Current Trends in Natural Sciences
Vol. 11, Issue 21, pp. 304-314, 2022
https://doi.org/10.47068/ctns.2022.v11i21.033
Current Trends in Natural Sciences (on-line)
ISSN: 2284-953X
ISSN-L: 2284-9521
Current Trends in Natural Sciences (CD-Rom)
ISSN: 2284-9521
ISSN-L: 2284-9521

CLIMATE ANALYSIS AND WEATHER FORECASTING WITH DATA MINING: THE CASE OF ADANA PROVINCE IN TURKEY

Mümine Kaya Keleş [1,*], Elif Kavak [1]

[1]Adana Alparslan Turkes Science and Technology University, Faculty of Computer and Informatics, Department of Computer Engineering, 01250, Adana, Turkey

Current Trends in Natural Sciences

Abstract
In recent years, with the effect of the climate analysis and weather forecasting are accepted as one of the most important natural topics. Adana, located in the Mediterranean region, has a Mediterranean climate and is one of Turkey's roughest cities. Because of its Mediterranean climate, different weather conditions are observed in every season of the year in Adana. In this study, it is aimed to make a monthly average weather forecast for the next 12 months in Adana. The dataset used in this study was collected from two different weather websites since January 2019 and includes Adana's daily weather values including maximum temperature, minimum temperature, humidity, and wind speed. Weather forecasting was performed on this dataset using linear regression method with Weka data mining tool. While predicting the weather using the linear regression method, it is provided that the relationship between the variables was calculated, and an equation between them was established to make predictions. According to the results obtained, it has been observed that good results cannot be obtained when day-based forecasting is requested, but almost the same results are obtained when the monthly average are requested. In addition, the results of the estimations were approximate, and as a result of the study, the Correlation Coefficient was found to be 0.9832.Additionally, it was concluded that while forecasting the weather, linear regression, which is a data mining technique, yields positive results when the forecast is made with the monthly average.

Keywords: Adana, Climate Analysis, Data Mining, Linear Regression, Weather Forecasting

1. INTRODUCTION

The weather is one of the most influential factors in life. The weather is a situation that can change from one day to the next. Weather conditions vary according to the climatic characteristics of the locations. Climate is the average number of meteorological events over a long period of time. Climate analysis is used for a better understanding of the climate and to see what weather conditions the location has by looking at the average daily temperature, weather events, and other features such as humidity and pressure. Weather forecasting is the forecasting of future weather conditions.

Adana is a province in Turkey located in the Mediterranean region and is the 6th most populous city. According to the latest data, it has a population of 2,258,718 and an area of the province is 13,844 km^2. Adana, a major city in Turkey, is one of the country's leading agricultural, commercial and cultural centers.

Adana is a province with Mediterranean climatic characteristics. While summers are hot and dry, winters are warm and rainy. Air flows from the sea and the Taurus Mountains to Çukurova, which is

*Corresponding author, E-mail address: mkaya@atu.edu.tr

Current Trends in Natural Sciences
Vol. 11, Issue 21, pp. 304-314, 2022
https://doi.org/10.47068/ctns.2022.v11i21.033
Current Trends in Natural Sciences (on-line)
ISSN: 2284-953X
ISSN-L: 2284-9521
Current Trends in Natural Sciences (CD-Rom)
ISSN: 2284-9521
ISSN-L: 2284-9521

a low-pressure center in summer. In this case, while humid air comes from sea on one hand, humidity increases due to dams and irrigation of the plain on the other hand. Thus, a humid hot air is observed in summer. Rainfall in Adana is usually caused by the encounter of slope rains and mobile air masses. Precipitation occurs naturally in the mountainous part. In addition, while snow is rarely seen on the plain, snowfall in the mountains starts early and can stay for months. Despite the fact that the plain is warm, the climatic conditions in the province vary greatly according to the altitude and surface forms. In this study, forecasting was applied with WEKA. In this section, forecasting applications made with Weka are examined.

In the study of Duran (2005), it is observed that meteorological weather forecasts can be made. In the dataset used, a 40-day dataset was used for the first time. The normalized values are given as the ANN and ANFIS input values. Thus, the weather forecast for the next day can be realized. When a random 100-day data forecast was made between 14.02.2003 and 09.02.2014, it was seen that while the rule base of the ANFIS classifier was 16, the 54-day forecast was correct. According to this study, thanks to the program developed in MATLAB, much more accurate predictions were made over 50%.

In the study of Kumar (2013), a decision tree was used to predict events such as fog, rain and thunder, because the decision tree evaluation can be digitized and is simple to use. In this study, WEKA was used to facilitate weather forecasting. The data used were collected over one year. In their study, they made predictions considering three parameters. As a result, 46 out of 72 test samples were classified appropriately, giving a .0584 kappa statistic. In addition, the authors mentioned that software equipped with a decision tree can provide artificial intelligence to the machine.

In the study of Jayasingh et al. (2016), three various soft computing methods, namely, the multi-layer perceptions (MLP), support vector machine, and J48 decision tree were used in weather forecasting. It shows a comparison between Delhi's time-series data collected over five years and fed into the three models. After training the three models, the results were compared and it was determined that the J48 decision tree consistently outperformed the others. The J48 decision tree model can be developed to better predict rain, hurricanes, storms, and other natural disasters in the near future, as well as humans and domestic animals. It is believed to have saved hundreds of thousands of lives.

In the study of Sagaltici et al. (2018), the loss of solar radiation was estimated using data obtained from the GAPYENEV center. In this study, using artificial neural networks and other data mining techniques, the estimation of the missing data in the radiation values obtained from the data were estimated, and the most appropriate method was evaluated by comparing the results. According to the results, the nearest neighbor method was the most successful in estimating the missing data.

In the study of Cinaroglu and Unutulmaz (2019), forecasts were made using WEKA. In their study, data mining was found to be suitable for analyzing meteorological data. The accuracy rates of the modeling algorithms for each feature are shown with different feature selection methods. The key features of the forecast model for each precipitation event are also presented. For example, when estimation algorithms are used, it has been observed that the highest accuracy rate is achieved with the "Wrapper" feature selection method.

In the study of Abinaya and Janani (2020), rain forecasts were made using the WEKA tool. In this study, four classification algorithms in WEKA, support vector machine, decision tree, k-nearest neighbor, and naïve bayes were used for rain forecasting. In this study, when the results are examined, it is seen that the decision tree algorithm that gives the best result for forecasting.

Today, forecasting has become a necessity in many areas such as weather, economy, agricultural, and military relations. In this study, weather forecasting is a time series problem. Although a one-dimensional dataset can be used in studies on weather conditions, a multi-dimensional dataset can

*Corresponding author, E-mail address: mkaya@atu.edu.tr

Current Trends in Natural Sciences
Vol. 11, Issue 21, pp. 304-314, 2022
https://doi.org/10.47068/ctns.2022.v11i21.033
Current Trends in Natural Sciences (on-line)
ISSN: 2284-953X
ISSN-L: 2284-9521
Current Trends in Natural Sciences (CD-Rom)
ISSN: 2284-9521
ISSN-L: 2284-9521

also be used. In general, linear or non-linear data mining methods are used by approaching this kind of problem solving as solving the time series problem. Because weather forecasting is a time-series problem, forecasts may be for the near future or the future.

2. MATERIALS AND METHODS

2.1. Dataset

In this study, the dataset containing the weather data from Adana province was used. Data in the dataset were obtained from two separate weather websites. The websites where the data is obtained were www.tr.freemeteo.com and www.havaturkiye.com. The daily maximum temperature value, daily minimum temperature, maximum speed of constant wind, total amount of precipitation, dew point, snow depth, pressure and description were taken from www.tr.freemeteo.com and the rest of the data were taken from the other site. While these data were obtained from two separate sites, attention was paid to the selection of the websites, ensuring the same air temperature and similarity to other available features. Data extraction from the web pages was performed manually. The data were collected from January 2019 and up to November 2021. There are 11 attributes in the 3-year datasets obtained. A total of 1042 days of data were obtained. In these datasets, there are the daily maximum temperature value, daily minimum temperature, maximum speed of constant wind, total amount of precipitation, dew point, humidity, UV degree, snow depth, pressure, and descriptive values.

2.2. WEKA

The University of Waikato in New Zealand developed the Waikato Environment for Knowledge Analysis (WEKA) program, which is an open source data mining program. It was developed on Java language. In addition, with WEKA, which has a completely modular design, many operations can be performed for requirements, such as data preprocessing, data visualization and business intelligence applications with machine learning algorithms. In WEKA, algorithms can be applied directly to a dataset or can be called from written Java code.
It offers graphical user interfaces. WEKA is released under the GNU General Public License and is free to use. Algorithms that cannot be found in WEKA are loaded later in the settings section. In addition, special documentation and writing rules for development with Java are also presented by the same research group. After WEKA software is installed, it comes with *.arff* extension support. The files must be converted to this format for the data to be analyzed. However, the files with the .csv extension can be opened in WEKA, and it is more correct to use the *.arff* file format. In addition, when the software is first opened, the .csv file can be converted into *.arff* file by selecting "ArffViewer" from the "tools" option.

2.3. Data Mining

With the advancement of technology and its ability to be used everywhere, many jobs that were previously physical are now performed from places such as computers, mobile phones, and tablets. Because of most transactions with these electronic devices, some data are accumulated on the other side. These data may be meaningless if they are unprocessed. By contrast, data mining can be defined as obtaining previously unknown, valid and applicable information from data stacks through a dynamic process.
It is not convenient to use traditional statistical methods when analyzing big data. As a result, special methods for processing and analyzing the data are required. To satisfy these needs, data mining

*Corresponding author, E-mail address: mkaya@atu.edu.tr

Current Trends in Natural Sciences
Vol. 11, Issue 21, pp. 304-314, 2022
https://doi.org/10.47068/ctns.2022.v11i21.033
Current Trends in Natural Sciences (on-line)
ISSN: 2284-953X
ISSN-L: 2284-9521
Current Trends in Natural Sciences (CD-Rom)
ISSN: 2284-9521
ISSN-L: 2284-9521

techniques have arisen. Although data mining can be viewed as a series of statistical methods, it can be performed using mathematical disciplines, modeling techniques, database technology, and various computer programs.

There are three commonly used methods for data mining analysis. These methods include classification, clustering and association rules (ARs).

In the classification method, data can be divided into special or general categories, and each category can be assigned as belonging to a class. Practically deciding on the processes can be used as a classification problem. A relationship can be established between the classification method and the values of the classified data and other classified data. If there are examples in the class, it is supervised learning, but if there are no known examples for the class, the classification process is unsupervised learning.

The purpose of the clustering method is to examine the similarity of the data to each other according to their values and group them accordingly. Similar data were collected from one cluster, and dissimilar data were collected from another cluster.

Association rules are data mining methods that determine the co-occurrence of events. Association rules, together with these methods, reveal rules with certain possibilities for the realization of the situation.

2.3.1. Data Preprocessing Methods

For data mining to be applied correctly, it is important that the data are of good quality. The data must comply with the strict criteria to be applicable. To increase the reliability of the study, the data obtained must first be preprocessed. If data are not preprocessed, incorrect data will result in an incorrect output. In addition, data preprocessing is time consuming. The necessity of preprocessing large amounts of data has made effective techniques for automatic data preprocessing important. As a result, data preprocessing is used to make more meaningful inferences from the given dataset, to prevent data that will cause errors in the analysis of the data, and to understand the data and achieve meaningful data analysis.

In many data mining applications, more than one data preprocessing technique may be required. To obtain accurate results, it is important to determine the methods to apply first. The basic techniques of data preprocessing are data cleaning, data merging, data conversion, and data reduction.

Data cleaning requires operations such as completing the missing values, identifying outliers, correcting the data called noisy data, and eliminating the inconsistencies between the data in the dataset to make the data fit for meaningful data analysis. Different methods can be used to fill in missing data values. Records with missing values were removed from the dataset. The average of the values of the variable can be used instead of the missing values, or the mean of the variable can be used for the values belonging to the same class. Additionally, the most appropriate value can be used based on the available data.

Generally, data from different databases may need to be combined for data mining. On the other hand, data merging involves joining data from these different places into a single database. When merging data, schema merge errors occur in the database or redundancy may occur if a variable is derived from another table. These redundancies can be investigated using correlation analysis.

The data used in the data conversion technique were converted into a form suitable for data mining. In this technique, one or more different operations can be used for correction, merging, generalization and normalization. Data normalization is the most widely used process.

*Corresponding author, E-mail address: mkaya@atu.edu.tr

Current Trends in Natural Sciences
Vol. 11, Issue 21, pp. 304-314, 2022
https://doi.org/10.47068/ctns.2022.v11i21.033
Current Trends in Natural Sciences (on-line)
ISSN: 2284-953X
ISSN-L: 2284-9521
Current Trends in Natural Sciences (CD-Rom)
ISSN: 2284-9521
ISSN-L: 2284-9521

Data reduction methods are applied to obtain a reduced sample of the dataset with a smaller volume. Thus, more effective results are obtained when data mining techniques are applied to the reduced dataset of these processes.

In this study, firstly, data merging was used while preprocessing the data. Data from two separate web sites were combined. The missing data in the dataset are filled in the blank values by taking the monthly average over the excel file. The attributes that were found to be contradictory to each other and that negatively affected the results were deleted. Thus, more reliable results were obtained with the obtained dataset are applied to the reduced dataset. The flowchart of data preprocessing methods used in this study is shown in Figure 1.

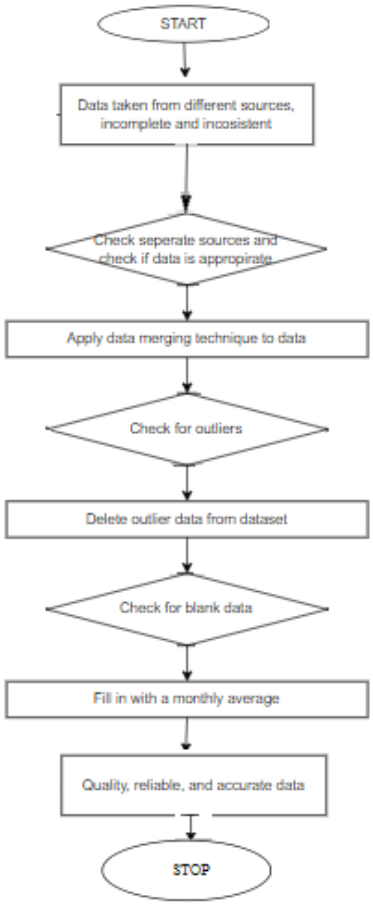

Figure 1. Flowchart of data preprocessing methods

2.3.2. Regression

Regression is a statistical method used to examine the relationships between the variables in a dataset. In this method, the relationship between a dependent variable and one or more independent variables was determined. Thus, if the independent variables change, it is necessary to understand how the dependent variable changes are applied to the reduced dataset.

Linear regression is the most basic technique used to model the relationship between different variables. Linear regression attempts to create an equation that can be used to estimate the value of

*Corresponding author, E-mail address: mkaya@atu.edu.tr

Current Trends in Natural Sciences
Vol. 11, Issue 21, pp. 304-314, 2022
https://doi.org/10.47068/ctns.2022.v11i21.033
Current Trends in Natural Sciences (on-line)
ISSN: 2284-953X
ISSN-L: 2284-9521
Current Trends in Natural Sciences (CD-Rom)
ISSN: 2284-9521
ISSN-L: 2284-9521

one variable relative to the value of the other by examining the relationship between two different variables. This is usually the first step in a complex analysis.
Linear regression was analyzed in two parts. These parts are referred to as simple and multiple regressions. Linear regression was used to determine the best-fit line over a set of points. Here, using a straight line, a relationship was established between the dependent variable and one or more independent variables. While "*x*" variable is considered as independent variable in Equation 1 and 2, "*y*" variable is considered as dependent variable. Equation 1 is used when there is only one independent variable, and Equation 2 is used when there are multiple independent variables. The "*b*" value used in these equations represents the bias point, that is, the intersection point. "*w*" represents the weight of that independent variable, that is, the regression coefficient, and the "*e*" value indicates the error value.

$$y = b + w_1 x_1 + e \quad (1)$$

$$y = b + w_1 x_1 + w_2 x_2 + w_3 x_3 + e \quad (2)$$

2.3.3. Correlation Coefficient

The correlation coefficient is a measure of the dependence between the two random variables. The value of the correlation coefficient varied between -1 and +1. A value of"0" indicates that there is no relationship between the variables. The greater the relationship, the closer the value is to one. Negative values indicate that the relationship between the two variables is negative, while positive values indicate positive values. The definition of Pearson's product-moment correlation coefficient "($\rho_{x,y}$)" between two independent variables X and Y, with mathematical expectation values, "μ_x"and "μ_y" and standard deviations "σ_x"and "σ_y"is shown in Equation 3.

$$\rho_{x,y} = (Cov(X,Y)/\sigma_x\sigma_y) = (E((X-\mu_x)(Y-\mu_y))/(\sigma_x\sigma_y) \quad (3)$$

3. RESULTS AND DISCUSSIONS

In this study, two studies were conducted using WEKA software. In the first study, WEKA was used to better understand the climatic conditions experienced in Adana. In the second study, it was used to forecast the weather conditions in Adana.

3.1. Adana Climate

Adana is a province with Mediterranean climatic characteristics. Summers in Adana are hot and dry, while Adana's winters are warm and rainy. Between 2019-2021 years, the highest temperature in Adana was 45 °C during the day and -4 °C at night.

*Corresponding author, E-mail address: mkaya@atu.edu.tr

Current Trends in Natural Sciences
Vol. 11, Issue 21, pp. 304-314, 2022
https://doi.org/10.47068/ctns.2022.v11i21.033
Current Trends in Natural Sciences (on-line)
ISSN: 2284-953X
ISSN-L: 2284-9521
Current Trends in Natural Sciences (CD-Rom)
ISSN: 2284-9521
ISSN-L: 2284-9521

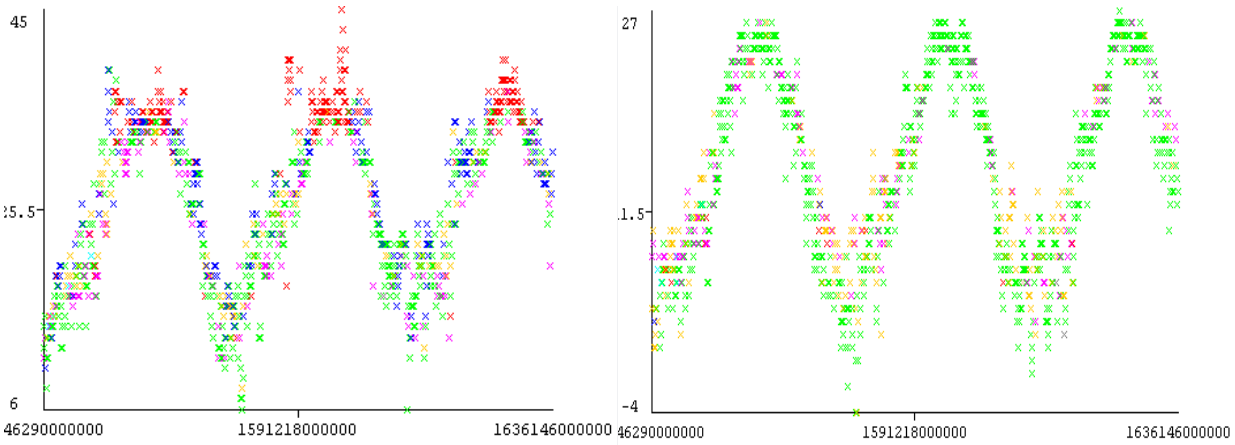

Figure 2. Daily maximum temperature of Adana and daily minimum temperature of Adana

On the hottest day in Adana, this date is 03.09.2020. The highest and lowest temperatures on this date were 45 °C and 25 °C, respectively as shown in Figure 2. This is a sunny day as shown in Figure 3. The wind intensity was 22 *Km/h* and the humidity was 57. The coldest day was on 11.02.2020. From the data, the maximum temperature reached during the day was 6 °C, and the minimum temperature was -4 °C. The maximum constant wind speed was 15 *Km/h* and the humidity was 58.5. The day was rainy as shown in Figure 4.

```
Plot : Master Plot
Instance: 612
                                timestamp : 1.5990804E12
                Daily_minimum_temperature : 25.0
                Daily_maximum_temperature : 45.0
             Maximum_speed_of_constant_wind : 22.0
Total_amount_of_precipitation_per_day : 0.0
                               Dew_Point : 66.7
                               UV_Degree : 9.5
                                Humidity : 57.0
                              Snow_Depth : 0.0
                                Pressure : 29.7
                             Description : sunny
```

Figure 3. Hottest day on Adana

```
Plot : Master Plot
Instance: 407
                                timestamp : 1.5813684E12
                Daily_minimum_temperature : -4.0
                Daily_maximum_temperature : 6.0
             Maximum_speed_of_constant_wind : 15.0
Total_amount_of_precipitation_per_day : 0.0
                               Dew_Point : 15.9
                               UV_Degree : 3.0
                                Humidity : 58.5
                              Snow_Depth : 0.0
                                Pressure : 30.1
                             Description : rainy
```

Figure 4. Coldest day on Adana

The maximum snow thickness per day in Adana is 7.1 cm and today is 24.12.2019. The maximum daily precipitation was 117.8 *cm*, and the day with the maximum precipitation was 25.12.2019. It was cloudy for 516 days over 3 years in Adana as shown in Figure 5 and Figure 6. The most common weather event in Adana over the three years was cloudy. The next most common weather event in Adana was sunny for 203 days.

*Corresponding author, E-mail address: mkaya@atu.edu.tr

Current Trends in Natural Sciences
Vol. 11, Issue 21, pp. 304-314, 2022
https://doi.org/10.47068/ctns.2022.v11i21.033
Current Trends in Natural Sciences (on-line)
ISSN: 2284-953X
ISSN-L: 2284-9521
Current Trends in Natural Sciences (CD-Rom)
ISSN: 2284-9521
ISSN-L: 2284-9521

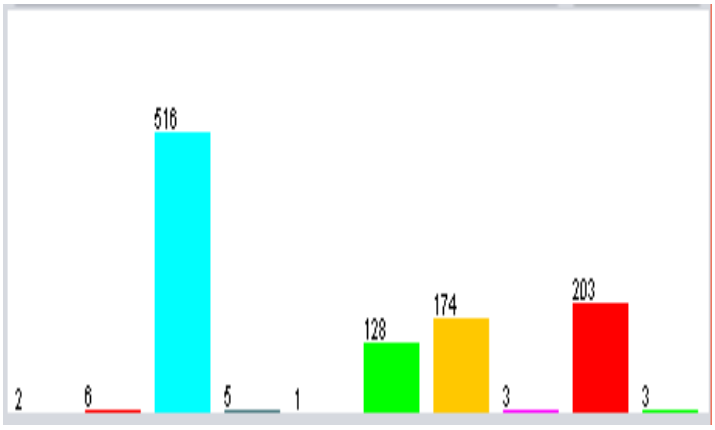

Figure 5. Distribution weather condition on Adana

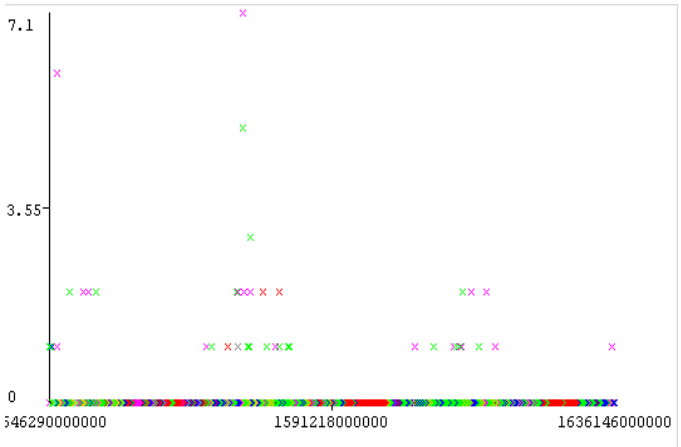

Figure 6. Daily amount of precipitation on Adana

3.2. Weather Forecasting on Adana

Weather forecasting is an important topic in general. Weather, which affects the lives of many people, is a factor that can affect a significant part of daily life. Therefore, it is important to predict the weather conditions in advance.

As mentioned above, weather forecasting is a time-series problem. In this study, using WEKA, forecasts were made on the dataset of Adana's 3-year weather data. It was made from the forecast section at the explorer interface in WEKA. These estimations were made by choosing linear regression, which is a data mining technique, from the forecast part of WEKA. When estimating with linear regression, the correlation coefficient obtained was found as 0.9832. This value indicates that the data have a positive relationship with each other.

If the forecasts were chosen on a daily basis, repeating values would begin to be observed after the first period. However, in this estimation process, accurate results can be obtained as shown in Figure 7 if it is desired to be obtained for only one day. In this study, on the other hand, estimations were made on a monthly basis because the closest results were obtained when it was desired to make a monthly estimation for each month as shown in Figure 8. As shown in Figure 9, the relationship between the predicted and actual values was observed for each attribute. As can be seen from the closeness of the values, there is not much difference between them, indicating that more probable

*Corresponding author, E-mail address: mkaya@atu.edu.tr

Current Trends in Natural Sciences
Vol. 11, Issue 21, pp. 304-314, 2022
https://doi.org/10.47068/ctns.2022.v11i21.033
Current Trends in Natural Sciences (on-line)
ISSN: 2284-953X
ISSN-L: 2284-9521
Current Trends in Natural Sciences (CD-Rom)
ISSN: 2284-9521
ISSN-L: 2284-9521

results were obtained. As can be seen in the figures below, it is correct to make monthly forecasts for Adana with WEKA.

30.10.2021	15	29
31.10.2021	18	26
01.11.2021	18	25
02.11.2021	17	20
03.11.2021	15	25
04.11.2021	14	27
05.11.2021	12	29
06.11.2021	13	31
07.11.2021*	14.958	29.5894
08.11.2021*	16.1881	30.2011
09.11.2021*	16.5214	29.7372
10.11.2021*	16.959	29.3328
11.11.2021*	16.2694	29.2808

Figure 7. Day-based forecast results

01.12.2021*	7.9433	21.5799	13.0395	40.0441
01.01.2022*	7.0005	18.0892	20.1093	41.301
01.02.2022*	6.462	21.4083	14.4789	40.5107
01.03.2022*	11.4042	23.615	23.6872	53.4708
01.04.2022*	11.4639	25.1572	18.7589	43.4952
01.05.2022*	19.743	32.3615	22.9388	71.5926
01.06.2022*	22.6019	34.6747	19.2423	65.1646
01.07.2022*	25.2749	34.5292	25.9125	84.9372
01.08.2022*	22.2751	36.0086	19.243	61.4865
01.09.2022*	28.0014	38.4867	22.674	91.9054
01.10.2022*	14.5677	31.4029	13.8916	40.7453
01.11.2022*	16.0093	25.0112	22.1336	66.4047

Figure 8. Monthly-based forecast results

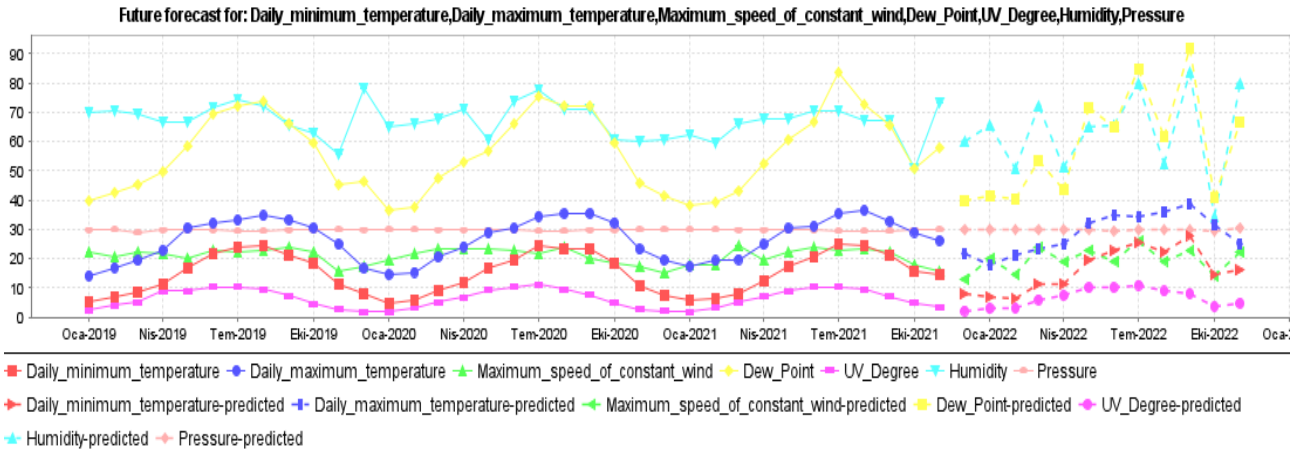

Figure 9. Graphic of monthly-based forecast

*Corresponding author, E-mail address: mkaya@atu.edu.tr

Current Trends in Natural Sciences
Vol. 11, Issue 21, pp. 304-314, 2022
https://doi.org/10.47068/ctns.2022.v11i21.033
Current Trends in Natural Sciences (on-line)
ISSN: 2284-953X
ISSN-L: 2284-9521
Current Trends in Natural Sciences (CD-Rom)
ISSN: 2284-9521
ISSN-L: 2284-9521

4. CONCLUSIONS

The aim of this study was to better understand the climate of Adana Province and to make a weather forecast for Adana. To carry out this study, the weather dataset of Adana province was used, and the WEKA program was used to make predictions with these analyses.

In the comparison made to better understand the climatic conditions of Adana, it has been seen more clearly how hotter Adana is, and how much less snowfall and rain it receives. Although there have been mostly cloudy days in both provinces for the three years, it has been seen as a result of the analysis that Adana is warmer.

The linear regression technique was used in the estimation made with WEKA. It has been observed that good results cannot be obtained when day-based forecasting is desired. However, similar results were obtained when the monthly average was desired. Thus, it has been seen that the linear regression technique used while forecasting the weather in WEKA gives positive results when the forecast is made with a monthly average.

In some studies, when estimation is desired, the classification method has been used, and it has been observed that the most appropriate method is the decision tree algorithm. Dursun (2005) performed predictions using ANFIS and ANN. Successful results were obtained in this study, which was conducted in the MATLAB environment. In this study, regression was used instead of classification, and good predictions were made for the future.

In future studies, it will be concluded that linear regression is more appropriate when the regression method is used for estimation. In addition, the time series in the dataset should be considered when making the forecast, and it is very important that the data are related to each other. WEKA's uncomplicated structure makes it easy to use, and it has many features for data mining. Thus, a more comfortable working space for data mining is required.

5. REFERENCES

Abinaya, P. L., Janani, N. (2012). Rainfall Forecasting Using Weka Data Mining Tool.*International Research Journal of Engineering and Technology (IRJET)*.7(3), 5346-5348.

Alan, M. A. (2012). Veri madenciliği ve lisansüstü öğrenci verileri üzerine bir uygulama [An application on data mining and graduate student data].*Dumlupınar University Journal of Social Sciences*, 33.

Anonymous, İlimizi tanıyalım. [Let's get to know our city.](n.d.). Türkiye Cumhuriyeti Çevre, Şehircilik ve İklim Değişikliğihttps://adana.csb.gov.tr/ilimizi-taniyalim-i-1222

Anonymous, Correlation (n.d.). https://tr.wikipedia.org

Cebeci, Y. E. (2019). Hava durumu tahmini için veri madenciliği tabanlı bir model geliştirilmesi [Developing a data mining-based model for weather forecasting] (Unpublished master's thesis). University of Istanbul Teknik, Istanbul.

Cinaroglu, E., Unutulmaz, O. (2019). A data mining application of local weather forecast for Kayseri Erkilet Airport. *Politeknik Dergisi*, 22(1), 103-113.

Dursun, Ö. O., (2005). Meteorolojik verilerin akıllı yöntemlerle sınıflandırılması [Classification of meteorological data with smart methods](Unpublished master's thesis).Fırat University, Elazığ.

Jayasingh, S. K., Mantri, J. K., Gahan, P. (2016). Comparison between J48 Decision Tree, SVM and MLP in Weather Forecasting. *International Journal of Computer Science and Engineering*, 3(11), 42-47.

Kumar, R. (2013). Decision tree for the weather forecasting. *International Journal of Computer Applications*, 76(2), 31-34.

Oğuzlar, A., (2012). Veri Ön İşleme [Data Preprocessing].*Journal of Erciyes University Faculty of Economics andAdministrative Sciences*, 33, 66-67.

Sagaltici, D., Alay, F. D., Efil, C., İlhan, N, (2018). Veri Madenciliği Yöntemleri İle Meteorolojik Verilerden Kayıp Güneş Işınım Değerlerinin Tahmini [Loss from Meteorological Data with Data Mining Methods Estimation of Solar Irradiance Values].*Harran University Journal of Engineering*, 2, 49-53.Taş, B.,

*Corresponding author, E-mail address: mkaya@atu.edu.tr

Current Trends in Natural Sciences
Vol. 11, Issue 21, pp. 304-314, 2022
https://doi.org/10.47068/ctns.2022.v11i21.033
Current Trends in Natural Sciences (on-line)
ISSN: 2284-953X
ISSN-L: 2284-9521
Current Trends in Natural Sciences (CD-Rom)
ISSN: 2284-9521
ISSN-L: 2284-9521

Taş, B.,. (2020) Doğrusal Regresyon [Linear Regression], Retrieved from https://bernatas.medium.com/do%C4%9Frusal-regresyon-linear-regression-8f562c19aadf

Topaloğlu, F., (2007) Veri madenciliği ile meteorolojik parametrelerin analizi ve zirai meteoroloji haritasının çıkarılması[Analysis of meteorological parameters and agricultural meteorological mapping with data mining] (Unpublished master's thesis).Fırat University, Elazığ.

Yıldız, M., & Şeker, S. (2016). Data Mining Tools. *YBS Encyclopedia*, 3(4).

*Corresponding author, E-mail address: mkaya@atu.edu.tr