G Model
SUSCOM-308; No. of Pages 7

ARTICLE IN PRESS

Sustainable Computing: Informatics and Systems xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Sustainable Computing: Informatics and Systems

journal homepage: www.elsevier.com/locate/suscom

Estimation of wheat planting date using machine learning algorithms based on available climate data

Abdülkadir Gümüşçü[a,*], Mehmet Emin Tenekeci[b], Ali Volkan Bilgili[c]

[a] Harran University, Osmanbey Campus Electrical-Electronics Eng. Dept, Şanlıurfa, Turkey
[b] Harran University, Osmanbey Campus Computer Eng. Dept, Şanlıurfa, Turkey
[c] Harran University, Osmanbey Campus Soil Sciences and Plant Nutrition Dept, Şanlıurfa, Turkey

ARTICLE INFO

Article history:
Received 14 June 2018
Received in revised form 10 January 2019
Accepted 18 January 2019
Available online xxx

Keywords:
Planting date estimation
Metereological data
Feature selection
Classification
Genetic algorithm

ABSTRACT

Agricultural applications supported with information technologies increase plant production, protect soil and reduce labor, which is crucial for sustainable agriculture. Impacts of planting dates on production are very well known. In the current study machine learning algorithms have been used in determining planting date. The proposed method aims to help farmers to obtain higher yield providing them with accurate planting date. For this purpose, metereological data was used as an input. For each year, metereological information (Daily Maximum Air Temperature, Daily Relative Humudity, Daily Average Air Temperature, Daily Minimum Air Temperature and Daily Precipitation) in the first 300 days were used to determine three different planting dates; early, normal and late for wheat crop. For estimation of planting date, classification algorithms of k Nearest Neighbor (kNN), Support Vector Machine (SVM) and Decisions Trees were used. Performances of different algorithms were calculated with leave one out cross validation approach. In order to eleminate extremely high processing time because of high dimension of the data set and improve estimation performance, genetic algorithm was used to reduce the number of features. For the estimations performed using both all features and also the features selected with genetic algorithm the highest accuracies were obtained using kNN method with classification accuracy rates of 37% and 92%, respectively. Overall, the results showed that wheat planting date could be determined successfully from climate information obtained in the first 300 days with the help of machine learning techniques combined with feature selection using genetic algorithm, which will prevent low productivity, financial and labor loss as a result of inaccurate planting date.

1. Introduction

Global warming is one of the top problems which human being will face in future and Turkey is considered as one of the countries which will suffer from global warming most. Because of global warming, the shifts in time periods of the seasons and climatic events occurring untimely and uncharacteristically highly impact seasonal industries. Food and agriculture industry gained importance with increasing population is also negatively affected from global warming [1], therefore it is necessary to manage agricultural lands with more productive management models and agricultural management processes such as fertilizing, irrigation, harvesting should be improved for more efficient usage of agricultural lands. Especially impacts of planting dates on production are very well known [2]. In present situation, for the planting dates of agricultural products is relied upon a fixed planting time period or upon the estimation of current climate conditions. Fixed planting dates will not provide successful results. In these times when the effects of global warming are seen more, the shifts in the beginnings and endings periods of seasons makes the fixed planting date unreasonable and disadvantageous [3]. Donatelli et al. reported that climate change decrease the productivity of agricultural products and suggested that the easiest way to reduce this impact was to modify the planting date [3]. In addition to that, they adapted a different approach proposing to consider climate data throughout the year in determining the date of plantation.

Cammarano et al. performed studies on wheat crops and investigated how the water management and planting date impact crop productivity [4]. They also developed strategies in the case of serious climate changes that may occur in the future modeling climate conditions between years of 2030 and 2070. Dobor et al. performed studies to determine planting date using information about soil temperature and soil moisture in Hungary and they found that the

* Corresponding author.
E-mail addresses: agumuscu@harran.edu.tr (A. Gümüşçü), etenekeci@harran.edu.tr (M.E. Tenekeci), vbilgili@harran.edu.tr (A.V. Bilgili).

https://doi.org/10.1016/j.suscom.2019.01.010
2210-5379/

Please cite this article in press as: A. Gümüşçü, et al., Estimation of wheat planting date using machine learning algorithms based on available climate data, Sustain. Comput.: Inform. Syst. (2018), https://doi.org/10.1016/j.suscom.2019.01.010

Year 1	1*300 daily min. temperature	1*300 daily max. temperature	1*300 daily average temperature	1*300 daily relative humidity	1*300 daily precipitation	Planting Date Group
Year 2	1*300 daily min. temperature	1*300 daily max. temperature	1*300 daily average temperature	1*300 daily relative humidity	1*300 daily precipitation	Planting Date Group
Year 3	1*300 daily min. temperature	1*300 daily max. temperature	1*300 daily average temperature	1*300 daily relative humidity	1*300 daily precipitation	Planting Date Group
Year 51	1*300 daily min. temperature	1*300 daily max. temperature	1*300 daily average temperature	1*300 daily relative humidity	1*300 daily precipitation	Planting Date Group

Fig. 1. Structure of the data set.

planting date was 12 days earlier than previous years because of global warming [5].

The studies performing planting date can be sorted into four main groups; the first group is to determine a fixed planting date based on observations [6]. Secondly the studies optimizing planting date [1], Thirdly is rule based planting date determination [7,8] and finally determining planting date to improve the product using climate data [9].

Climate data has been used in multiple studies related with estimation of planting date, which lead to the studies be handled at regional level [9]. Different artificial learning models can be constructed by new data sets that will be formed for other regions.

The methods forming a model considering the impacts of the relationships among variables on the outputs and using this model for future decisions are called as artificial learning techniques. Artificial learning techniques are often used in the areas like agriculture [10], medicine [11] and engineering [12,13]. In this study they have been used for the estimation of planting date. According to the review of earlier studies, it is apparent to say that a relationship exists between variables of the climate data set and planting date [14]. As a solution to this problem, the machine learning methods that can find the relations between features and output in a most effective way have been used and the successful results obtained have been presented.

By being used in various areas, machine learning algorithms have made significant contributions to development of the concept of machine learning methods. Especially machine learning algorithms are often used in process of identification and estimation procedures. In addition to identification processes like fingerprint, iris, walking, and gender the estimations processes such as temperature, diseases, emotion can be performed with machine learning algorithms. Machine learning algorithms form a model with relationship between quality and results with determined training set and the success of the model can be determined using a test set. Machine learning algorithms is an strong and flexible method to produce solutions for the problems that are not being handled with traditional approaches. Therefore, the use of machine learning algorithms are increasing over time and they are being used in many sectors of social life, which make them an important factor in solving the problems and giving decisions. Data sets are data collections in which information about events, experiments or truths are collected. When used with machine learning algorithms, data set provide foresights and classifications for future.

In this study, estimation of planting dates in agricultural products using machine-learning algorithms is proposed. A model was developed for determining plantation date based on climate data, which will be impacted from the climate changes due to global warming as little as possible.

The proposed method makes prediction of planting date taking the climate conditions of earlier months into consideration and gives information about whether an early, a normal or a late planting will be made in corresponding year at the beginning of planting period by making connections between yearly climate data and planting time. Thus farmers will have a general knowledge about planting date prior to planting period, which will prevent probable agricultural yield losses especially in these days when food concept is significant.

A data set was formed from 51 years yearly climate data to estimate planting date for wheat crop. Expert opinions have been utilized for determining planting dates. Planting date estimation was performed using three different machine learning algorithms such as k nearest neighbor, support vector machine and decisions trees and the performance results obtained were tested using one out cross validation.

2. Materials and methods

2.1. Data set

The data set used in the study consisted of the climate data of Şanlıurfa city of Turkey and it was formed based on the most relevant planting dates determined based on expert opinion. In the study as agricultural production, wheat, a common staple crop of the region and the country, was preferred. The climate data was obtained from Turkish State Meteorological Services and consisted of 53 years long term climate data of Şanlıurfa city between 1965 and 2017 years. Due to missing data in different two years, the study was performed using only climate data of 51 years. In order to standardize day numbers among years, leftover days in the month of February were not included into the data set. So that all years were presumed as 365 days. For yearly data, information about daily minimum temperature (°C), daily average temperature (°C), daily maximum temperature (°C), daily relative humudity (%) and daily total rainfall rate (mm) data in the first 300 days of each years were defined as feature.

In the study, information about planting date as early, normal and late planting will be provided to the farmers in advance (in 300th day of the year) with machine learning methods depending on the climate data of the past 300 days. So that a total 1500 features (300 days x 5 daily climate variables) were recorded in each year (Fig. 1).

In addition, within each data set there is information about the day showing the most ideal planting date based on corresponding climate data within the same year. This information was obtained from the experts who were provided with daily climate data aforementioned above. Descripitive statistics of the data set are given in Table 1.

With these information obtained, planting date were separated into three different groups as early, normal and late based on the threshold values given in Table 2. These threshold values were determined using multi threshold otsu algorithms [15] which separated class labels homogeneously.

Please cite this article in press as: A. Gümüşçü, et al., Estimation of wheat planting date using machine learning algorithms based on available climate data, Sustain. Comput.: Inform. Syst. (2018), https://doi.org/10.1016/j.suscom.2019.01.010

Table 1
Descriptive statistics of the data set.

Climate variables	Average value	Standart deviation	Min	Max
Daily Maximum Temperature (°C)	26,384	10,707	−2,200	46,800
Daily Relative Humudity (%)	48,414	19,878	10,000	98,700
Daily Average Temperature (°C)	20,267	9,812	−5,500	38,200
Daily Minimum Temperature (°C)	14,429	8,510	−9,600	32,500
Daily Precipitation Amount (mm)	1,046	4,193	0,000	117,100

Table 2
Groups of planting date.

Information about days based on expert opinion	Classess
X<324	Early planting
324<=X<335	Normal planting
335<X	Late planting

X: Planting date provided by the expert.

Based on the conditions given in Table 2, the information of planting date was classified into three groups. Accordingly 13 samples were grouped in the early planting date class, 24 in the normal planting date class and 14 samples in the late planting date class. So that based on the climate conditions of the first 300 days of the year, estimations of planting date would be provided to the farmers as early, normal or late planting and they could prepare their field accordingly.

2.2. Feature selection methods

Feature selection is a process of finding the most relevant variables for an estimation model and it is used to determine and remove irrelevant, unneeded and redundant features that do not impact the accuracy of the model. Mathematically, feature selection is considered as optimization problem. Aim of the optimization is to reach to the result aimed fast and successfully with the least number of features and the aim of the feature selection methods is to perform more successful classification procedure by removing the least related features with the result. Feature selection process is often used in machine learning as a pre-process step. Earlier studies indicate that there are different approaches to determine the best feature subsets, each of which may have own advantages and disadvantages. There are three different common feature selection methods in literature, which are filters, wrappers, and embedded techniques [16]. In this study feature selection processes with genetic algorithm, one of the wrappers feature selection models was applied.

2.2.1. Feature selection using genetic algorithm

The Genetic Algorithm is a meta - heuristic algorithm and based on the fact that in nature finding the best suitable genes for generations to better adapt to the environment [17]. Feature selection using genetic algorithm aims to select the most relevant features in order to minimize the error of the model. This procedure is fitness function that differs for each problem. The value of this function is brought to maximum or minimum value according to the type and structure of the problem.

In the method the number of population to be formed is randomly determined and a fitness value is calculated according to this formed population. Later a new population is formed by mutation and crossing and a fitness value is calculated for new population. According to fitness value, this process continues iteratively until the most successful population is found. In the study, the features that are suited to determined population were used.

In the study there are 1500 features (300 * 5) for each sample. The use of all these feature will require significant processing time as well as affect the accuracy negatively. For the genetic algoritm used to reduce the number of feature, 200 populations were used and 250 generations were realized. Termination value was determined as 0.

̈For the feature selection the genetic algorithm (GA) generates a random number as either 1 or 0 for all the features in the data set. 1 and 0 respectively mean ̈featureänd n̈ot feature.̈ I.e. if a binary number of 10000001 is generated by GA for a data set with 8 features; it can be said that only the first and last features are selected at this step. When classification error rate reaches to zero, GA stops to search."

2.3. Classification methods

With classification method a model is created by establishing a relationship between features given for every sample and result labels in the data set and this model is used to make estimations for the samples whose result label is unknown. In the study in order to assess the accuracy of the determination of planting date using climate information, statistical measures of performances given in equations from 1 to 5 such as Sensitivity, Specifity, Accuracy, Precision and F-Score were used;

$$Sensitivity = TP/(TP + FN) \quad (1)$$

$$Specifity = FP/(FP + TN) \quad (2)$$

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (3)$$

$$Precision = TP/(TP + FP) \quad (4)$$

$$F\text{-}Score = 2 * Sensitivity * Precision/Sensitivity + Precision \quad (5)$$

Where ; True positive (TP) = the number of cases correctly identified as positive

False positive (FP) = the number of cases incorrectly identified as positive

True negative (TN) = the number of cases correctly identified as negative

False negative (FN) = the number of cases incorrectly identified as negative

In the study three different classification techniques; K nearest neighbor (kNN), Support Vector Machines (SVM) and Decision Trees have been used. Leave one out cross validation approach was applied to determine success parameters of machine learning methods. The machine learning algorithms used in the study are summarized below.

2.3.1. k Nearest neighbor (kNN)

kNN method is one of the commonly used methods in various areas because it is easy to use and classification algorithm is uncomplicated [18]. In the training stage of kNN algorithm, samples in training data set is related with classes placing them on coordinate plane. The samples to be used for testing is later placed on coordinate plane established. Class of test data is determined looking at the test data at the nearest neigbors. The class of test data is estimated that it belongs to the class where neighbor samples are belonging to the most. In order to run kNN algorithm a k value need to be entered by the user.

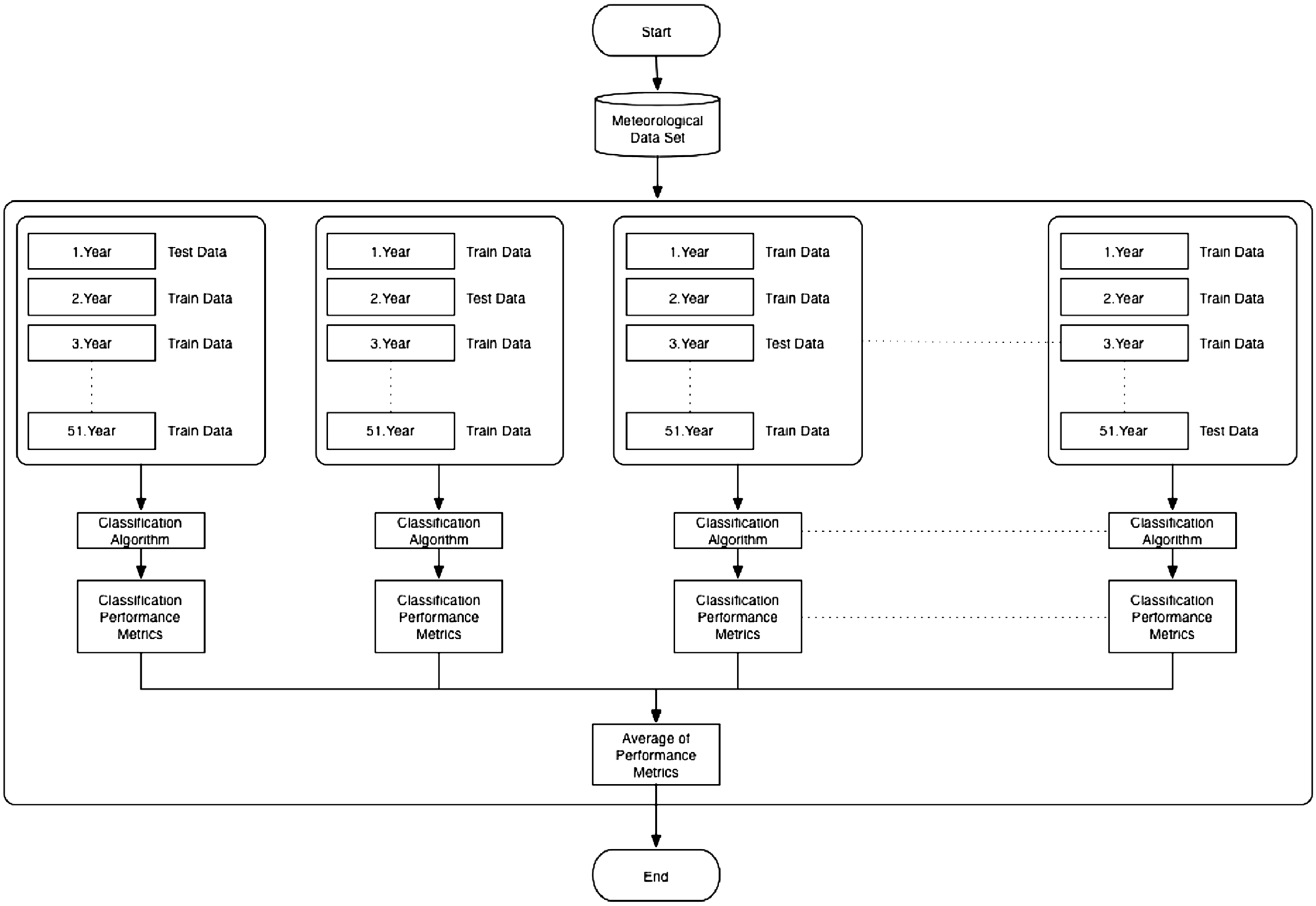

Fig. 2. Flow chart of classification using all feautures.

Enabling the algorithm run requires the user to enter a k value externally. In addition to k value, distance criteria need to be determined by the user. In this study k value was selected as 1 and neighbor distance measurement was calculated according to euclidean distance.

2.3.2. Support vector machines (SVM)

Support vector machines are commonly used especially in reciprocal classification. This method was developed by Vapnik for the solution of classification problems [19]. The most important advantage of Support Vector Machines is their low processing time. Becuase of this character, it is prefered in the data sets with high number of samples.

2.3.3. Decision trees

Decision trees are used for classification and other estimation purposes and it is one of the common machine learning algorithms due to ease of interpretation and understanding. A decision tree model consists of parts such as root, node and leave.

The samples in test set is classified with tree structure formed according to values of attributes in order for calculation of rate of success of the tree structure formed with training data. The decisions tree algorithm used in the study is C4.5 algorithm [20].

3. Estimation of planting date with machine learning methods

In this study procedures shown in Fig. 2 were applied to data set formed and the performance metrics were calculated using leave one out cross validation approach

As seen in Fig. 2 one of 51 sample was determined as test set and the rest 50 samples were used as training data set to form machine learning model. This situation was repeated in 51 different iteration untill all samples become test data. Performance metrics for the data set was calculated using leave one out cross validation taking average of classification performance metrics obtained in each iteration.

In addition to these procedures, prior to classification process performance metrics were also calculated reducing the number of features applying wrappers feature selection with genetic algorithm. Steps in the study are shown in Fig. 3.

For calculation of the performance parameters of classification algorithms, a computer with 8 GB RAM, 2.8 GHz and Intel Core i7 process was used.

4. Results

In this study, wheat planting date has been estimated using 51 years metereological data obtained from climate stations located in GAP regions. For estimation procedure, three different machine learning algorithms, namely kNN, Decision Tree and SVM, have been used. The metrics showing the performance of different classification algorithms are given in Table 3.

The performances of the methods were poor because of the data set with the number of variables were higher than the number of samples, which is the case of a high p and low n problem. Current climate data set contains 51 sample and 1500 features. To avoid this, before classification of climate data set, as a pre-processing process wrappers feature selection with genetic

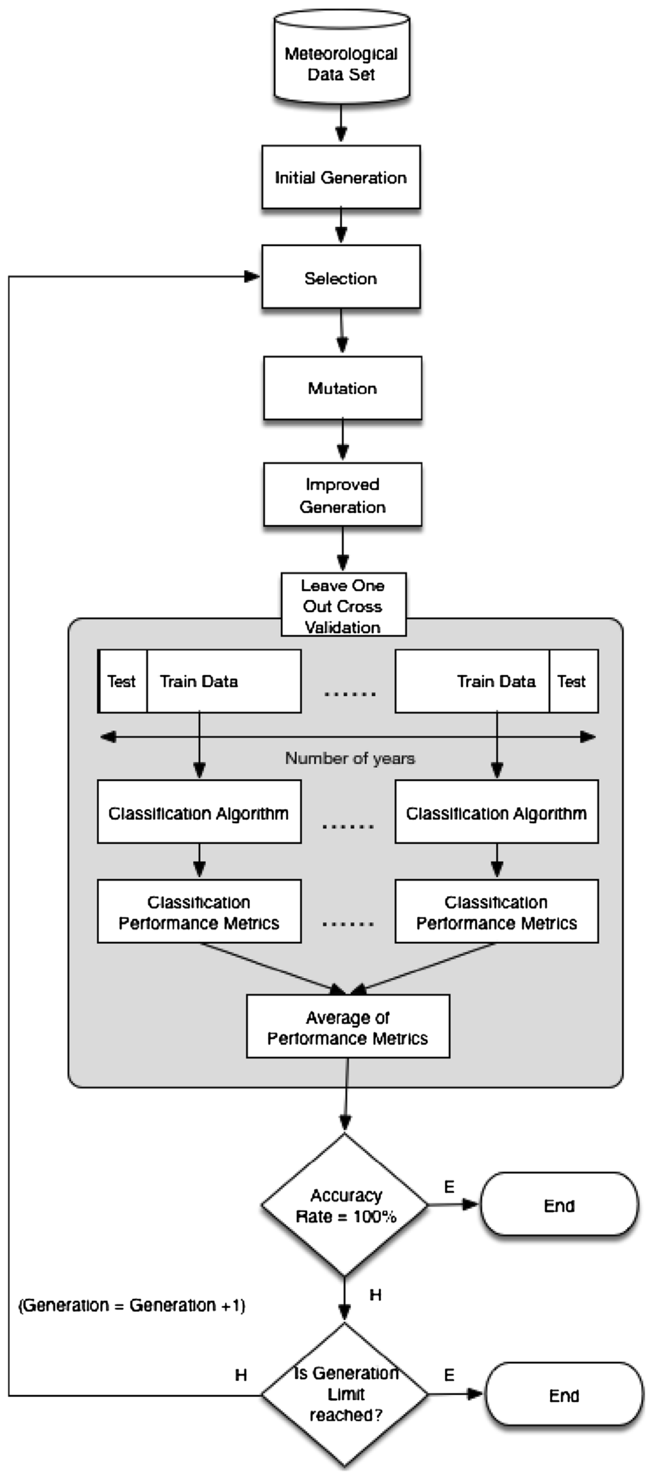

Fig. 3. Flow chart of classification using features determined feature selection with genetic algorithm.

Table 3
Classification performance metrics obtained using all features.

Classification Algorithms	Accuracy	Sensitivity	Specificity	Precision	F-Score
kNN	0,372	0,458	0,296	0,366	0,407
Decision Tree	0,176	0,291	0,074	0,218	0,250
SVM	0,352	0,500	0,222	0,363	0,421

Table 4
Classification performance metrics features selected using feature selection with genetic algorithm.

Classification Algorithm	Accuracy	Sensitivity	Specificity	Precision	F-Score
kNN	0,921	0,916	0,925	0,916	0,916
Decision Tree	0,862	0,916	0,814	0,814	0,862
SVM	0,470	0,666	0,296	0,457	0,542

algorithms have been applied and Table 4 summarizes the results of applying feature selection algorithms procedure to the datasets prior to classification using algorithms kNN, Decisions Tree and SVM

After feature selection procedure with the application of genetic algoritm, the number of features selected are given in Table 5. Table 5 also shows how many features are selected from which attributes.

Accuracies of classification algorithms have been improved after application of feature selection using genetic algorithm increasing classification success rate. Genetic algorithm is an optimization method performing iterative process. Amount of improvements obtained in each generation for kNN, Decision Tree and SVM classification algorithms were shown in Fig. 4.

Also as seen in Fig. 4, 250 generations has not been reached in Decision Tree and SVM classifiers which is because of the fact that there is no difference between new generations and previous generations.

5. Discussion

The studies on estimation of relevant planting date for agricultural crops especially for wheat crop using machine learning algorithms are quite limited and there is no available data set dealt with that. The methods already applied in determining planting dates generally include either the determination of planting date depending on weather forecasts made on present days or it is performed according to fixed dates within the year. Given the global warming-related climate changes, these traditional methods of determining planting date cause labor and financial loss. The method proposed in this study is thought to prevent such losses. In this study wheat planting time periods were estimated using 51 years of climate data measured in Şanlıurfa city, Southeastern Anatolia Region, Turkey through machine learning algorithms. For classification of the data set, classification algorithms of kNN, Decision Tree and SVM have been used.

Classification success criteria without attribute selection were found to be unacceptable. The results of classification procedures

Table 5
The number of features selected by genetic algorithm.

Classification Algorithm	Daily Min. Temp.	Daily Max. Temp.	Daily Avr. Temp.	Daily Rel. Humi.	Daily Preci. Amo.	Total selected feature
kNN	163	143	146	156	125	733
Decision Tree	146	150	141	143	134	714
SVM	146	133	148	136	131	694

Please cite this article in press as: A. Gümüşçü, et al., Estimation of wheat planting date using machine learning algorithms based on available climate data, Sustain. Comput.: Inform. Syst. (2018), https://doi.org/10.1016/j.suscom.2019.01.010

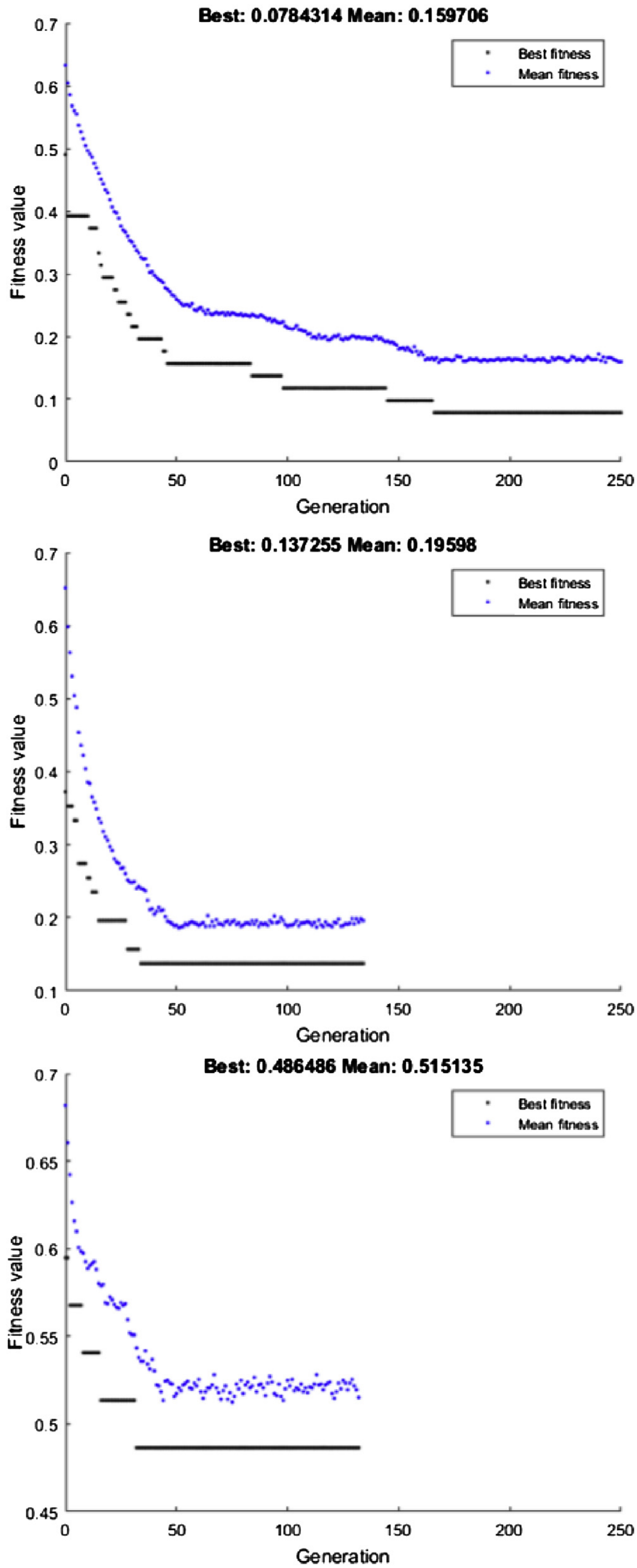

Fig. 4. Accuracy assessment of classification algorithms, kNN, Decision Tree and SVM from top to bottom.

used in the study revealed that feature selection algorithms are required in order to achieve best classification results. The small number of samples and the large number of features in the data set are the most important parameters affecting the success rate of machine learning methods [16]. Therefore the reason of obtaining performance metrics in poor quality without feature selection can be interpreted as having 51 samples in the data set versus 1500 attributes. Evalution of the performance parameters without feature selection showed that the best accuracy for estimations were produced using kNN classification algorithm with 0.372 accuracy rate and using SVM with f-score value of 0.421 (Table 3).

Comparison of success metrics obtained without attribute selection showed that Decision Trees had the lowest performance. Attributes in the dataset were considered to be numeric, which prevented the decision trees from effectively thresholding in the training phase and separating the classes successfully. The most successful results in classification without feature selection were obtained by kNN method. kNN was relatively less impacted by structure of the data set showing that kNN had a more robust training model against high p low n problem.

Classification using features selected with feature selection provided expected results and improved success metrics significantly. In particular, is the value of the amount of daily rainfall in the dataset in many days for the GAP region, which is a dry region in summer, was assigned as 0. Thus, feature selection method have successfully improved performance metrics by eliminating unnecessary and unrelated attributes associated with class information. The wrapper model feature selection algorithms improved the success rates, but it had the disadvantage of long operation times [16]. Therefore, the feature selection process by genetic algorithm was applied. As seen in Fig. 4, the lowest error rates were achieved in generations that proceed iteratively with the genetic algorithm.

In classifications after the application of feature selection, the SVM provided the lowest performance. Although, the feature selection was applied into the data set, the number of features could only be reduced around between 600 and 750. Therefore SVM that distinguishes among classes with linear kernel was not able to perform a successful classification. In addition, SVM method that provides successful results especially in binary classification was seriously affected negatively when the number of classes increased. On the other hand, the feature selection methods made the most important contributions for the decision trees method. Particularly, in the educational stage of the decision tree classification, the reduction in the number of features lead to the selection of the first branch more accurately and the prevention of unnecessary branching, which made the feature selection method be more successful in decision trees.

6. Conclusions

The proposed method estimated the most relevant planting date of wheat crop using metereological data in the first 300 days of the year obtained from long term climate data. For the estimation of wheat planting date which were seperated into three groups as early, normal and late, different classification algorithms (kNN, SVM and Decision Tree) were used and their performances were compared with one out cross validation approach. Feature selection with genetic algorithms used to reduce the number of features in the data set improved the performances of all classification algorithms at varying levels. According to performance metrics of different classification algorithms with and without feature selection, the most succesfull results were obtained with kNN method. Up to 92.5% accuracy performances of the classifiers obtained revealed that climate parameters recorded in the first 300 days of the year were highly related with wheat planting date and could be successfully used in deciding the most appropriate dates for wheat plantation in order to eleminate or reduce yield lossess due to late planting. Our results also demostrated a modelling approach for farmers that can be effective in minimising adverse effects of climate change in sustainable agriculture.

Please cite this article in press as: A. Gümüşçü, et al., Estimation of wheat planting date using machine learning algorithms based on available climate data, Sustain. Comput.: Inform. Syst. (2018), https://doi.org/10.1016/j.suscom.2019.01.010

Acknowledgement

We would like to thank to Turkish State Metereological Services for providing us with the valuable metereological data used in the study and Expert Ali İlkhan, Expert Memduh Serdar Polat from GAPTAEM for sharing his experience and valuable contributions

References

[1] M. Waongo, P. Laux, H. Kunstmann, Adaptation to climate change: the impacts of optimized planting dates on attainable maize yields under rainfed conditions in Burkina Faso (in English), Agric. For. Meteorol. 205 (June (1)) (2015) 23–39.
[2] A.A. Rasouli, B.S. Sarraf, G. Kamali, S. Fateh, Effect of planting date choice on the vulnerability of winter wheat to climate change: case study of cool temperate, northwestern provinces of Iran, Biosci. Biotechnol. Res. Commun. 10 (1) (2017) 172–180.
[3] M. Donatelli, A.K. Srivastava, G. Duveiller, S. Niemeyer, D. Fumagalli, Climate change impact and potential adaptation strategies under alternate realizations of climate scenarios for three major crops in Europe, Environ. Res. Lett. 10 (July (7)) (2015), 075005.
[4] D. Cammarano, J. Payero, B. Basso, L. Stefanova, P. Grace, Adapting wheat sowing dates to projected climate change in the Australian subtropics: analysis of crop water use and yield, Crop Pasture Sci. 63 (10) (2012) 974–986.
[5] L. Dobor, Z. Barcza, T. Hlasny, T. Arendas, T. Spitko, N. Fodor, Crop planting date matters: estimation methods and effect on future yields (in English), Agric. For. Meteorol. 223 (June) (2016) 103–115.
[6] N. de Noblet-Ducoudre, et al., Coupling the Soil-Vegetation-Atmosphere-Transfer Scheme ORCHIDEE to the agronomy model STICS to study the influence of croplands on the European carbon and water budgets (in English), Agronomie 24 (September–November (6–7)) (2004) 397–407.
[7] D.P. Holzworth, et al., APSIM - Evolution towards a new generation of agricultural systems simulation, Environ. Model. Softw. 62 (December) (2014) 327–350.
[8] K. Waha, L.G.J. van Bussel, C. Muller, A. Bondeau, Climate-driven simulation of global crop sowing dates, Glob. Ecol. Biogeogr. 21 (Febuary (2)) (2012) 247–259.
[9] C.J. Kucharik, Contribution of planting date trends to increased maize yields in the central United States (in English), Agron. J. 100 (March–April (2)) (2008) 328–336.
[10] G. Russ, R. Kruse, Machine learning methods for spatial clustering on precision agriculture data (in English), Eleventh Scandinavian Conference on Artificial Intelligence (Scai 2011) 227 (2011) 40–49.
[11] N. Lavrac, Machine learning for data mining in medicine (in English), Artif. Intell. Med. 1620 (1999) 47–62.
[12] S. Chien, et al., Machine learning in engineering automation (in English), Mach. Learn. (1991) 577–580.
[13] J. Dyck, Machine learning for engineering (in English), in: 2018 23rd Asia and South Pacific Design Automation Conference (Asp-Dac), 2018, pp. 422–427.
[14] M. Trnka, et al., Agroclimatic conditions in Europe under climate change, Glob. Chang. Biol. 17 (July (7)) (2011) 2298–2318.
[15] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Syst. Man Cybern. 9 (1) (1979) 62–66.
[16] I.I. Yvan Saeys, Pedro Larrañaga, A review of feature selection techniques in bioinformatics, Bioinformatics 23 (19) (2007) 2507–2517.
[17] J.H. Holland, Genetic algorithms, Sci. Am. 267 (1992).
[18] T. Cover, P. Hart, Nearest neighbor pattern classification, IEEE Trans. Inf. Theory 13 (1) (1967) 21–27.
[19] V.N. Vapnik, Statistical Learning Theory: Inference from Small Samples, Wiley, New York, 1998.
[20] J.R. Quinlan, C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers Inc, San Francisco, CA, USA, 1993.

Please cite this article in press as: A. Gümüşçü, et al., Estimation of wheat planting date using machine learning algorithms based on available climate data, Sustain. Comput.: Inform. Syst. (2018), https://doi.org/10.1016/j.suscom.2019.01.010