

Received December 24, 2019, accepted December 29, 2019, date of publication January 1, 2020, date of current version January 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2963590

# Integration of Google Play Content and Frost Prediction Using CNN: Scalable IoT Framework for Big Data

RANA M. AMIR LATIF<sup>1</sup>, SAMIR BRAHIM BELHAOUARI<sup>2</sup>, SAQIB SAEED<sup>3</sup>,  
LAIQA BINTE IMRAN<sup>1</sup>, MAZHAR SADIQ<sup>1</sup>, AND MUHAMMAD FARHAN<sup>1</sup>

<sup>1</sup>Department of Computer Science, COMSATS University Islamabad–Sahiwal, Sahiwal 57000, Pakistan

<sup>2</sup>Division of Information and Computing Technology, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar

<sup>3</sup>Department of Computer Information Systems, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, Dammam 34212, Saudi Arabia

Corresponding author: Samir Brahim Belhaouari (sbelhaouari@hbku.edu.qa)

**ABSTRACT** The forecast of frost occurrence requires complex decision analysis that uses conditional probabilities. Due to frost events, the production of crops and flowers gets reduced, and we must predict this event to minimize the damages. If the frost prediction results are accurate, then the damage caused by frost can be reduced. In this paper, an ensemble learning approach is used to detect frost events with Convolutional Neural Network (CNN). We have used this to get more efficient and accurate results. Frost events need to be predicted earlier so that the farmer can take on-time precautionary measures. So, for measurement and analysis of Google Play, we have scrapped a dataset of the Agricultural category from different genres and collected the top 550 application of each category of Agricultural applications with 70 attributes for each category. The prediction of frost events prior few days of an actual frost event with an accuracy of 98.86%.

**INDEX TERMS** CNN, frost event, ensemble, prediction, Google Play store, scraping, measurement, and analysis.

## I. INTRODUCTION

Frost's event occurs due to changes in environmental and biological processes. When there is an accurate prediction of temperature with the warning of the critical temperature, it provides significant benefits to predict frost events, and hence we can be able to protect our crops on time. Sometimes the damage is not visible at the start, but when the farmer harvests the crop, the quality and quantity of crop reduce. Frost affects the leaves, taste, and production of the crop and even the flowers which remain small. The frost has damaged the beauty and crop production in countries like Argentina, Australia, Georgia, Mexico, China, South Korea, Iran, USA, Italy, Germany, Berlin, and many others [1]. Temperature is one of the primary factors affecting the plant's growth and geographical distribution. When the atmosphere decreases the level of temperature, it causes damage to plant species and affects the growth rate of a plant [2]. Low temperature like chilling and freezing can occur in all plants, but the damage

is different in every aspect. Many crops experience physiological damage when the temperature is below 12.5 degrees, but the damage above 0 degrees is an injury on the crop. An injury occurs in all plants when the temperature is low, and there is an ice formation in the plant. Frost damage has a drastic effect on an entire plant or, sometimes, only a specific organ [3], [4].

Agricultural capability for food production in each region depends on weather and climate conditions. The protection of plants from the effects of lethally low temperatures is essential in agriculture, especially for the horticultural production of high-value fruits and vegetables [5]. Plant physiology, seed germination, pollination, growth, photosynthesis, and material transport within the plant are sensitive to climate conditions in growth stages. One of the meteorological issues for many agricultural systems is frost occurrence [6]. Frost also causes a significant impact on electrical lines, engine turbine blades, heat exchangers, and wind turbines. To minimize damaging effects, we must act in the operational stage of agriculture and the planning of agriculture. These issues let us know that there is a considerable need for understanding and accurately predicting the formation of frost so that

The associate editor coordinating the review of this manuscript and approving it for publication was Sohail Jabbar<sup>1</sup>.

actions must be taken on time [7]. The formation of frost has three stages that are: frost layer with full growth interval, frost layer having simple growth interval, and crystal growth interval. This type of problem is enormous in areas where the meteorological data is not available. In this type of case, there is a small amount of data to build a predictive model with high accuracy. Due to heat shocks, there were physiological impacts on wheat production that has to be modeled to reduce its loss [8], [9]. The prediction of frost events needs some complex decision analysis that uses probabilities and economics. For the starting of the active frost prediction model, it is essential to predict the time of temperature fall [10].

In Urban areas, the temperature is slightly lower than the rural areas as there are more fields, grounds, and open environment. Whereas in rural areas, there is a higher temperature because of the environment as having a lot of home site to site and vehicles. Due to this, we are having more problems with frost damage in urban areas [11], [12]. There is also an important thing that our crops grow in urban areas.

There are several applications available on the Google play store related to the agricultural domain. In different genres, we have several applications that support agricultural for different purposes. So, in this research work, we analyze and assess the applications, which can facilitate our users related to frost events data collection, suitable for machine intelligence and analytics [13].

A third-party application attached to the google play store is very significant, which allows users to take benefit from it.

The thousands of users use these applications on their mobiles. Several applications are uploaded with the thousands of programmers daily. The validation is the fact that for assessing and detecting applications [14]. The third-party programmers have access to the original code. Google has not any in-person access to these applications because the programmer uploads these applications directly on the Google platform and does not give any access to the Google authorities. On these applications, Google places many rules and procedures that should be followed by programmers [15].

## II. LITERATURE REVIEW

The authors have proposed a prediction of frost events using machine learning techniques and IoT devices in [16], [17]. The authors have used Bayesian networks and Logistic Regression that predict the minimum temperature for the next day. They have proposed a network of IoT sensors that collect the temperature data and send the data further to provide frost forecast output. Winter wheat in China was affected by a spring frost, and they have faced much damage. The risk of frost is smaller in the northern area of China and the warmer southern region. When the temperature rises, it can reduce the chance of frost event occurrence. The authors have predicted the frost event, and for prediction, they have used the Regression method in [18]. In [19], the author has discussed the effect of frost on wheat and listed some models as well as they have observed field conditions that are damaging the crop. They have estimated the effect of frost on different crops.

The artificial neural networks model was employed to predict the behavior of temperature, and this was implemented in a greenhouse to control frost damage [5]. The author has used the data of 30 years, which is not the right approach. Prediction of frost events can also be made using statistical modeling. In South Korea, frost warning systems have been deployed using logistic regression and decision trees [20]. In [21], the author had said that in 2013, in Argentina, the frost events have destroyed a large production of peach orchards. Temperature data is collected using a low power wireless mesh network, and then the prediction is performed by using a regression tree [22]. They have predicted frost events in peach orchards.

In 2017, the regression kriging method had been used to predict, spatial analysis of frost risk to determine viticulture suitability [23]. In this approach, the authors have done the average weather data. In Germany, due to climate change and spring frost, the production of sweet cherries has been damaged, and to predict spring frost, they have taken daily data of air temperature and calculated the results by Phenological models [24]. In [25], the author wrote a paper for the assessment of spring frost effect on beech, and remote sensing is used to collect the data on which they have applied support vector machine (SVM) to predict spring frost. In [26], the author has experimentally investigated the sensitivity of frost on vegetables and reproductive series of herbs by using the logistic regression model that shows the relationship between leaf frost damage and the effect of temperature on a leaf.

In 2017, M.H. Jorenoosh said that whenever there are low-temperature many fruits, vegetables, and crops of the tropical region are experiencing physiological damage [27]. Protection of plants from the effects of lethally low temperatures is essential in agriculture, especially in the horticultural production of high-value fruits and vegetables. They have done a statistical analysis to predict the frost event in [28]. Robert, in [22], proposed a web-based fuzzy logic system. This system was developed to predict air and to observe wind speed in order to predict frost. The warning levels have been generated so that they can save the production of crops.

In [29], the author has applied a machine learning algorithm for frost prediction. He has built a system to predict the air temperature and soil temperature in order to predict frost events in north Italy. The author has provided 24-hour weather prediction in [30]. In this system, they have used the Multi-Layered Perceptron Network (MPLN). Simulations are done numerically, which are used to predict the weather. In [31], the author has shown the models of weather forecasting to study frost events in the US. They have used the advanced research model for soil data and specific topography map resolution [32].

Prediction of weather temperature in a crop field is a crucial issue, whether in the urban area or rural area. In [33], they have used the ANN model, but with another approach, their accuracy is good, but some confusions also need to be resolved in terms of data. They have mainly sensors to get

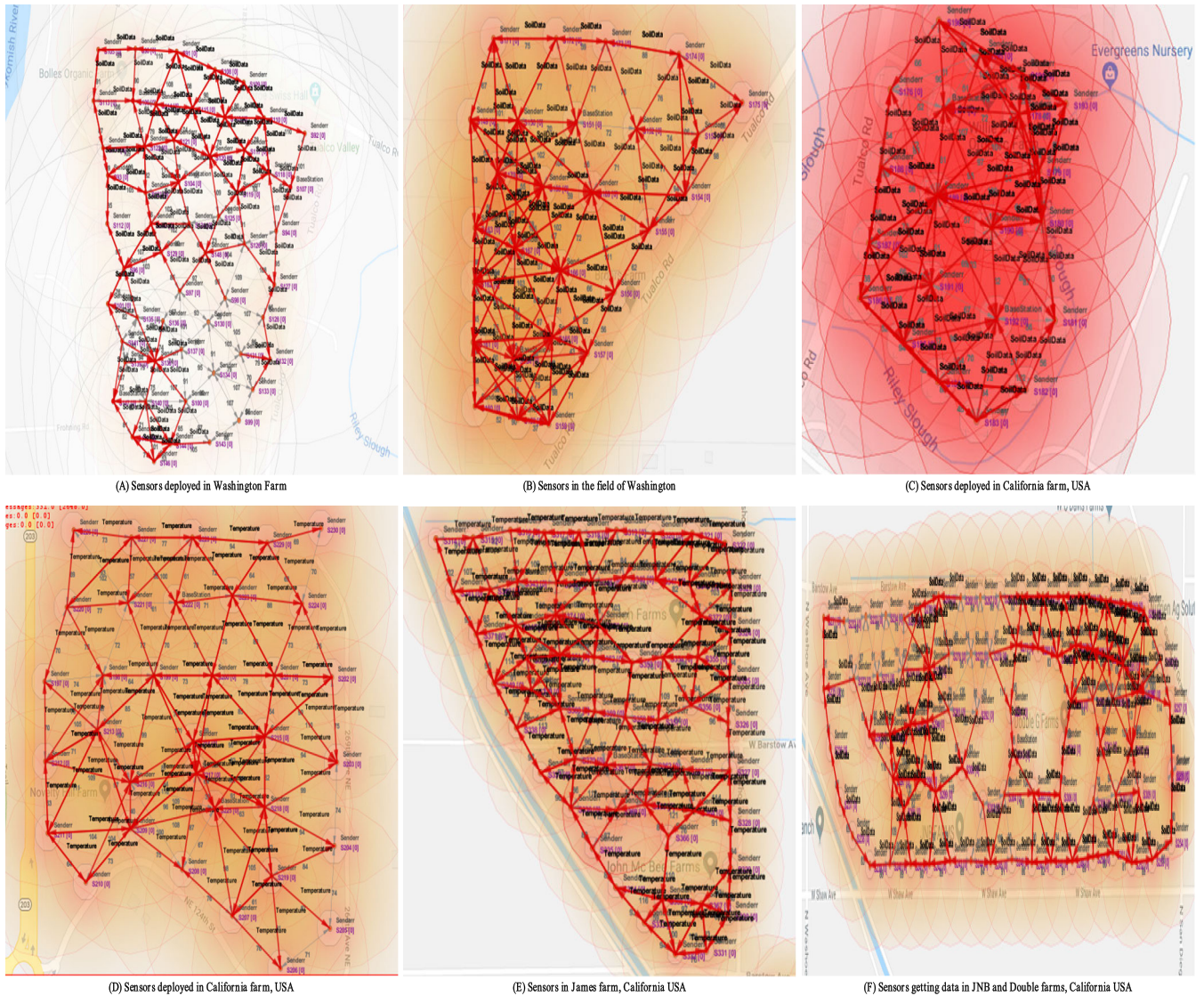


FIGURE 1. Differece sensors deployed in the fields of the USA, as shown in (A), (B), (C), (D), (E), and (F).

weather data from weather stations. The author has calculated the frequency and strength of frost in [27].

In this paper, authors have developed an Android-based application for the management and the production of organic manures (MoAPOM). MoAPOM application is primarily for the small and medium-scale farmers who want to improve the crop production MoAPOM has been designed on the traditional production method of organic manure using a human to human communication was converted into the machine to human communication and the intelligent computer-aided system [34], [35]. Five types of modeled and organic manure production, which is earthworm compost, green manure, poultry waste, cattle dung, and kitchen waste. After the validation of the MoAPOM application for offline status, which uses for sharing portability and convenient GUI, ease of download, timeliness, many tutorials, and offline usage [36].

Internet of Things (IoT) devices produce the bulk of data, and it leads to cloud storage. This data is used to analyze the stock requirements for the crop, market, analysis of the crops, and fertilizer requirement [37], [38]. Then based on different data mining techniques, the prediction can be performed and reaches towards the farmer via a mobile application. The purpose of this application is to increase the crop production and agricultural cost that is more useful to the farmers [39].

### III. PROBLEM STATEMENT

Frost’s event is the condition that exists when the temperature of the earth’s surface decreases at night, and the temperature level comes to zero degrees centigrade. These events damage the whole production of the crop and flowers due to which the farmer and the whole country faces economic loss. Frost’s events are challenging to predict, and previous methods are not so much accurate. Through the literature review, we came

to know that previous methods have a very noisy structure of the data. The people had mostly predicted the frost one day earlier or some hours earlier before the frost, but the frost needs to be predicted earlier like a few days before the frost events so that the farmer can take precautionary measures on time and the use of in-depth learning approach provides the accuracy and earlier prediction to make well-founded decisions.

#### IV. MATERIALS AND METHODS

We have deployed our Wireless Sensor Network (WSN) devices in many fields. We have collected the soil data and temperature data as well as we are predicting the surface temperature to predict frost events.

##### A. DATA COLLECTION

When we take the dataset for the prediction, the most crucial step is to preprocess the data. The initial data has much noise, and it is necessary to remove noise from the data so that our results can be more efficient. Sometimes the dataset has many parameters, and we do not want them, so we should neglect those and train our model on a dataset that is related to our problem. We are getting the data through WSN devices, and sensors have been deployed in the fields of the USA virtually. Those are getting the data and transferring the data to the station, and then different techniques are applied to get the prediction. The farmer has deployed two sensors in the field that is the soil sensor and temperature sensor. These sensors are transferring the data to the base station, and the base station is sending the data to the central station, as shown in Figure 1.

In this research, the working is on two values, like above zero degrees and below zero degrees. When the temperature is above zero degrees, then there will be no frost, but when the temperature is below zero degrees, then there is a possibility of the frost. Labels are essential for the supervised model, but the dataset has no labels, in this case, we must define the labels. The whole dataset has two parts that are 80% for the training data and 20% for the testing data. Input data creates vectors at the time of training. The input form for the model can be defined, as shown in Equation 1.

$$I = (I^{(1)}, I^{(2)}, \dots, I^{(m)}) \quad (1)$$

where “m” is the number of input samples. In each sample, a year, month, date, and soil temperature values are incorporated. The output form is also like the input one. In this O, that is representing the output, as shown in Equation 2.

$$O = (O^{(1)}, O^{(2)}, \dots, O^{(m)}) \quad (2)$$

##### B. DEVELOPMENT OF THE FROST DAMAGE PREDICTION MODEL

Convolutional Neural Network (CNN) is used to predict, and the following sections have the details.

##### 1) CONVOLUTIONAL FUNCTION

In Python using TensorFlow, a function is used “conv2d” to convolve in the CNN model, but it accepts the 2D image as input. As the data is of one-dimensional time series, so it must be changed to “conv1d”. While adopting this function, it covers many properties like some filter, width, stride as input. The kernel size can be 1 or 2 depending upon the requirements. The Gaussian distribution is used to initialize the value of filters and biases. The output of this function is in the form of matrices, and in the next layer as input is used. The next layer is the pooling layer, which alters the computation of the activation function.

##### 2) CONVOLUTIONAL NEURAL NETWORK ARCHITECTURE

As we are using a conv1d function, so the CNN model is built to make the predictions. First, the data is input into the frost prediction model, and it uses conv1d function for the convolution. In the dataset, there are 512 hidden layers and ten neurons in the output layer. The data is in linear form, and it does not accept the non-linear data, so it is necessary to introduce the activation function. There are different activation functions like LReLU, Sigmoid, ReLU, softmax, and Tanh. In this model, we are using ReLU and LReLU function for the computation of the activation function. Equation 3 and Equation 4 define these functions.

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad (3)$$

$$f(x_i) = \begin{cases} a_i x_i & \text{for } x < 0 \\ x_i & \text{for } x \geq 0 \end{cases} \quad (4)$$

Equation 3 represents ReLU function, and equation 4 represent the LReLU function. In this equation, both  $x$  and  $x_i$  are the results of convolution computation in the convolutional layer. There is a hyperparameter represented as  $a_i$ , and its value is minimal while training this parameter can also be fine-tuned to get the best result. The ReLU activation function because it is the most useful activation function, its computation is secure, and it takes less time to train or to run the data. Hence by looking at the properties of the ReLU function, the first convolutional layer is used, and then it is shifted towards LReLU function in the next layer. To increase the computation power and to reduce the feature dimension, the feature of the convolutional layer in the pooling layer will use. In this model, we have used the MaxPooling1D function as the data is one dimension, and the pool size that it is used to  $1 \times 1$  that will subsample its input with maximum function. In the layer’s system, the output of one layer becomes the input of the next layer by using the convolutional and pooling layer. The output of the pooling layer is the input for the fully connected layer. In the fully connected layer, we have used the ReLU function. We have used a dropout strategy to avoid over-fitting in the fully connected layer with the value to 0.2. The accuracy cab is improving with this model because the error rate will become less, and Equation 5 defines how we

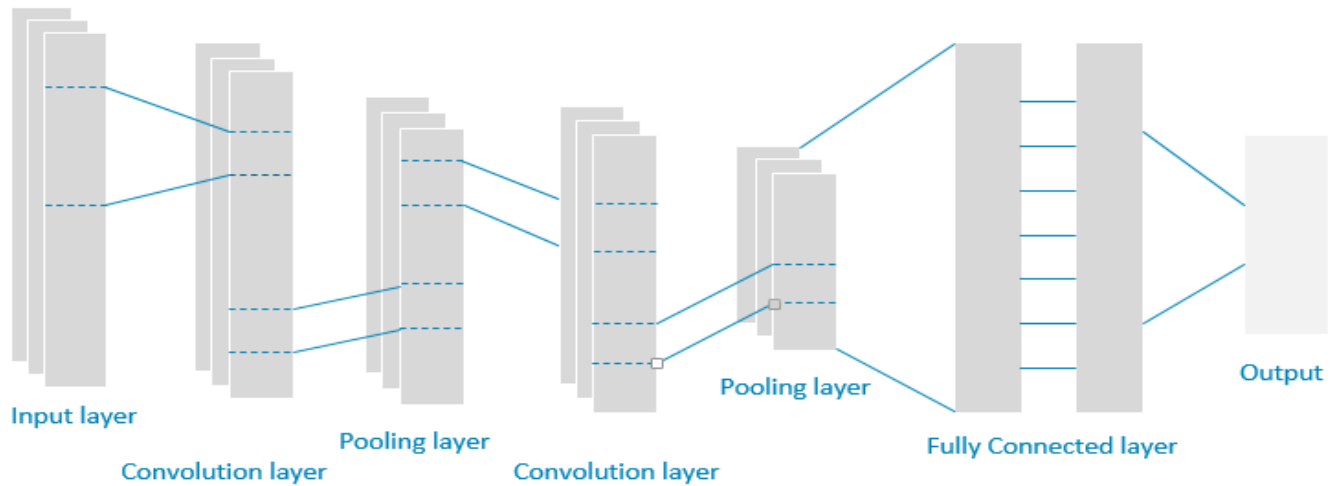


FIGURE 2. The architecture of the CNN model used for frost detection.

have calculated the error.

$$E = \sum_{x=1}^i |y^{(x)} - p^{(x)}| \tag{5}$$

In this equation  $y^{(x)}$  is the actual value,  $i$  is the number of samples and  $p^{(x)}$  is the predicted value, we have taken the difference between the actual value and the predicted value. When we must minimize the loss, there are many ways like backpropagation, but in this, we have used Adam’s approach for the computation of optimization as defined in Equation 6.

$$V' = V - \eta \cdot \frac{\partial J(V)}{\partial V} \tag{6}$$

In equation 6  $V'$  is the updated value which we have gained after computation and  $\eta$  is the learning rate. Figure 2 shows the architecture of the CNN model. In this figure, we have convolutional layers, and we have used a small number of convolutional layers to make the model efficient and fast computation. After that, in the center, we have a pooling layer, and the pooling layer output is the input of a fully connected layer. We have used two fully connected layers, and the size of our kernel is  $2 \times 2$  with 32, 64, 96 filters. The size of the pooling layer is also 2. The dataset contains many samples in which 80% is the training data, and 20% is the test data. In the end, after a fully connected layer, we are getting our output in the form of graphs, matrix with accuracy.

### 3) ENSEMBLE LEARNING ARCHITECTURE

Ensemble learning has two types, i.e., bagging and boosting. We are using bagging, which is a robust ensemble learning algorithm that combines the strengths of multiple models. The bootstrap sampling method is employed further to create the subsets of observations from the original dataset. The size of the data is the same as the original, and it is used to gain distribution. The size of the subset is smaller than

the original dataset. We have created a base model on each of these subsets, and then the model has run in parallel, and they are independent of each other. In the end, we have gotten the final prediction by combining the predictions from all the models, and we have created three CNN models. In the first model, we have one layer, in the second layer, we have two layers, and in the third, we have three convolutional layers with three pooling layers. In bagging, we have two algorithms that are bagging the meta estimator and Random Forest, both provide the result with less error rate.

Overall the process of CNN with the ensemble approach has several steps. We have performed for the forecasting of the frost event, and Figure 3 shows these steps. At first, preprocessing has been performed on data, in which it divides and transforms into a test sample and training sample. In the next step, the training is performed on the data with three models and getting the predictions of the frost event. Then ensemble learning is applied with the bagging technique, and after training the data, we are getting the predictions. In aggregation, it extracts the result by taking the mean of  $k$  predictions.

## V. RESULTS AND DISCUSSIONS

In Figure 4,  $Y_{test}$  is the test data, and  $X_{test}$  is the test data in the model. We have used the dataset from the different stations in the USA.  $X$  and  $Y$  are two different stations from where the data is collected for testing purposes. The prediction of frost shows in terms of a graph, where the blue line is indicating the train data, whereas the orange line is describing the test data. When the temperature goes below zero degrees, frost event takes place, but when the temperature is above zero degrees, then there is no frost. Figure 5 is showing the error rate of the model, and we have calculated this by taking the difference between actual data and predicted data.

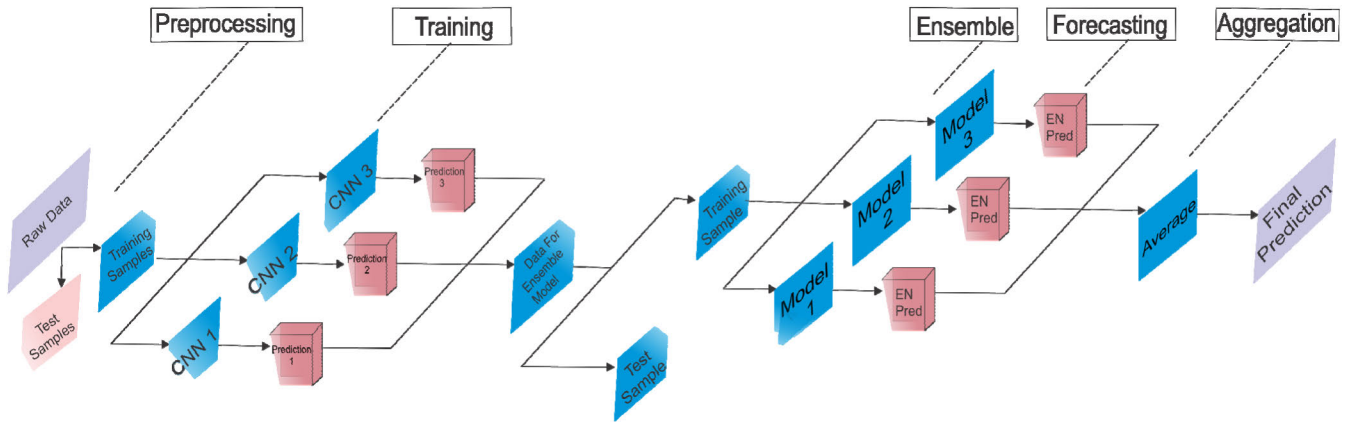


FIGURE 3. CNN architecture using ensemble learning approach.

TABLE 1. The result of the model using different data sources.

Data Source	1	2	3	4
Iterations = 20000				
Accuracy	96.67%	97.8%	96.2%	95.8%
Iterations = 50000				
Accuracy	97.3%	98.86%	98.2%	98.6%

TABLE 2. Prediction accuracy of frost events among different countries and its crop-wise damage.

Countries	Damage of frost event	Prediction accuracy of frost
Australia	Fruits and vineyards	78%
Mexico	Crops and Plants	95%
South Korea	Crops	78 to 80%
Argentina	Peach Orchards	85%
China	Wheat	86%
Iran	Fruits, vegetables and Ornamental Crops	80%
USA	Leaves, Flowers of subalpine plants, Fruits	93%
Italy	Beech forest	87%
Germany	Flowers, Plants, Fruits (Cherries)	90 to 93%

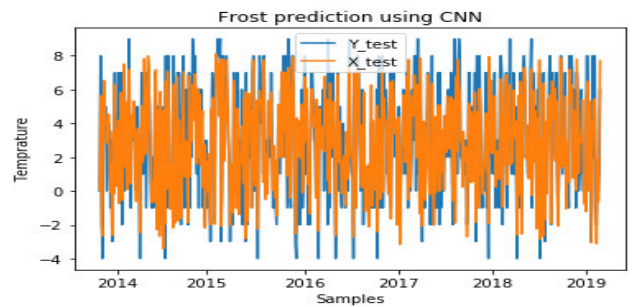


FIGURE 4. Training and prediction of frost events using CNN for five years.

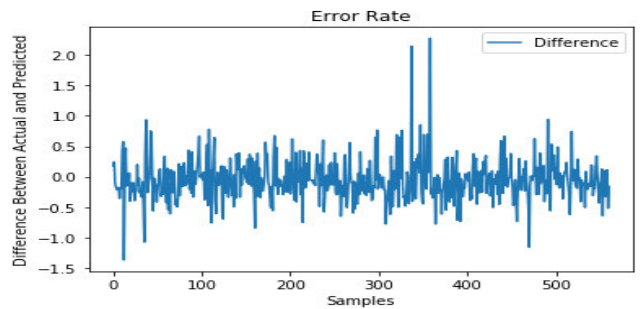


FIGURE 5. The error rate of the CNN model using five years of data.

We have predicted the frost event by using one-year data from Feb 2018 to Jan 2019, as shown in Figure 6. The blue line is training data, whereas the orange line is representing the test data. Figure 7 is showing the error rate of the model on a one-year dataset.

We have used different data sources, so the result varies according to their data and the number of iterations. We have concluded that the configuration of our model is more feasible and efficient, as shown in Table 1.

It is shown in Figure 8 that we have 98.86% of accuracy in predicting the frost events. It shows the linear regression line while it is not discussed in the citing paragraph. The behavior of the data is linear. In previous work, the accuracy was 95%. In 50000 iterations, it gives 98.86% accuracy. In 20000 to 30000 iterations, it gives 97.3% to 98.86% accuracy on different datasets.

Table 2 has shown the issue of frost event and its damage to plants, crops, fruits, and so on within many countries. Some researchers have achieved an accuracy of 93% to 95% in different countries, but we have achieved 98.86% accuracy in US virtual data.

After the prediction of the frost event, the farmers receive a notification message. It can be in the form of SMS and voice calls, as shown in Figure 9. For this service, we have used Nexmo's SMS API that enables us to send and receive text messages to and from users worldwide. It is a beneficial and high-quality service with an extremely reliable global network.

## VI. GOOGLE PLAY STORE USE CASE

Google play store is the primary hub for the application, where several developers upload millions of applications.

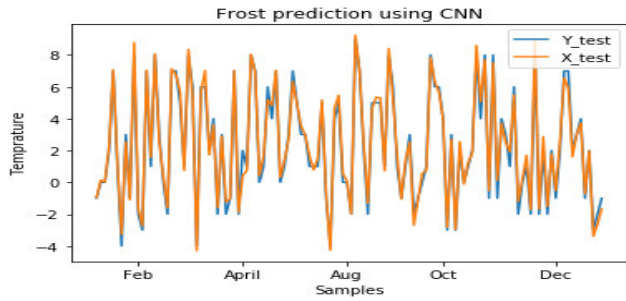


FIGURE 6. Training and prediction of frost event using one-year data.

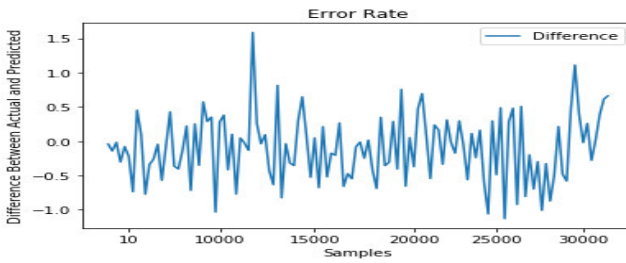


FIGURE 7. The error rate of a model on a one-year dataset.

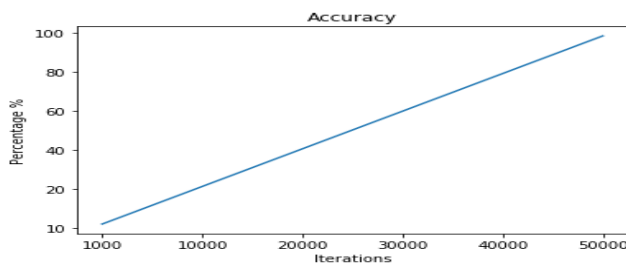


FIGURE 8. Percent accuracy vs. the number of iterations.

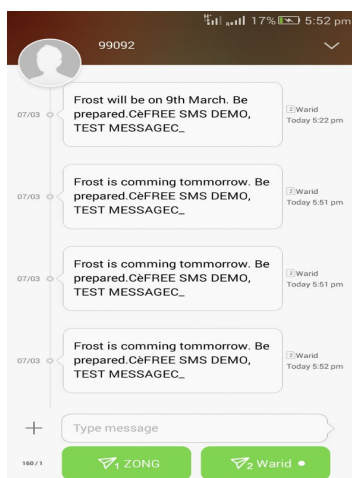


FIGURE 9. Sample SMS notifications on a farmer's cell phone.

Likewise, millions of users download these applications without checking the authenticity and duplicity of applications. It causes the user's loss due to damage to user trust in the Google play store. Aggregately, Google play store

1	title	installs	ratings	offersIAP	adSupported
2	Agriculture Student(BSC Agri Notes)AGRILEARNER	50,000+	595	FALSE	TRUE
3	Agricultural Business	100,000+	523	FALSE	TRUE
4	Agriculture Dictionary	100,000+	1126	FALSE	TRUE
5	agriculture for all	5,000+	17	FALSE	TRUE
6	Agriculture quiz	50,000+	208	FALSE	TRUE
7	Agriculture Engineering	1,000+	7	FALSE	TRUE
8	Agri Zone : All In one Agri App	10,000+	349	FALSE	TRUE
9	IFFCO Kisan- Agriculture App	500,000+	3159	FALSE	FALSE
10	Real Tractor Farming Simulator 2018	1,000,000+	4015	TRUE	TRUE
11	Agricultural Business App V3.0	5,000+	9	FALSE	TRUE
12	Agriculture Engineering 101	1,000+	29	TRUE	FALSE
13	Learn Agriculture by Pictures - 1000 Images	5,000+	101	FALSE	FALSE
14	Agronote - Farm Record	10,000+	137	TRUE	TRUE
15	Agriculture News Network	10,000+	79	FALSE	TRUE
16	Agri Book	50,000+	124	FALSE	TRUE
17	Agriculture quiz	1,000+	4	FALSE	TRUE
18	Farmersgrid - Agriculture & Farming	1,000+	35	FALSE	FALSE
19	Agriculture Students, Agri notes, Agri Books	1,000+	4	FALSE	TRUE
20	Grade 12 Agricultural Sciences Mobile Application	5,000+	44	FALSE	TRUE

FIGURE 10. Sample screenshot of agricultural applications dataset.

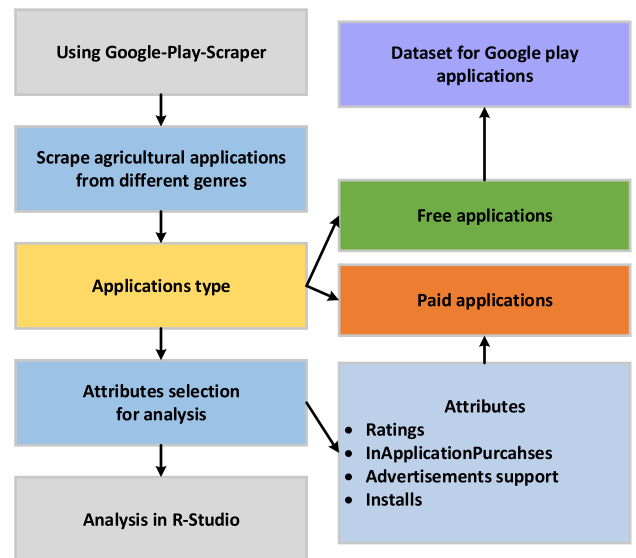


FIGURE 11. The methodology adopted for Google play store use case.

gives a small amount of information about such applications. We have made a Google play store dataset with a free and paid category of the agricultural applications using the Google-play-scraper. The scraper is used to collect the top 550 applications of every single type of agricultural application with a free and paid application category. Each application on the Google play store has 70 features, which are rankings, InAppPurchases, addSupported, and number of installs. Figure 10 shows the sample screenshot of the agricultural applications dataset. In this research results have been visualized the installs and ratings of the free and paid applications in the form of histogram, purchase rate of those free and paid applications also purchases rate of those applications which offers to get a percentage of the advertisement support with all categories of agricultural application as shown in Figure 11.

### A. INAPPLICATIONPURCHASES (IAP) OF FREE AGRICULTURAL APPS

The user of the application can get some credits, and products by performing any task like watching an advertisement

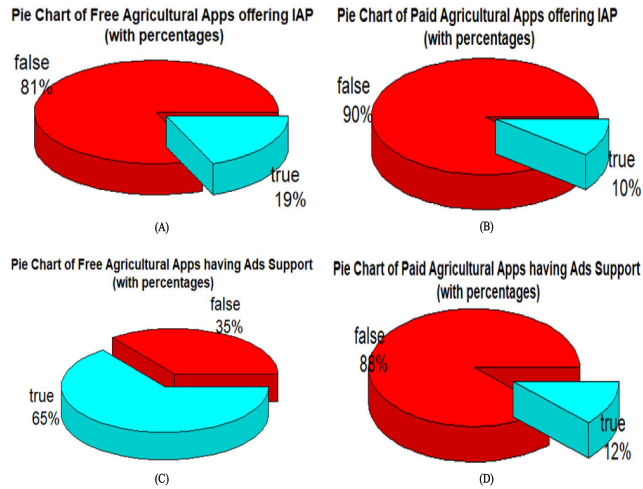


FIGURE 12. Different agricultural applications (A) Free IAP (B) Paid IAP (C) Free ads support (D) Paid ads to support.

video in-application purchase. The applications of the free category casually offer more IAP because, at the time of the install, there is no fee for the user of the free application. We have used the Pie3D charts for visualizing the percentage of defined attributes. 19% of the free applications of agriculture offer IAP, and 81% did not offer anything, as shown in Figure 12 (a).

**B. INAPPLICATIONPURCHASES (IAP) OF PAID AGRICULTURAL APPS**

The paid applications offer more usability to their users that is the way in the time of the installation of the paid applications, the specific amount of that application is charged from the user. We have analyzed that only 10% of paid agriculture applications to offer IAP, and 90% did not offer anything, as shown in Figure 12 (b).

**C. ADVERTISEMENTS FOR FREE AGRICULTURAL APPS**

The advertisement in the applications plays an essential role in the success of the application. In the Google play store, many free applications offer many frustrating advertisements. The user also gives helpful information sometimes. Most of the free applications offer many advertisements because, at the time of the installation, there is no fee charged from the user. In the evaluated results, we have analyzed that 65% of free agricultural apps offer advertisements, and 35% did not offer anything, as shown in Figure 12 (c).

**D. ADVERTISEMENTS FOR PAID AGRICULTURAL APPS**

The paid applications charge some specific amount from the user at the time of the installation; that is why they are offering not some irritating advertisements to the user end. In the evaluated results, we have analyzed that 12% paid agricultural apps to offer advertisements, and 88% did not offer anything, as shown in Figure 12 (d). In the paid agriculture applications, the advertisement ratio is lesser then free applications.

**E. RATING VALUE OF FREE AGRICULTURAL APPS**

In the casual routine, the user downloads the application for their personal use. If the application has a good impact on the use-ability factor, the user will give an excellent rating to that application. Otherwise, the user gives bad rrating. In this case, on the bases of the rrating of the application, we can judge the popularity of the application. The scale of the rrating of the application from the Google play store is between 0-5. On this parameter, we have analyzed that 0 rrating shows that the application use-ability is terrible, and if the application that has 5 rrating that shows the popularity of the application. In the evaluated results, we have visualized that most of the user gives 4.5 rrating to the free agriculture application, as shown in Figure 13. Here Y-axis represents the number of installs.

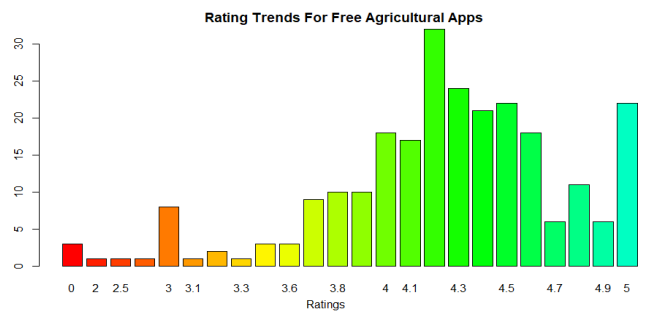


FIGURE 13. Free agricultural application rating trends.

**F. RATING VALUE OF PAID AGRICULTURAL APPS**

In the paid application, there is no more frustration for the user of the application, but the evaluated results were amazingly opposite. We have visualized that mostly peoples give 5 rrating to the paid application, and at the same time, many users give 0 rrating that has shown the unsatisfaction of the users, as shown in Figure 14.

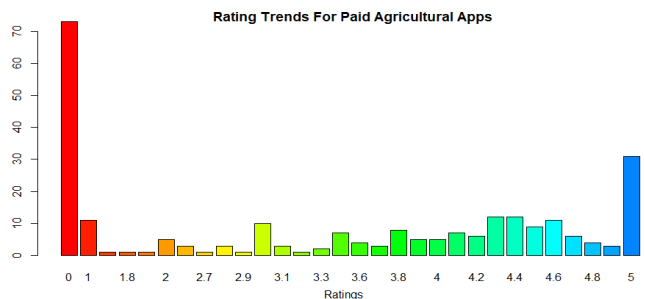


FIGURE 14. Paid agricultural application rating trends.

**G. FREE AND PAID AGRICULTURAL APPLICATIONS INSTALLS**

The number of installs plays an essential role in the application rating. In comparison, we have analyzed that free applications have a more significant number of installs, as shown



in Figure 15. The audience of the Android app focuses on free apps rather than paid apps. There can be many reasons for it, and one of the reasons is the availability of free apps.

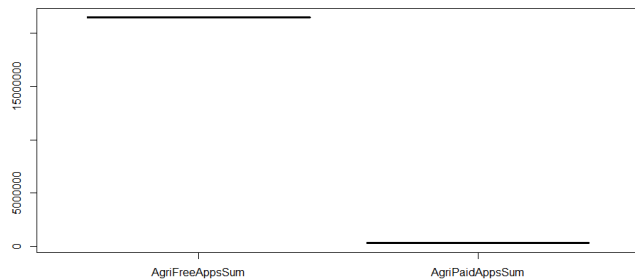


FIGURE 15. The number of installs for free and paid agricultural apps.

## VII. CONCLUSION AND FUTURE WORK

The forecast of frost occurrence requires complex decision analysis that uses conditional probabilities and economics. Due to frost events, the production of crops and flowers gets reduced, and we must predict this event in order to minimize the damages. In this paper, we have introduced a CNN model approach for frost event prediction, and a conv 1d function has been used to process the 1-dimensional data. We deployed our sensors on the field virtually, the soil data and temperature collected, and the frost events have been predicted using the CNN model. We have used three convolutional layers, which is a more efficient method. After that, we have applied the bagging technique using ensemble learning to make our predictions more accurate. We have used three models, in one model, there is one layer, in the second model, there are two layers, and in the third model, there are three layers. We have embedded these models into the ensemble model, taken the prediction by applying the bagging approach, and then test the model. We have scraped the Google play store dataset of agricultural applications. We have used four attributes for analysis like rankings, InAppPurchases, addSupported, and number of installs.

In the future, we will also combine this model with the RESNET model and others. We will create a new model by combining the two models, and this is expected to result in higher accuracy of prediction and computationally feasible.

## ACKNOWLEDGMENT

The authors would like to thank Qatar National Library, QNL, for supporting us in publishing our research.

## REFERENCES

- [1] Y. Huang, H. Yao, F. Zhao, and K. Reddy, "Detection of crop herbicide injury through plant hyperspectral remote sensing of chlorophyll fluorescence," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 5069–5072.
- [2] P. F. Verdes, P. M. Granitto, H. D. Navone, and H. A. Ceccatto, "Frost prediction with machine learning techniques," VI Congreso Argentino de Ciencias de la Computación, Tech. Rep., 2015.
- [3] G. Sgubin, D. Swingedouw, G. Dayon, I. G. de Cortázar-Atauri, N. Ollat, C. Pagé, and C. Van Leeuwen, "The risk of tardive frost damage in French vineyards in a changing climate," *Agric. Forest Meteorol.*, vols. 250–251, pp. 226–242, Mar. 2018.
- [4] T. Watteyne, A. L. Diedrichs, K. Brun-Laguna, J. E. Chaar, D. Dujovne, J. C. Taffernaberry, and G. Mercado, "PEACH: Predicting frost events in peach orchards using IoT technology," *EAI Endorsed Trans. Internet Things*, vol. 2, no. 5, Dec. 2016, Art. no. 151711.
- [5] L. Ghielmi and E. Eccel, "Descriptive models and artificial neural networks for spring frost prediction in an agricultural mountain area," *Comput. Electron. Agricult.*, vol. 54, no. 2, pp. 101–114, Dec. 2006.
- [6] P. Moller-Acuna, R. Ahumada-Garcia, and J. Reyes-Suarez, "Prediction of frost episodes based in agrometeorological information and machine learning techniques," in *Proc. IEEE Int. Conf. Automat. (ICA-ACCA)*, Oct. 2016, pp. 1–7.
- [7] S. M. Kotikot, A. Flores, R. E. Griffin, A. Sedah, J. Nyaga, R. Mugo, A. Limaye, and D. E. Irwin, "Mapping threats to agriculture in East Africa: Performance of MODIS derived LST for frost identification in Kenya's tea plantations," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 72, pp. 131–139, Oct. 2018.
- [8] I. Maqsood, M. R. Khan, and A. Abraham, "An ensemble of neural networks for weather forecasting," *Neural Comput. Appl.*, vol. 13, no. 2, pp. 112–122, 2017.
- [9] D. Martino and P. Abbate, "Frost damage on grain number in wheat at different spike developmental stages and its modelling," *Eur. J. Agronomy*, vol. 103, pp. 13–23, Feb. 2019.
- [10] C. Malmqvist, K. Wallertz, and U. Johansson, "Survival, early growth and impact of damage by late-spring frost and winter desiccation on douglas-fir seedlings in Southern Sweden," *New Forests*, vol. 49, no. 6, pp. 723–736, Nov. 2018.
- [11] H.-J. Gim, C.-H. Ho, J. Kim, and E. J. Lee, "Urbanization may reduce the risk of frost damage to spring flowers: A case study of two shrub species in South Korea," *PLoS ONE*, vol. 13, no. 2, Feb. 2018, Art. no. e0191428.
- [12] K. Rautiainen, T. Parkkinen, J. Lemmetyinen, M. Schwank, A. Wiesmann, J. Ikonen, C. Derksen, S. Davydov, A. Davydova, J. Boike, M. Langer, M. Drusch, and J. Pulliainen, "SMOS prototype algorithm for detecting autumn soil freezing," *Remote Sens. Environ.*, vol. 180, pp. 346–360, Jul. 2016.
- [13] A. Ahmad, A. Paul, and M. M. Rathore, "An efficient divide-and-conquer approach for big data analytics in machine-to-machine communication," *Neurocomputing*, vol. 174, pp. 439–453, Jan. 2016.
- [14] J. Businge, M. Openja, D. Kavalier, E. Bainomugisha, F. Khomh, and V. Filkov, "Studying Android app popularity by cross-linking GitHub and Google play store," in *Proc. IEEE 26th Int. Conf. Softw. Anal., Evol. Reeng. (SANER)*, Feb. 2019, pp. 287–297.
- [15] J. Lu, W. Hu, M. Song, X. Zhan, and X. Liu, "Mobile medical service system based on portable devices," in *Proc. IEEE 17th Int. Conf. High Perform. Comput. Commun., IEEE 7th Int. Symp. CyberSpace Saf. Secur., IEEE 12th Int. Conf. Embedded Softw. Syst.*, Aug. 2015, pp. 1530–1535.
- [16] A. L. Diedrichs, F. Bromberg, D. Dujovne, K. Brun-Laguna, and T. Watteyne, "Prediction of frost events using machine learning and IoT sensing devices," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4589–4597, Dec. 2018.
- [17] A. Ahmad, A. Paul, M. M. Rathore, and H. Chang, "Smart cyber society: Integration of capillary devices with high usability based on cyber-physical system," *Future Gener. Comput. Syst.*, vol. 56, pp. 493–503, Mar. 2016.
- [18] L. Xiao, L. Liu, S. Asseng, Y. Xia, L. Tang, B. Liu, W. Cao, and Y. Zhu, "Estimating spring frost and its impact on yield across winter wheat in China," *Agric. Forest Meteorol.*, vols. 260–261, pp. 154–164, Oct. 2018.
- [19] K. Barlow, B. Christy, G. O'Leary, P. Riffkin, and J. Nuttall, "Simulating the impact of extreme heat and frost events on wheat crop production: A review," *Field Crops Res.*, vol. 171, pp. 109–119, Feb. 2015.
- [20] H. Lee, J. A. Chun, H.-H. Han, and S. Kim, "Prediction of frost occurrences using statistical modeling approaches," *Adv. Meteorol.*, vol. 2016, pp. 1–9, Mar. 2016.
- [21] K. Brun-Laguna, A. L. Diedrichs, J. E. Chaar, D. Dujovne, J. C. Taffernaberry, G. Mercado, and T. Watteyne, "A demo of the PEACH IoT-based frost event prediction system for precision agriculture," in *Proc. 13th Annu. IEEE Int. Conf. Sens., Commun., Netw. (SECON)*, Jun. 2016, pp. 1–3.
- [22] A. Ahmad, M. Babar, S. Din, S. Khalid, M. M. Ullah, A. Paul, A. G. Reddy, and N. Min-Allah, "Socio-cyber network: The potential of cyber-physical system to define human behaviors using big data analytics," *Future Gener. Comput. Syst.*, vol. 92, pp. 868–878, Mar. 2019.
- [23] M. Webb, A. Pirie, D. Kidd, and B. Minasny, "Spatial analysis of frost risk to determine viticulture suitability in Tasmania, Australia," *Austral. J. Grape Wine Res.*, vol. 24, no. 2, pp. 219–233, Apr. 2018.

- [24] P. J. Caradonna and J. A. Bain, "Frost sensitivity of leaves and flowers of subalpine plants is related to tissue type and phenology," *J. Ecol.*, vol. 104, no. 1, pp. 55–64, Jan. 2016.
- [25] M. Bascietto, S. Bajocco, F. Mazzenga, and G. Matteucci, "Assessing spring frost effects on beech forests in central apennines from remotely-sensed data," *Agricult. Forest Meteorol.*, vol. 248, pp. 240–250, Jan. 2018.
- [26] D. F. G. Junco, D. F. Diaz Caro, M. S. C. Forero, and I. A. R. Ruge, "Agrometeorological monitoring station based microcontroller and Bluetooth communication," in *Proc. IEEE 2nd Colombian Conf. Autom. Control (CCAC)*, Oct. 2015, pp. 1–5.
- [27] A. Paul, A. Ahmad, M. M. Rathore, and S. Jabbar, "Smartbuddy: Defining human behaviors using big data analytics in social Internet of Things," *IEEE Wireless Commun.*, vol. 23, no. 5, pp. 68–74, Oct. 2016.
- [28] M. Jorenoosh and A. Sepaskhah, "Prediction of frost occurrence by estimating daily minimum temperature in semi-arid areas in Iran," *Iran Agricult. Res.*, vol. 37, no. 1, pp. 19–32, 2018.
- [29] S. Gu, L. Yao, C. Tunc, A. Akoglu, S. Hariri, and E. Ritchie, "An autonomic workflow performance manager for weather research and forecast workflows," in *Proc. Int. Conf. Cloud Autonomics Computing (ICCAC)*, Sep. 2016, pp. 111–114.
- [30] A. Castañeda-Miranda and V. M. Castaño, "Smart frost control in greenhouses by neural networks models," *Comput. Electron. Agricult.*, vol. 137, pp. 102–114, May 2017.
- [31] F. G. Antivilo, R. C. Paz, M. Echeverria, M. Keller, J. Tognetti, R. Borgo, and F. R. Juñent, "Thermal history parameters drive changes in physiology and cold hardiness of young grapevine plants during winter," *Agricult. Forest Meteorol.*, vol. 262, pp. 227–236, Nov. 2018.
- [32] S. Jabbar, K. R. Malik, M. Ahmad, O. Aldabbas, M. Asif, S. Khalid, K. Han, and S. H. Ahmed, "A methodology of real-time data fusion for localized big data analytics," *IEEE Access*, vol. 6, pp. 24510–24520, 2018.
- [33] M. Fuentes, C. Campos, and S. García-Loyola, "Application of artificial neural networks to frost detection in central Chile using the next day minimum air temperature forecast," *Chilean J. Agricult. Res.*, vol. 78, no. 3, pp. 327–338, Sep. 2018.
- [34] M. Imran, S. Jabbar, N. Chilamkurti, and J. J. Rodrigues, "Enabling technologies for social Internet of Things," *Future Gener. Comput. Syst.*, vol. 92, pp. 715–717, Mar. 2019.
- [35] A. Ahmad, A. Paul, M. Rathore, and H. Chang, "An efficient multidimensional big data fusion approach in machine-to-machine communication," *TECSACM Trans. Embedded Comput. Syst.*, vol. 15, no. 2, pp. 1–25, Jun. 2016.
- [36] O. Arogundade, A. Abayomi-Alli, I. Fatoye, C. Adejuyigbe, and V. Olowe, "Development of an Android based mobile application for the production and management of organic manure (MoAPOM)," *J. Organic Agricult. Environ.*, vol. 6, 2019.
- [37] A. Ahmad, M. Khan, A. Paul, S. Din, M. M. Rathore, G. Jeon, and G. S. Choi, "Toward modeling and optimization of features selection in big data based social Internet of Things," *Future Gener. Comput. Syst.*, vol. 82, pp. 715–726, May 2018.
- [38] S. Din, A. Ahmad, A. Paul, and S. Rho, "MGR: Multi-parameter green reliable communication for Internet of Things in 5G network," *J. Parallel Distrib. Comput.*, vol. 118, pp. 34–45, Aug. 2018.
- [39] J. Ruan, H. Jiang, X. Li, Y. Shi, F. T. S. Chan, and W. Rao, "A granular GA-SVM predictor for big data in agricultural cyber-physical systems," *IEEE Trans. Ind. Informat.*, vol. 15, no. 12, pp. 6510–6521, Dec. 2019.



**RANA M. AMIR LATIF** received the B.S.C.S. degree from COMSATS University Islamabad–Sahiwal, Sahiwal, Pakistan, where he is currently pursuing the M.S. degree. He is also a Research Assistant with the Department of Computer Science, CUI, under the supervision of Dr. M. Farhan. He is also a Lecturer with the Barani Institute of Sciences, Sahiwal. His research interests include machine learning, visual graphs analysis, deep learning, and the IoT.



**SAMIR BRAHIM BELHAOUARI** received the master's degree in telecommunications and network from the Institut Nationale Polytechnique of Toulouse, France, in 2000, and the Ph.D. degree from the Federal Polytechnic School of Lausanne-Switzerland, in 2006. He is currently an Associate Professor with the Division of Information and Communication Technologies, College of Science and Engineering, HBKU. He also holds several positions at the University of Sharjah, Innopolis University, Petronas University, and the EPFL Federal Swiss School.



**SAQIB SAEED** received the B.Sc. degree (Hons.) in computer science from the International Islamic University Islamabad, Pakistan, in 2001, the M.Sc. degree in software technology from the Stuttgart University of Applied Sciences, Germany, in 2003, and the Ph.D. degree in information systems from the University of Siegen, Germany, in 2012. He is currently a Certified Software Quality Engineer with the American Society of Quality. He is also an Assistant Professor with the Department of Computer Information Systems, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia. His research interests include human-centered computing, software engineering, information systems management, and digital business transformation. He is a member of the advisory boards of several international journals besides being a Guest Editor of several special issues.



**LAIQA BINTE IMRAN** received the B.S.C.S. degree from COMSATS University Islamabad–Sahiwal, Sahiwal, Pakistan, where she is currently pursuing the M.S. degree. Her research interests include Web engineering, data science, networking, and the IoT.



**MAZHAR SADIQ** was born in Pakistan. He received the bachelor's degree in computer science and the master's degree in computer science from Pakistan. After completing the master's degree, he went to Sweden for higher education. He received the Master of Science degree from the Blekinge Institute of Technology, Sweden, in 2007, and the Ph.D. degree in philosophy from the University of Jyväskylä, Finland, in 2016. Since March 2017, he has been an Assistant Professor with COMSATS University Islamabad–Sahiwal, Sahiwal, Pakistan. He has been teaching Human–Computer Interaction course to Bachelor of Computer Science students. He has supervised many Master of Computer Science students. He is the author of more than six articles. His research interests include usability, virtual reality, and augmented reality.



**MUHAMMAD FARHAN** received the B.S.C.S. degree from the Virtual University of Pakistan (VU), in 2007, the M.S.C.S. degree from the University of Management and Technology (UMT), Pakistan, in 2010, and the Ph.D. degree in computer science from the Department of Computer Sciences and Engineering, University of Engineering and Technology (UET), Pakistan, in 2017. He had worked as a Lecturer in Department of Computer Science, COMSATS University Islamabad–Sahiwal, Sahiwal, Pakistan. He also had worked as an Instructor of Computer Science at the Virtual University of Pakistan for about five years. He is currently an Assistant Professor with the Department of Computer Science, COMSATS University Islamabad–Sahiwal. He started his research career

with the publication of a conference paper in San Luis Potosí, Mexico, IEEE CPS. He has honored for winning travel grant provided by ACM and Microsoft for the presentation of a student paper in the ACM/SIGAPP Symposium (Human–Computer Interaction track) held by the University of Salamanca, Spain. He has published a good number of SCI-indexed impact factor journal articles, which are published by the *Journal of Real-Time Image Processing* by Springer, *Multimedia Tools and Applications* by Springer, the *International Journal of Distributed Sensor Networks* by SAGE Journals, the *EURASIA Journal of Mathematics, Science and Technology Education* by MODESTUM, *Life Science Journal* by Marshland Press and in various renowned journals of IEEE, Springer, Elsevier, and Hindawi. His research interests include data science, machine and deep learning, and the Internet of Things.

...