}{m}\sum_{j=1}^{m} n(s_{ij}) \tag{2}$$

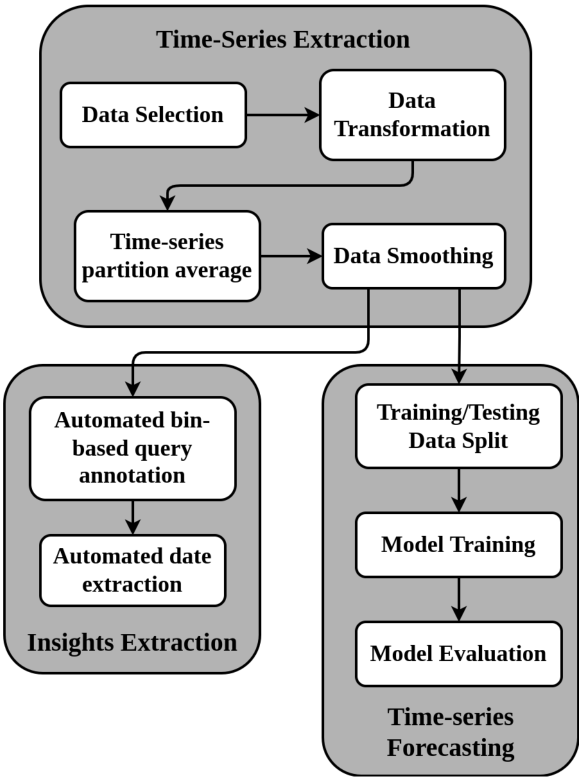


Figure 5: Block diagram of the Temporal-analysis phase

Where, t is the output time-series, t_i represents the average number of queries made on day i, m represents the total number of years during which the data is collected, s_{ij} represents the set of all query-calls made on the i^{th} day of the j^{th} year, and $n(s_{ij})$ represents the cardinality of set s_{ij}.

iii. Time-series partition average: Until this point, the obtained time-series corresponds to eight years' data (March 2013 to February 2021). In this sub-step, the complete time-series is partitioned year-wise (equation 3), and later, day-wise data points of the partitions are averaged to obtain a single-year-length time-series (equation 4).

$$t_l^k = t_{\left(l+(365\times(k-1))\right)} \tag{3}$$

$$f_l = \frac{1}{n}\sum_{k=1}^{n} t_l^k \tag{4}$$

$$f = (f_l : l \in N \mid l \leq 365) \tag{5}$$

Here, t^k represents the k^{th} partition of the time-series t, and n represents the total years the data corresponds to. f in equation 5 represents the final time series obtained after calculating the average per-day query count for all the years.

iv. Data smoothing: For better analysis and plotting results, the obtained time-series is smoothened using Savitzky–Golay filter [21] with the general parameter settings (input window-length parameter of 21 data points and input polynomial order one). Although the savgol filter has a computation time proportional to the window width, it has the advantage of maintaining the position and width of the peaks, which is helpful for our analysis. The smoothened per-day query-count time-series data is later plotted with the x-axis representing the yearly timeline and the y-axis representing the number of queries received (figure 7). The plotted time series represents the average per-day query call of the data points from all the target years.

(b) Insights extraction: In this module, we extract the temporal insights regarding the topic-wise problems corresponding to the target state and crop in the form of step series and the dates corresponding to the various phases of the demand for help. The module is divided into the following two parts.

i. Automated bin-based query annotation: The time-series plots obtained till the previous steps are generally bell-shaped curves (figure 7). The plots help extract insights regarding the farmers' behavior, but these do not provide the exact starting and ending dates of different phases of the farmers' problems. To categorize the per-day query-count time-series data points, we propose an annotation-function $\psi(t_i)$, to marginalize the data points to extract dates through automated binning (equation 6). With the proposed function, we annotate the time-series data points into one of the four categories, i.e., "High-frequency query call day", "Medium-frequency query call day", and "Low-frequency query call day", and "No query call day". For each input time series, $\psi(t_i)$ creates virtual bins, where the range of each bin is approximately one-third the range of the whole time series. Moreover, the output "step-series" can be mathematically described by equation 7.

$$\psi(f_i) = \begin{cases} \mathit{0}, & \text{if } f_i < 5 \\ \mathit{1}, & \text{if } 5 \leq f_i < \frac{m}{3} \\ \mathit{2}, & \text{if } \frac{m}{3} \leq f_i < \frac{2m}{3} \\ \mathit{3}, & \text{otherwise} \end{cases} \tag{6}$$

$$s = \big(\psi(f_i) : i \in N \mid i \leq 365\big) \tag{7}$$

In equation 6, f_i represents the time-series data point that is to be binned, and m represents the largest value present in the time series f. In equation 7, s represents the step-series obtained after the annotation. The value of $\psi(f_i)$ being *0* represents that the data point (particular day) of the step-series belongs to the "No query call day" category, *1* represents the "Low-frequency query call day" category, *2* represents the "Medium-frequency query call day", and *3* represents the "High-frequency query call day" category.

ii. Automated date-extraction: The dates corresponding to the "corners" present in the step-series provide essential insights regarding starting or ending of the phases of farmers' problems. In this step, all the dates associated with the "corners" in the step-series are extracted using algorithm 1. The algorithm rolls a variable over the step-series and stores all the dates where a transition is noted (Output sample table 2).

(c) Time-series forecasting: The whole process of developing the 7-day forecasting models for the query-count time-series corresponding to the target topic from the selected combination of state and crop can be divided into the following three sub-steps.

i. Training/Testing data split: First, the obtained time series corresponding to seven years (January 2014 - December 2020) is divided into two parts. The first six years (2014-2019) of data points are

Algorithm 1: Automated date-extraction algorithm

Input : s (step-series)

Output: d (list of starting and ending of the different phases of the farmers' problems)

$d \leftarrow \phi$;

$i \leftarrow 1$;

while $i \leq 364$ **do**

    if $s[i] \neq s[i+1]$ **then**

        $d.append\_at\_end(i, s[i])$;

        $d.append\_at\_end(i+1, s[i+1])$;

    $i \leftarrow i + 1$;

end

$last\_rec \leftarrow d.pop\_first\_element()$;

$d.append\_at\_end(last\_rec)$;

used for training the forecasting models, and the time-series corresponding to the last year (January 2020 - December 2020) is used to evaluate the performance of the forecasting models on the unseen data.

ii. Model Training: In the present study, six forecasting models, including five Deep Learning-based models, are trained and tested. Deep learning is a category of machine learning algorithms that progressively employs multiple layers to extract higher-level features from the raw input. For example, lower layers may identify boundaries in image processing, while higher layers may determine the concepts relevant to a human, such as digits, letters, or faces. Deep neural networks are usually interpreted using the universal approximation theorem, or probabilistic inference [22]. The probabilistic interpretation emanates from the field of supervised learning. It features inference, training, and testing optimization ideas related to fitting and generalization.

Furthermore, RNNs are a class of networks useful in modeling sequence data. Derived from feedforward neural networks, RNNs display similar behavior to how human brains function [23]. Whereas TCN is a framework that employs simple convolutions and dilations, it is adaptive for sequential data with its temporality and large sensory fields [24].

The architectures of the models used in the study, based on the FFT, N-BEATS, RNN, LSTM, GRU, and TCN are described in [25], [26], [27], and [28] respectively. Each model is trained on the training dataset containing 2557 time-series data points.

iii. Model Evaluation: The forecasting performances of the models on the training dataset and the testing dataset are calculated separately. The latest 365 time-series data points corresponding to 2020 are used to test the models on the unseen data. To compare the models' performances, two metrics, i.e., Averaged MSE (AMSE) and Averaged MAE (AMAE) [29] are taken into account (equation 8 and 9).

$$AMSE = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} (Y_{ij} - \hat{Y_{ij}})^2 \tag{8}$$

$$AMAE = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} |Y_{ij} - \hat{Y_{ij}}| \tag{9}$$

Here, m represents the total number of time series the forecasting model is trained on, n is the number of output data points, $\hat{Y}_{ij}$ is the output of the forecasting model, and Y_{ij} is the expected output.

3. Experiments and results

In the presented study, we have focused on extracting four types of insights from the KCC dataset. These insights are obtained through geo-coordinates maps, time-series plots, step-plots, and time-series forecasts. The remainder of the section elaborates on each of the obtained insights in detail.

3.1. Geo-coordinates Mapping

To verify the spatial properties of the available dataset, initially, we plotted the geo-coordinates corresponding to all the query calls that are present in the farmer's helpline dataset (figure 6 (a)). From the figure, it is observed that farmers from all over the country use the KCC service. The figure gives the locations of the farmers who had asked questions in the eight-year time span (March 2013 to February 2021). Figure 6 (b) plots the geo-coordinates of the clustered query calls based on the call-density using DBSCAN clustering algorithm [30]. From the figure, it is noted that the farmers' calls are dense in the irrigated regions of India, i.e., the region of the alluvial soils of Indo-Gangetic Plains and red soil region of Krishna-Tungabhadra region and Kaveri region [31].

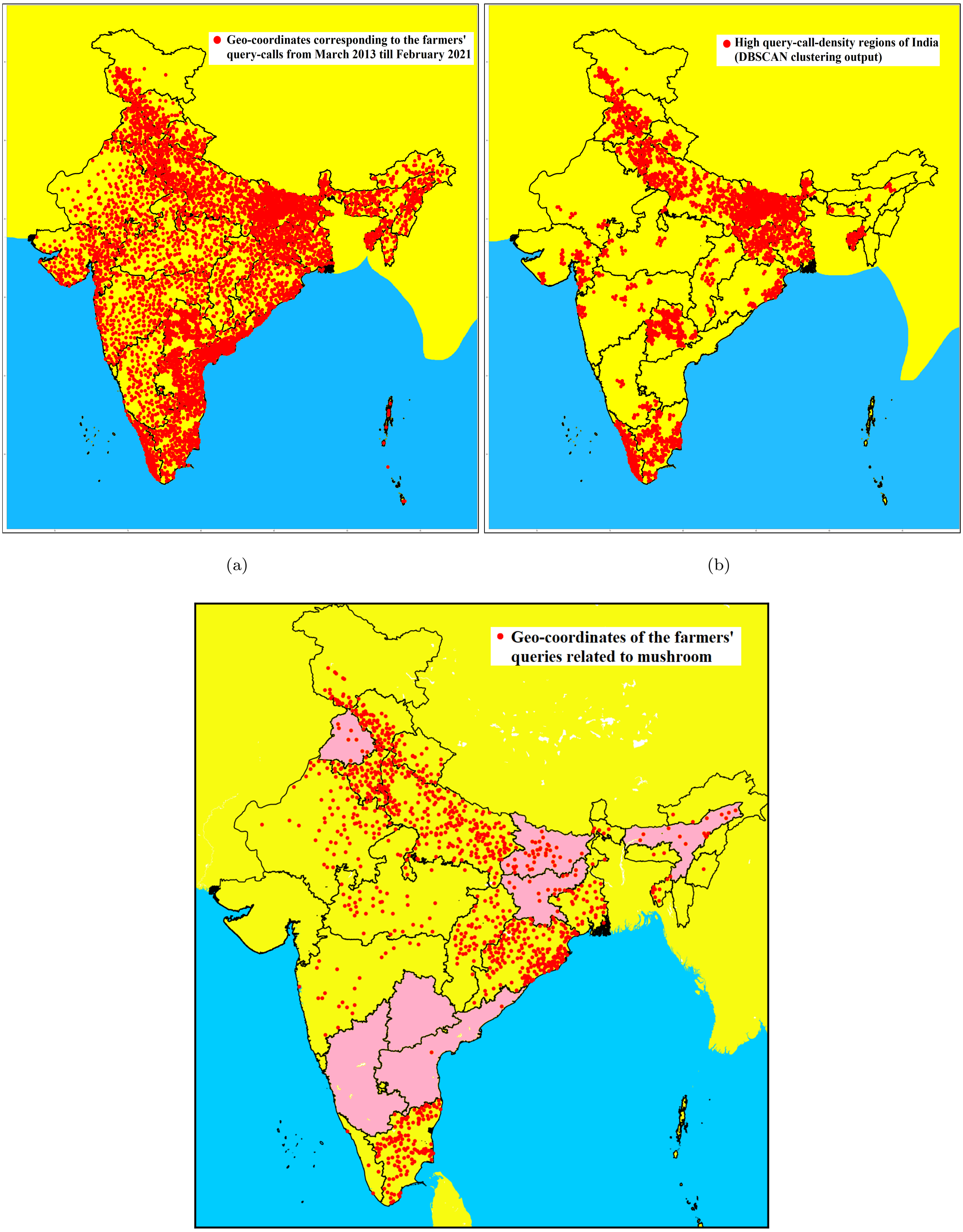
Furthermore, figure 6 (a) also gives information regarding the Indian regions where the farmers made very few query calls in the past eight years. The calls are lesser in number from the mountainous regions of the Himalayas (extreme northern India) and the Thar desert (extreme western India) due to less population density in these regions.

To show the potential of the proposed spatial analysis, we consider the crop of mushroom, which is proven to be both nutritionally and economically beneficial to Indian farmers ([32], [33]). Since, mushroom crop cultivation is relatively more profitable [34], and the Indian government's recent focus is on increasing farmers' income [35], this crop is a potential candidate for enterprise diversification in the commercial farming systems. The geo-coordinates of the query-calls corresponding to the mushroom crop are shown in figure 6 (c).

Although this crop's nutritional and economic benefits are high, farmers from some states (with a suitable environment) have still not accepted this crop. The reason may be lack of awareness, insufficient training, poor access to technology, cultural barriers, and low risk-bearing ability of farmers. From the spatial analysis, we pointed out a group of 7 states (i.e., Punjab, Bihar, Jharkhand, Assam, Telangana, Andhra Pradesh, and Karnataka) where the crop can be grown and can be used for the betterment of farmers [36], [37].

3.2. Time-series plots

In the present study, we extract the temporal insights related to the farmers' queries regarding the rice crop from the state of Uttar Pradesh. Figure 7 plots the daily average number of query-calls corresponding to the five query types (Plant protection, Weed management, Seeds/Varieties, Water management, and Fertilizer/Nutrient usage) related to the combination of the target crop and state. The figure illustrates that farmers of the Uttar Pradesh state start asking queries regarding the seeds/variety at the beginning of the crop season, i.e., mid-April. Furthermore, it is observed that the queries related to plant protection are substantial in number when compared to other types of queries. The plant protection queries start at the beginning of the season (end of April) and end after all the



(a) (b) (c)

Figure 6: Geo-coordinates plot of the farmers' query calls made from March 2013 to February 2021: (a) all query-calls made in the period, (b) High query-call-density regions of India, (c) geo-coordinates related to mushroom queries, potential states for the mushroom crop are coloured in pink.

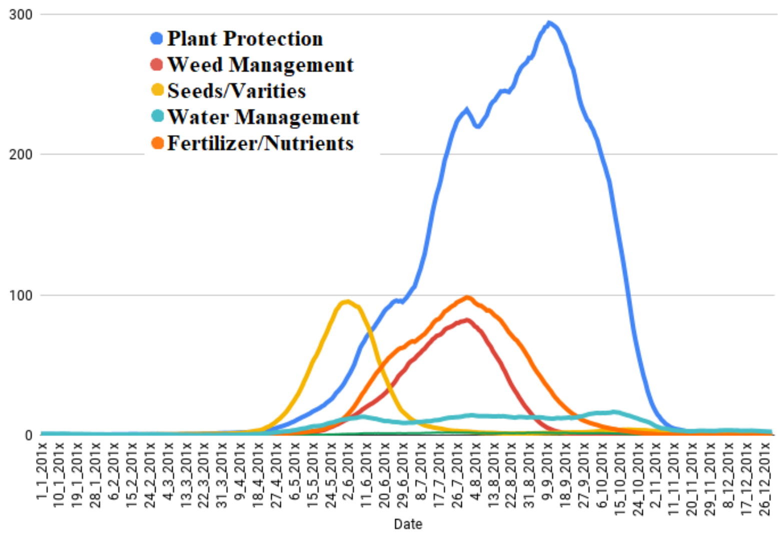


Figure 7: Time-series plot corresponding to the number of query-calls made by farmers regarding several topics (plant protection, weed management, seeds/variety, water management, and fertilizers/nutrients) regarding rice crop from Uttar Pradesh state.

other types of queries (mid of November) [38]. Furthermore, the queries related to fertilizer/nutrient usage and weed management seem to occur simultaneously, with similiar intensity. Whereas the queries related to water management are comparatively low in number, the intensity seems consistent throughout the crop.

3.3. Automated bin-based query annotation

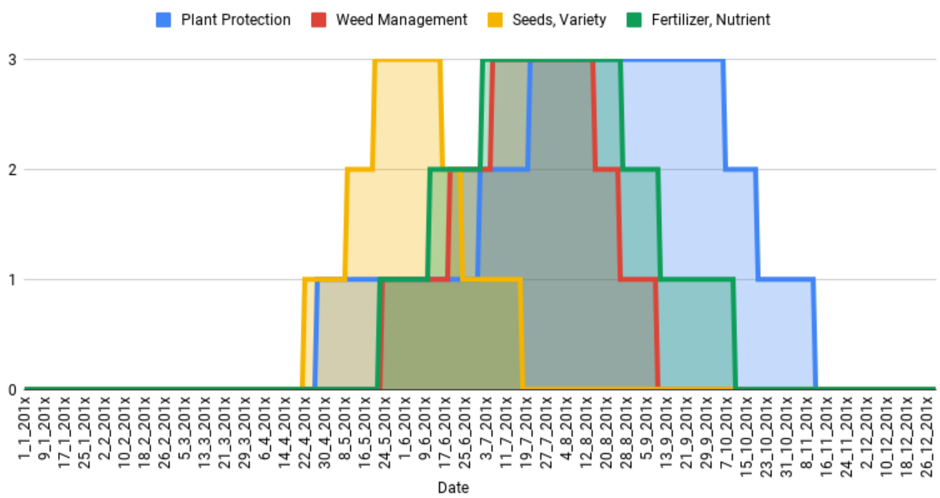


Figure 8: Step-plot corresponding to the queries made by farmers of Uttar Pradesh regarding the rice crop.

For extracting the dates corresponding to the different phases of the farmers' problems, the time-series data points are annotated into one of the four categories, i.e., "No query call days", "Low-Frequency Call Days", "Medium-Frequency Call Days", and "High-Frequency Call Days". The proposed automated bin-based query annotation algorithm is for this process. An example of the algorithm output is shown in figure 8 (the output corresponding to the figure 7 time-series input). The discrete steps in the step plots help obtain the exact starting and ending date of the different frequency periods.

From figure 8, it is observed that the queries related to seed/variety start around the 22$^{\text{nd}}$ of April and are asked till 19$^{\text{th}}$ of August (matching with the *kharif* season sowing window) [39]. Table 2 includes the output of the

Query-type	Frequency-type	Start-date	End-date	Frequency-bin
Seeds & Variety	Low Frequency	22-Apr	08-May	5-31
	Medium Frequency	09-May	19-May	31-63
	High Frequency	20-May	15-Jun	63-95
	Medium Frequency	16-Jun	23-Jun	31-63
	Low Frequency	24-Jun	17-Jul	5-31
	Random-calls	18-Jul	21-Apr	0-5
Fertilizer & Nutrient Usage	Low Frequency	22-May	10-Jun	5-32
	Medium Frequency	11-Jun	01-Jul	32-65
	High Frequency	02-Jul	26-Aug	65-98
	Medium Frequency	27-Aug	10-Sep	32-65
	Low Frequency	11-Sep	10-Oct	5-32
	Random-calls	11-Oct	21-May	0-5
Weed Management	Low Frequency	23-May	18-Jun	5-27
	Medium Frequency	19-Jun	05-Jul	27-54
	High Frequency	06-Jul	15-Aug	54-82
	Medium Frequency	16-Aug	25-Aug	27-54
	Low Frequency	26-Aug	09-Sep	5-27
	Random-calls	10-Sep	22-May	0-5
Plant Protection	Low Frequency	27-Apr	30-Jun	5-97
	Medium Frequency	01-Jul	20-Jul	97-195
	High Frequency	21-Jul	06-Oct	195-293
	Medium Frequency	07-Oct	19-Oct	97-195
	Low Frequency	20-Oct	11-Nov	5-97
	Random-calls	12-Nov	26-Apr	0-5

Table 2: Extracted dates from the step-series corresponding to the queries made by the farmers from the Uttar Pradesh state regarding the Rice crop

date-extraction algorithm with the input step-series of Uttar Pradesh - Rice (figure 8). The first row of the table represents the "Low-frequency query call days" frequency type, corresponding to the "Seeds/Variety" query type, starting from 22$^{\text{nd}}$ April, till 8$^{\text{th}}$ May, moreover, the number of calls made by the farmers in this phase is in range of 5-31.

3.4. Query-count time-series forecasting

In the present study, the performance of six forecasting models are evaluated on the time series corresponding to the query calls regarding four topics (i.e., plant protection, weed management, fertilizer usage, and seeds/varieties) for the rice crop from the state of Uttar Pradesh. For the comparison, 24 models (4 topics $\times$ 6 techniques) are trained, figure 9 illustrates the model-wise performance comparison. Each forecasting model is designed to generate 7-data points (days) time series using the past year of data (365 time series data points). Furthermore, since each forecasting model is trained and tested on four different time series, AMSE and AMAE are calculated and compared

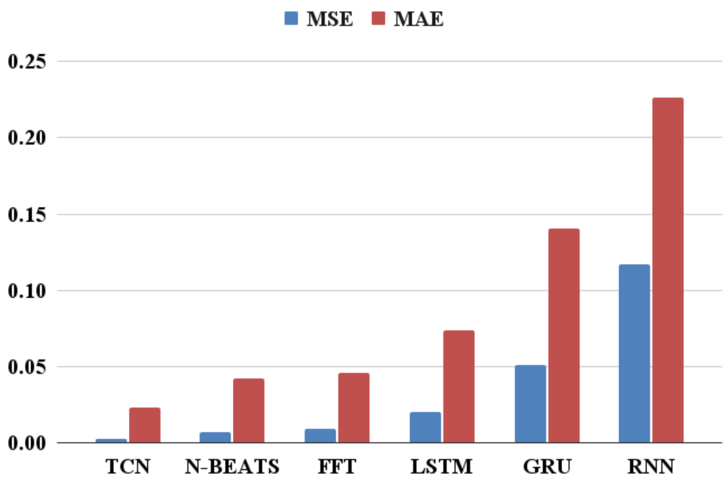


Figure 9: AMSE and AMAE comparison of the forecasting models on the testing data

to assess the models' performances. For the training of the models, time-series data corresponding to the period of March 2013 till December 2019 is considered, and time-series data generated from January 2020 till December 2020 is taken for testing purposes. From the obtained results, it is observed that the TCN-based models have outperformed the other models (with AMSE = 0.0025, AMAE = 0.023), followed by the N-BEATS and FFT-based models (table 3). Moreover, in the study, the three sequential deep learning models, including LSTM, GRU, and RNN, have shown the highest error rates among the other forecasting models. The forecasting output of the TCN-based model corresponding to the time series of the target topics is shown in figure 10.

S.No.	Model	AMSE	AMAE
1.	TCN	0.0025	0.0230
2.	N-BEATS	0.0072	0.0421
3.	FFT	0.0087	0.0455
4.	LSTM	0.0200	0.0735
5.	GRU	0.0506	0.1405
6.	RNN	0.1173	0.2260

Table 3: AMSE and AMAE of the Deep Learning-based forecasting models

4. Discussion

The section elaborates on the possible reasons behind the extracted insights along with multiple decision-making applications . Overall, the recommendations are directed towards optimization of the resources, including maximizing addressing the producers' problems while minimizing the state-level financial expenditure on agricultural extension, marketing, and various other production-related interventions.

4.1. Geo-coordinated mapping of the mushroom-crop related queries

The geo-coordinates map plots corresponding to the queries related to the mushroom crop (figure 6 (c)) highlight that in the states of Haryana, Uttar Pradesh, and Tamil Nadu, farmers from across the state have been asking

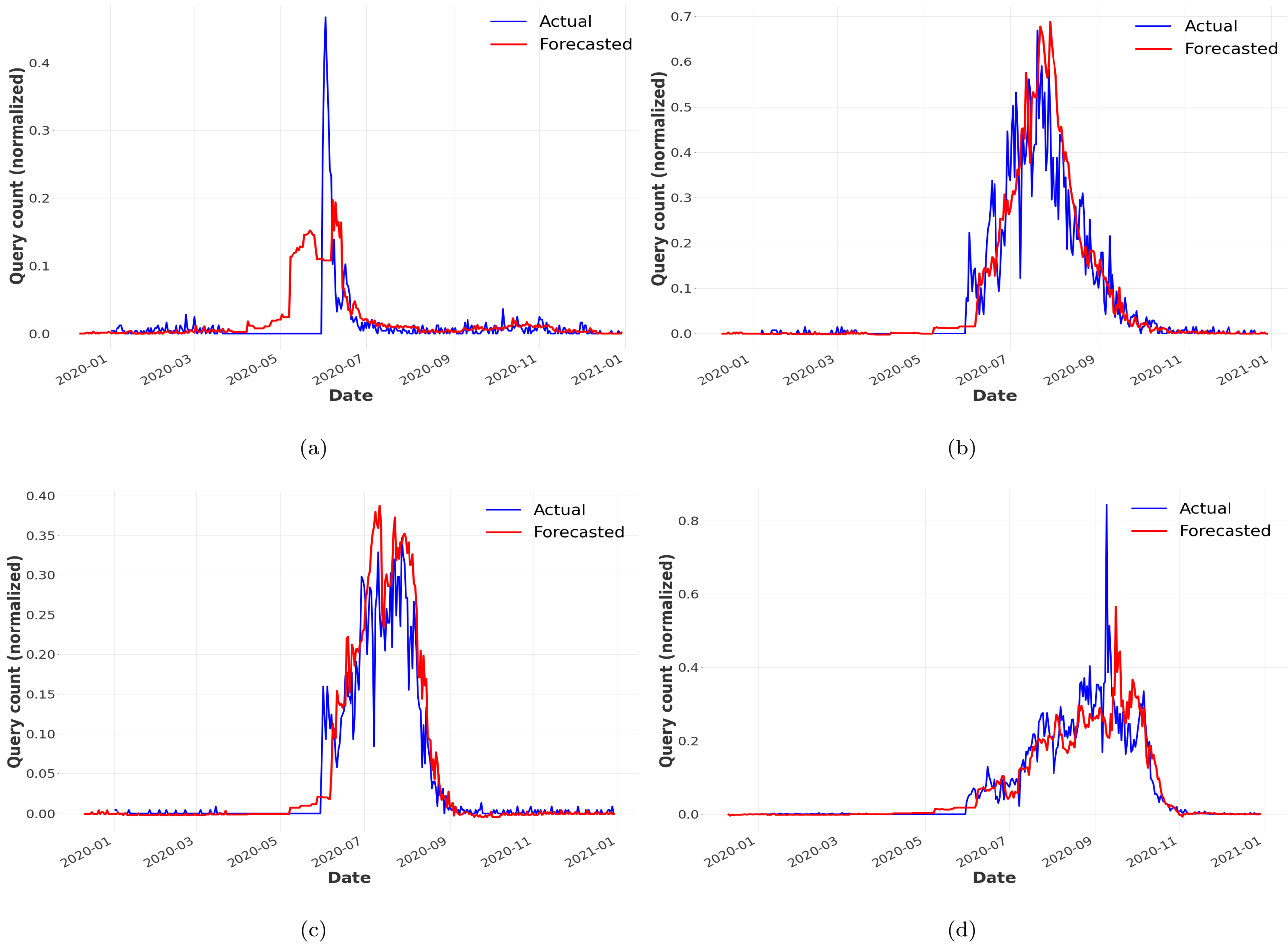


Figure 10: Forecasting output of the TCN-based models corresponding to the query-count time-series (2020) of rice crop from the Uttar Pradesh state for the topic a) Seeds and Varieties b) Fertilizer Usage c) Weed Management d) Plant Protection

questions regarding the target crop. It seems that this behavior of farmers is because of the high demand for the mushroom crop in metropolitan cities. For example, the Delhi-NCR region is in high demand for mushroom crop [40]. Therefore, the nearby states, including Haryana and Uttar Pradesh-located farmers, are interested in cultivating the crops for better income. Similarly, due to the high demand for the mushroom crop in the Chennai metropolitan, farmers from Tamil Nadu state also demand help in growing mushroom crop increasingly [41]. Whereas, in some states, questions were asked from particular locations only. For example, in the Uttarakhand state, the southern district farmers are demanding help, and it appears that farmers from the northern part are not showing interest in the mushroom crop. This is because the state's capital city Dehradun has a high demand for the mushroom crop [42], and is situated in the south of the state. Hence, farmers from the only southern part of the state are interested in supplying the crop due to transpiration limitations in the state's mountainous region.

The extracted insights recommend directing the agricultural extension resources to the highlighted parts of the states instead of targeting each district. Furthermore, the marketing of mushroom-related products (including seed, compost, etc.) must be focused on the target regions to optimize the marketing investments.

In figure 6 (c), we pointed out a few states where the crop can potentially benefit the producers economically and provide nutrient-rich diets to the people. In this direction, the recommendation is that the national-level decision-makers should focus on the states, including Punjab, Bihar, Jharkhand, Assam, Telangana, Andhra Pradesh, and Karnataka, to introduce the crop to the producers.

4.2. Time-series and Step-series plots of queries corresponding to the rice crop from Uttar Pradesh state

Figure 7 illustrates the time series plot of farmers' query call count corresponding to the rice crop from the Uttar Pradesh state. From the plot, it is found that the queries related to "seeds and varieties" are asked earlier than the other types of queries. Furthermore, during the sowing, farmers apply a basal dose of fertilizers and at the same time, they are also concerned about weed problems that may emerge soon after crop establishment [43]. Therefore, the queries related to "fertilizer/nutrient" and "weed management" seem to come and go approximately at the same time of the year. At last, due to the fact that most of the insects, pests, and diseases infest the crops at the later stage of the crop life cycle, the "plant protection"-related queries come in the later stage of the season.

Upon investigating the time-series plot of the per-day query count from the state of Uttar Pradesh regarding the crop of rice, it was found that the number of calls related to plant protection is much higher than the rest of the query types (Figure 7). One possible reason behind such behavior is that compared to other cereal crops, rice suffers more from pest and disease attacks [44]. Moreover, in rice, the Basmati varieties are relatively more susceptible to pests and diseases [45], which is grown in Uttar Pradesh state on a large scale.

The presented concept of the step-series and the dates extracted from it can play a significant role in temporal decision-making of the state-level agricultural extension and marketing activities. As the existing studies indicate that extension mediums differ in cost-effectiveness [46], the available mediums of the extension activities can be classified into three major groups based on the intervention cost and its effectiveness. The first group (G1) of extension mediums including low costing extension means, for example, newspaper/magazine articles. The second group (G2) of extension interventions consisting of the medium-cost means, for example, radio and television-based extension. Finally, the third group (G3) including the cost-intensive and effective extension means comprising one-to-one or group training activities [47].

Since the dates extracted in the study correspond to three different query-calls-frequency classes (table 2, excluding random calls days), the planning of the state-wise extension activities groups (G1, G2, and G3) can be done accordingly. From the extracted dates, it is recommended to carry out the G1 activities in the state of Uttar Pradesh related to the seeds and varieties from the starting date of the first low-frequency period (22nd April) till the end of the second low-frequency period (17^{th} July). Similarly, the G2-related activities are recommended to be executed in the period of 9^{th}May till 23^{rd}June. Furthermore, it is suggested to execute G3 extension activities when the demand for help regarding this topic is at its peak, i.e., 20^{th}May till 15^{th}June. This planning scheme is expected to optimize the state-level financial expenses as the intervention of the extension programs are designed according to the cost and effectiveness of the mediums and the need of the producers. The cut-off dates optimize the expenditure on such activities by reducing the resource burden on the state exchequer and on agro-industries. Whereas all three types of extension activities come into play at the time when the farmers from the target region need help the most (figure 11). Similarly, decisions can be made for the extension regarding Fertilizer usage, Weed Management, and Plant Protection related topics. Furthermore, these insights are also beneficial in deciding the agricultural-technology intervention time. Moreover, the interventions related to the introduction of the new variety, fertilizer, or any other related technology can be done in the suggested windows.

In addition, the same is helpful for the agro-industries in designing marketing strategies. Like the extension means, the marketing mediums can also be divided into classes based on cost-effectiveness. Furthermore, each marketing activity group can be executed according to its requirement in the corresponding window.

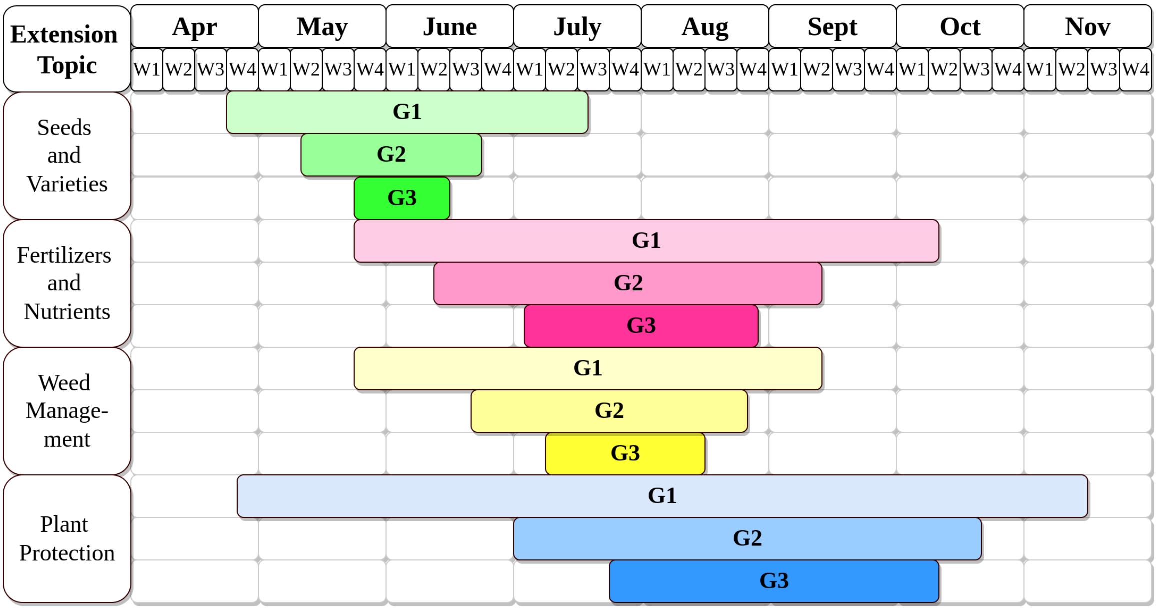


Figure 11: Timeline representation of the recommended topicwise agricultural extension planning scheme for the rice crop in the Uttar Pradesh state. Here the groups G1, G2, and G3 represents the low, medium and high-costing extension mediums, respectively.

4.3. Deep learning-based query-count time-series forecasting

The performance comparison results indicated that TCN-based models give better performance than other models, similar results are obtained in other studies [48]. The dates extracted from the step-series estimate the various query-calls-frequency periods. Obtaining an accurate forecast of the number of calls can play a significant role in resource management in various fields. The forecast helps manage the workforce of the agricultural helpline centers as it can be used to determine the demand and trend of help for the next 7-days. Moreover, the forecast can also be used in short-term accurate extension and marketing planning regarding the selected crop. Furthermore, such models play a crucial role in designing early warning systems, recommender systems, and expert systems for the producers of the target region.

5. Conclusion

To cope with the increasing global demand for food, the latest technological advancements must be incorporated into the farming sector. Lack of an efficient system to extract insights regarding farmers' problems on a large scale calls for introducing a cost-effective and reliable platform for the same. In this direction, the proposed work presents new methodologies for spatial and temporal analysis of the helpline data to extract hidden patterns in the farmers' behavior on a national scale.

In the study, the presented framework of AgriMine is used to obtain insights from one of the world's largest farmers' helpline networks, i.e., India's Kisan Call Centers. The derived spatial inference from the proposed approach gives compelling knowledge concerning the location of farmers requesting help regarding the mushroom crop. Moreover, the obtained temporal information presents an in-depth perception of the crop-wise time for the year the farmers of the target region request help. Furthermore, from the executed forecasting experiments, it is observed that TCN-based models give the best forecasting results among the other considered Deep Learning-based models with AMSE of 0.0025 and AMAE of 0.0230. The extracted geo-temporal information from AgrMine is beneficial for policymakers and agri-industrialists to design plans and agriculture-based products, respectively. Besides, the obtained patterns are also crucial for marketing specialists, extension workers, and scholars to develop optimized marketing policies, farmers' training, and agriculture-related analysis. In future work, the authors intend to incorporate Natural Language Processing in the analysis to perform text-based mining and hybrid approach-based modeling to design forecasting models based on the available dataset.

References

[1] S. C. Zipper, Agricultural research using social media data, Agronomy Journal 110 (1) (2018) 349–358. doi:http://dx.doi.org/10.2134/agronj2017.08.0495.

[2] S. Wolfert, L. Ge, C. Verdouw, M.-J. Bogaardt, Big data in smart farming–a review, Agricultural Systems 153 (2017) 69–80. doi:https://doi.org/10.1016/j.agsy.2017.01.023.

[3] A. Kamilaris, A. Kartakoullis, F. X. Prenafeta-Boldú, A review on the practice of big data analysis in agriculture, Computers and Electronics in Agriculture 143 (2017) 23–37. doi:https://doi.org/10.1016/j.compag.2017.09.037.

[4] J. Muangprathub, N. Boonnam, S. Kajornkasirat, N. Lekbangpong, A. Wanichsombat, P. Nillaor, Iot and agriculture data analysis for smart farm, Computers and electronics in agriculture 156 (2019) 467–474. doi:https://doi.org/10.1016/j.compag.2018.12.011.

[5] C. Brewster, I. Roussaki, N. Kalatzis, K. Doolin, K. Ellis, Iot in agriculture: Designing a europe-wide large-scale pilot, IEEE communications magazine 55 (9) (2017) 26–33. doi:https://doi.org/10.1109/MCOM.2017.1600528.

[6] MoSPI, Agricultural Surveys, Government of India (2019 (accessed December 20, 2021)).
URL http://mospi.nic.in/agriculture-surveys

[7] MoSPI, National Sample Survey Office (NSSO) (2019 (accessed December 20, 2021)).
URL http://mospi.nic.in/NSSOa

[8] P. Rosenfeld, S. Booth-Kewley, J. E. Edwards, Computer-administered surveys in organizational settings: Alternatives, advantages, and applications, American Behavioral Scientist 36 (4) (1993) 485–511. doi:https://doi.org/10.1177\%2F0002764293036004006.

[9] S. Godara, D. Toshniwal, Sequential pattern mining combined multi-criteria decision-making for farmers' queries characterization, Computers and Electronics in Agriculture 173 (2020) 105448. doi:https://doi.org/10.1016/j.compag.2020.105448.

[10] B. Chavan, R. Garg, R. Bhargava, Role of 24 hour telephonic helpline in delivery of mental health services, Indian journal of medical sciences 66 (5/6) (2012) 116. doi:http://dx.doi.org/10.4103/0019-5359.114196.

[11] R. G. Fukkink, S. Bruns, R. Ligtvoet, Voices of children from around the globe; an international analysis of children's issues at child helplines, Children & Society 30 (6) (2016) 510–519. doi:https://doi.org/10.1111/chso.12150.

[12] T. Shafiei, L. J. Biggs, R. Small, H. L. McLachlan, D. A. Forster, Characteristics of women calling the panda perinatal anxiety & depression australia national helpline: a cross-sectional study, Archives of women's mental health 21 (6) (2018) 801–812. doi:https://doi.org/10.1007/s00737-018-0868-4.

[13] C.-H. Chan, H.-K. Wong, P. S.-F. Yip, Exploring the use of telephone helpline pertaining to older adult suicide prevention: A hong kong experience, Journal of affective disorders 236 (2018) 75–79. doi:https://doi.org/10.1016/j.jad.2018.04.093.

[14] V. K. Viswanath, C. G. V. Madhuri, C. Raviteja, S. Saravanan, M. Venugopalan, Hadoop and natural language processing based analysis on kisan call center (kcc) data, in: 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 2018, pp. 1142–1151. doi:https://doi.org/10.1109/ICACCI.2018.8554531.

[15] K. K. Verma, N. Shrivastava, A novel tdmkcc model for decision making in e-agriculture., International Journal of Advanced Research in Computer Science 8 (5) (2017). doi:http://dx.doi.org/10.26483/ijarcs.v8i5.3909.

[16] S. Godara, D. Toshniwal, Deep learning-based query-count forecasting using farmers' helpline data, Computers and Electronics in Agriculture 196 (2022) 106875. doi:https://doi.org/10.1016/j.compag.2022.106875.

[17] GOI, Union Budget, Ministry of Finance, Government of India (2021 (accessed March 10, 2022)).
URL https://www.indiabudget.gov.in/

[18] GOI, Annual Report 2020-21, Department of Agriculture, Cooperation and Farmers' welfare, Government of India (2021 (accessed March 10, 2022)).
URL https://agricoop.nic.in/sites/default/files/Web%20copy%20of%20AR%20%28Eng%29_7.pdf

[19] GOI, Kisan Call Center, Ministry of Agriculture, Government of India (2004 (accessed December 1, 2021)).
URL https://dackkms.gov.in/account/login.aspx

[20] N. I. Centre, Open government data platform india, https://data.gov.in/ (2021).

[21] W. H. Press, S. A. Teukolsky, Savitzky-golay smoothing filters, Computers in Physics 4 (6) (1990) 669–672. doi:https://doi.org/10.1063/1.4822961.

[22] L. Deng, D. Yu, et al., Deep learning: methods and applications, Foundations and trends® in signal processing 7 (3–4) (2014) 197–387.

[23] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural computation 9 (8) (1997) 1735–1780.

[24] S. Bai, J. Z. Kolter, V. Koltun, An empirical evaluation of generic convolutional and recurrent networks for sequence modeling, arXiv preprint arXiv:1803.01271 (2018).

[25] J. Herzen, F. Lässig, S. G. Piazzetta, T. Neuer, L. Tafti, G. Raille, T. V. Pottelbergh, M. Pasieka, A. Skrodzki, N. Huguenin, M. Dumonal, J. Kościsz, D. Bader, F. Gusset, M. Benheddi, C. Williamson, M. Kosinski, M. Petrik, G. Grosch, Fast fourier transform forecasting model (2019).
URL https://unit8co.github.io/darts/generated_api/darts.models.forecasting.fft.html

[26] B. N. Oreshkin, D. Carpov, N. Chapados, Y. Bengio, N-beats: Neural basis expansion analysis for interpretable time series forecasting, arXiv preprint arXiv:1905.10437 (2019). doi:https://doi.org/10.48550/arXiv.1905.10437.

[27] J. Herzen, F. Lässig, S. G. Piazzetta, T. Neuer, L. Tafti, G. Raille, T. V. Pottelbergh, M. Pasieka, A. Skrodzki, N. Huguenin, M. Dumonal, J. Kościsz, D. Bader, F. Gusset, M. Benheddi, C. Williamson, M. Kosinski, M. Petrik, G. Grosch, Recurrent Neural Networks forecasting model (2019).
URL https://unit8co.github.io/darts/generated_api/darts.models.forecasting.rnn_model.html

[28] S. Bai, J. Z. Kolter, V. Koltun, An empirical evaluation of generic convolutional and recurrent networks for sequence modeling, arXiv preprint arXiv:1803.01271 (2018). doi:https://doi.org/10.48550/arXiv.1803.01271.

[29] G. Bontempi, S. Ben Taieb, Y.-A. L. Borgne, Machine learning strategies for time series forecasting, in: European business intelligence summer school, Springer, 2012, pp. 62–77. doi:https://doi.org/10.1007/978-3-642-36318-4_3.

[30] D. Birant, A. Kut, St-dbscan: An algorithm for clustering spatial–temporal data, Data & knowledge engineering 60 (1) (2007) 208–221. doi:https://doi.org/10.1016/j.datak.2006.01.013.

[31] R. Jain, P. Kishore, D. K. Singh, Irrigation in india: Status, challenges and options (2019). doi:http://dx.doi.org/10.5958/2455-7145.2019.00050.X.

[32] P. Cheung, The nutritional and health benefits of mushrooms, Nutrition Bulletin 35 (4) (2010) 292–299. doi:https://doi.org/10.1111/j.1467-3010.2010.01859.x.

[33] M. R. M. Usman, S. Barhate, M. Abullais, M. Usman, A review on drumstick tree (moringa pterygosperma gaertn): Multiuse tree with higher economical values, Int J Curr Pharm Rev Res 3 (1) (2012) 15–22.
URL https://www.researchgate.net/publication/274893933_A_Review_on_Drumstick_Tree_Moringa_pterygosperma_Gaertn_Multiuse_Tree_with_Higher_Economical_Values

[34] J. Raman, S.-K. Lee, J.-H. Im, M.-J. Oh, Y.-L. Oh, K.-Y. Jang, Current prospects of mushroom production and industrial growth in india, Journal of Mushroom 16 (4) (2018) 239–249.

[35] A. Kar, P. Kumar, P. Prakash, D. Singh, J. Dabas, S. Gowda, N. Deepak, V. Pooniya, R. Bana, et al., Doubling farmers' income in delhi state: An introspection, The Indian Journal of Agricultural Sciences 90 (8) (2020).

[36] B. Kaur, J. Singh, B. Garg, J. Singh, S. Singh, Causes and impact of labour migration: A case study of punjab agriculture, Agricultural Economics Research Review 24 (347-2016-16993) (2011) 459–466. doi:http://dx.doi.org/10.22004/ag.econ.119397.

[37] R. Saini, R. Kumar, Credit for investment: Its magnitude and sources in punjab agriculture, Indian Journal of Economics and Development 14 (1a) (2018) 360–365. doi:http://dx.doi.org/10.5958/2322-0430.2018.00082.3.

[38] R. Prasad, et al., Textbook of field crops production., Textbook of field crops production. (2002).

[39] N. Kumar, R. Chhokar, R. Meena, A. Kharub, S. Gill, S. Tripathi, O. Gupta, S. Mangrauthia, R. Sundaram, C. Sawant, et al., Challenges and opportunities in productivity and sustainability of rice cultivation system: a critical review in indian perspective, Cereal Research Communications (2021) 1–29doi:https://doi.org/10.1007/s42976-021-00214-5.

[40] S. Mahantesh, O. Ahlawat, K. Manikandan, et al., Mushroom consumption and purchasing behaviour in india: a study among selected respondents., Mushroom Research 23 (2) (2014) 225–231.

[41] V. P. Sharma, S. K. Annepu, Y. Gautam, M. Singh, S. Kamal, Status of mushroom production in india, Mushroom research 26 (2) (2017) 111–120.

[42] S. Dey, A. S. Noel, P. Tripathi, Study on marketing of mushroom (button mushroom) in dehradun district of uttarakhand (2022).

[43] R. S. Bana, D. Singh, M. S. Nain, H. Kumar, V. Kumar, S. Sepat, Weed control and rice yield stability studies across diverse tillage and crop establishment systems under on-farm environments, Soil and Tillage Research 204 (2020) 104729. doi:https://doi.org/10.1016/j.still.2020.104729.

[44] K. Krishnaiah, N. Varma, Changing insect pest scenario in the rice ecosystem—a national perspective, Rice Knowl. Manag. Portal (2011) 1–28.
URL http://www.riceportal.in/sites/default/files/ris/research-themes/Changing%20Insect%20Pest%20Scenario%20in%20the%20Rice%20Ecosystem_A%20National%20Perspective.pdf

[45] R. Singh, U. Singh, G. Khush, Basmati rice of india, Aromatic rices 135 (2000).
URL http://books.irri.org/8120414209_content.pdf

[46] S. Raina, H. Chahal, S. Kher, Analysing agriculture extension services for media mixes for transfer of technology, Journal of Rural Development 35 (3) (2016) 465–481.
URL http://nirdprojms.nirdpr.in/index.php/jrd/article/view/104697/0

[47] A. Gulati, P. Sharma, A. Samantara, P. Terway, Agriculture extension system in india: Review of current status, trends and the way forward (2018).
URL http://hdl.handle.net/11540/8400

[48] J. Yan, L. Mu, L. Wang, R. Ranjan, A. Y. Zomaya, Temporal convolutional networks for the advance prediction of enso, Scientific reports 10 (1) (2020) 1–15.