EGE AKADEMİK BAKIŞ / EGE ACADEMIC REVIEW

Cilt 18. Sayı 4. Ekim 2018
SS. 579-588
DOI: 10.21121/eab.2018442989
Received: 27.11.2017 • Accepted: 20.07.2018

Forecasting the Direction of Agricultural Commodity Price Index through ANN, SVM and Decision Tree: Evidence from Raisin

YSA, DVM ve Karar Ağacı ile Tarımsal Emtiaların Fiyat Endekslerinin Tahminlenmesi: Kuru Üzüm Örneği

Burcu Akın[1] *Orcid No: 0000-0001-6665-3213*
İkbal Ece Dizbay[2] *Orcid No: 0000-0003-2431-4269*
Şevkinaz Gümüşoğlu[3] *Orcid No: 0000-0001-8442-8167*
Erçin Güdücü[4] *Orcid No: 0000-0001-6497-9068*

Abstract

To be able to make appropriate actions during buying, selling or holding decisions, economic actors need accurate commodity price forecasts. This study focuses on forecasting raisin price by using predetermined volatile variables. Therefore, we seek for answers of three main questions. Do the social & political issues effect raisin price in countries that have internal disturbance? By using volatile variables, can we represent or predict price index thoroughly? Lastly, which method has the best prediction performance; Artificial Neural Networks (ANN), Decision Tree or Support Vector Machine (SVM)? In accordance with these purposes, ANN, decision tree and SVM methods are implemented for proposed model and their prediction performances are compared. Experimental results showed that accuracy performance of SVM method was found significantly better than ANN method and decision tree.

Keywords

Commodity market • Artificial neural networks • Decision tree • Support vector machines • Social & political issues

Jel Codes

Q02 • C53

Öz

Emtia fiyat endekslerinin başarılı bir şekilde tahminlenmesi ekonomik aktörlere doğru alım satım kararları verebilmeleri için fayda sağlamaktadır. Türkiye'deki ticaret borsalarında işlem gören tarımsal ürünlerden biri olan kuru üzüm fiyatlarının oynak değişkenler kullanılarak tahminlenmesinin incelendiği çalışmada üç temel soru üzerinde durulmuştur. İç karışıklığın yüksek olduğu ülkelerde sosyal ve politik olaylar kuru üzüm fiyatlarını etkiler mi? Oynaklığı yüksek olan değişkenler kullanılarak kuru üzüm fiyat endeksleri tahminlenebilir mi? Son olarak, bu tip bir çalışmada Yapay Sinir Ağları (YSA), Karar Ağacı ve Destek Vektör Makineleri (DVM) yöntemlerinden hangisinin tahmin performansı daha yüksektir? Bu amaçla oluşturulan tahmin modeline YSA, KA ve DVM yöntemleri uygulanmış ve yöntemlerin tahmin performansları karşılaştırılmıştır. Uygulama sonuçları, oynak değişkenler ile sosyal ve politik olayların kuru üzüm fiyatlarının tahminlenmesinde kullanılabileceğini ve ilgili modelde DVM yönteminin en yüksek doğruluk oranını verdiğini göstermiştir.

Anahtar kelimeler

Emtia piyasası • Yapay sinir ağları • Karar ağacı • Destek vektör makineleri • Sosyal ve politik olaylar

Jel Kodları

Q02 • C53

Accurate commodity price forecasting provides visible effect on domestic and global economies. Producers, importers, exporters, traders and governments need to get reliable information about the agricultural commodity price trends or future forecasts. This knowledge avails them to provide against the fluctuations in price so they can make appropriate actions during buying, selling or holding decision. Especially for exportable agricultural products, it is important to determine product prices and yearly production amount as early as possible. That's because the purchase agreements are placed generally four to six months earlier and if the price hasn't determined at this point, it is possible to lose customers due to substitutable structure of such products.

Price index forecasting studies utilize macroeconomic indicators as GDP, interest rates or past prices while predicting. However, macroeconomic indicators of the countries that have uncertainty and usually cannot represent the real picture. In such countries due to large amounts of social and political incidents, they usually touch the short-run memories. Accordingly, in this study we prefer to use most volatile variables as explanatories that respond rapidly to these incidents. Main issue in our study is that, can volatile variables such as gold price, stock

[1] Burcu Akın, (Phd.), Independent Researcher.
[2] **Correspondence:** İkbal Ece Dizbay, (Lecturer Phd.), Yaşar University, Logistics Program, Izmir.
[3] Şevkinaz Gümüşoğlu, (Prof. PhD), Yaşar University, Dept. of Business Adm., Izmir.
[4] Erçin Güdücü, (Phd.), Izmir Commodity Exchange, Izmir.

Atıf/Citation: Akın, B., Dizbay, İ. E., Gumusoglu, S. & Guducu, E. (2018). Forecasting the direction of agricultural commodity price index through ANN, SVM and decision tree: Evidence from raisin. *Ege Academic Review*, 18(4), 579-588.

returns, exchange rates, oil price and social & political incidents represent or predict price index thoroughly. Numerous commodity and stock price studies reveal that methods as artificial neural networks and support vector machine outperformed time series forecasting models as ARIMA or linear regression (Kohzadi, Bahman, & K, 1996) Therefore, Decision Tree, Artificial Neural Networks (ANN) and Support Vector Machine (SVM) will be implemented for proposed model.

This paper focuses on forecasting the raisin price index of Turkey, which is one of the trading commodities. As shown in Figure 1 (2017) raisin is selected as among agricultural commodities since raisin is the second most important agricultural product after nuts and Turkey is the world's biggest raisin exporter with about 30% share. 85% of raisin produced in Turkey, is for export and Turkey has about 500 million dollars foreign exchange earnings depending upon this export volume.

Furthermore, raisin is a valuable agricultural product especially for about one hundred thousand producers and their families live in Aegean Region. These farmers produce about three hundred thousand tons of raisins in a year and this is equal to nearly 24% of Worlds raisin production as shown in Figure 2 (2017).

This study focused on three questions. First, do the social & political issues effect raisin price in countries that have internal disturbance? Second, can most volatile indicators such as gold price, stock returns, exchange rates and oil price help to forecast movement of raisin price? Third, which one has the best forecasting performance as artificial neural networks, decision tree or support vector machine? Therefore, directional prediction model that aims to forecast raisin price direction according to gold price, stock returns, exchange rates oil price and social & political issues introduces to decide timing of buying and selling actions. Decision Tree, ANN and SVM methods are established for the prediction and their performances are compared. To measure the performance of the methods, accuracy and f-measure are used. Detailed information about the data exists on Section 2. Predictive models appear on Section 3. Experimental results are discussed in Section 4 and Section 5 includes conclusion.

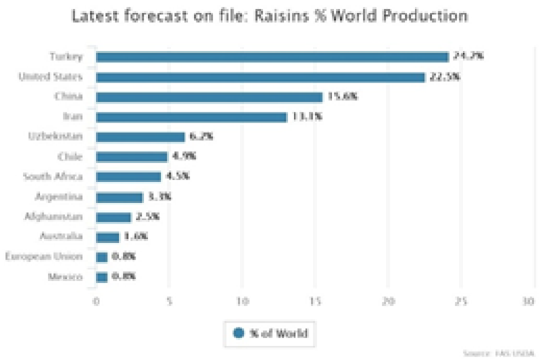

Figure 1: *Raisins % World production*

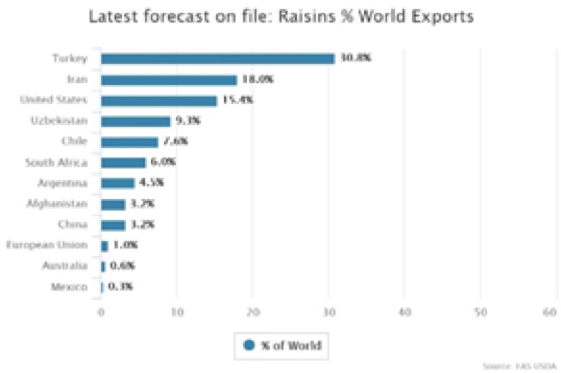

Figure 2: *Raisins % World Exports*

Related Studies

Recent research studies showed that relatively many empirical works have been undertaken on forecasting commodity price. Various forecasting techniques and explanatories have been used in order to predict commodities' price movement and provide the most accurate forecast. However, the studies that concern on political instabilities, gold price, stock returns, exchange rates and oil price are limited. Therefore, literature review includes researches that forecast commodity prices with ANN models and use various inputs apart from past prices.

Kohzadi et al. (1996) proposed an artificial neural network and time series model for predicting the upcoming price trend of live cattle and wheat. The models that used monthly data are compared according to the absolute mean error (AME), the mean absolute percentage error (MAPE) and the mean square error (MSE). The results revealed that a non-linear model as ANN is superior to ARIMA models for time series forecasting. Parisi et al. (2008) focused on estimating the upcoming trend of the price on gold market. Weekly datasets are used as train and holdout datasets of rolling neural network and recursive neural network models. The results showed that while rolling networks exceed the ARIMA model, recursive networks resulted almost equally to ARIMA. Hu et al. (2012) suggested neural network models such as multilayer perceptron BPNN, the Elman recurrent neural network (ERNN) and the recurrent fuzzy neural network (RFNN) in order to forecast crude oil futures prices. The results of comparison stated that RFNN is best-performed model and BPNN and ERNN were almost equal. Jammazi and Aloui (2012) worked on optimizing ANN architecture to forecast crude oil price, innovatively. Ultimately, they reached network architecture that performs better.

Haidar and Wolff (2009) used ANN to forecast the crude oil price. Their main objective was investigating the effects of data smoothing techniques on the results on ANN. In fact, they proved that some of those techniques are effective in improving the accuracy. Ye et al. (2002) present a wavelet neural network model in order to forecast the crude oil spot price. They used OECD industrial petroleum inventory levels as independent variable. Besides a significant relationship was appeared. Azadeh et al. (2012) suggested a flexible model based on ANN and fuzzy regression to estimate oil price. In their study, applicability of the model to complex, non-linear and uncertain datasets were proved. Li et al. (2010) focused on short-term price forecasting of tomato with ANN model. In addition, they compared the experimental results with time series model ARIMA. ANN model's better performance on forecasting the tomato price was resulted. Zou et al (2007) investigated a study that compare the directional forecasting performance of ANN, ARIMA and the linear combination models. The results showed that ANN model's performance is best and model can be used as an alternative forecasting model.

The price forecasting capabilities of ANN and ARIMA are compared by Jha and Sinha (2014). The study revealed that the neural network models have clear advantage over linear models for predicting the direction of monthly price change for different series. Chen, Rogoff, and Rossi (2010) worked on predictability in an aggregate commodity price index, which involves more than forty traded products. In addition to evidence that exchange rates predict commodity prices both in-sample and out-of-sample, they used five commodity currencies. Groen and Pesenti (2011) investigated the spot price predicting performances of commodity exchange rates and a broad cross-section of macroeconomic variables gains random walk or autoregressive. They determined that using commodity exchange rates and macro-economic variables cannot avail against random walk or autoregressive.

Hong and Yogo (2012) found evidence of limited in-sample predictability of returns on commodity futures. Gargano and Timmermann (2014) urged upon predictability of commodity price by means of macroeconomic and financial variables. They concluded that predictability of commodity price varies across economic states and commodity prices are most predictable during recessions. Sujit and Kumar (2011) investigated the dynamic relationship among exchange rate, stock returns, gold price and oil price. They established cointegration and Vector Auto Regression (VAR) models. The study revealed that exchange rate is highly affected by changes in other variables except stock market.

Samanta and Zadeh (2012) suggested a model to examine the comovements of several macro-variables in the world economy as gold and stock price, real exchange rate for dollar and the crude oil price. The results suggested that the possibility of cointegration among these variables is indicating comovements. Wang and Chueh (2013) aimed to examine the short and long-term dynamic interactions between oil prices, gold prices etc. They established that international gold and crude oil prices effect each other and have influence on interest rates. Malliaris and Malliaris (2013) proposed a model that explores interconnection of price behaviors of oil, gold and the euro using time series and neural network methods. Results showed that, although the markets for oil, gold and the euro are efficient, their relationships are limited. Besides these studies about forecasting the movement of stock price and comparing the techniques are existing on Chen et al. (2003), Ahmed (2008), Kara et al. (2011) and Yakut and Gemici (2017).

The contributions of this paper to related literature are, working with social and political issues (SPI) as an input to predict raisin price index, which is not used before and comparison of ANN, SVM and decision tree methods while forecasting the commodity price index.

Research Data

This section defines the data sources for the commodity price and predictor variables. Besides, a brief characterization of the data is provided.

Raisin is chosen for agricultural commodity price index that is aimed to predict in this study. Raisin prices are compiled by Izmir Commodity Exchange, which was established in Izmir in 1891 as the first commodity exchange of Turkey. Raisin is one of the most trading commodities so as the beginning we establish a model to forecast movement of raisin price index. We use end-of-day prices measured at close, denominated in Turkish Liras. The research data used is direction of daily closing price movement in raisin price index. The entire data set covers the period from January 4, 2015 to June 30, 2017. The total number of cases is 627 trading days. Directions of closing prices that used in this study are threefold; increasing direction, unchanged and decrease direction. The number of cases with increasing direction is 239, the number of cases with unchanged 140 and the number of cases with decreasing direction is 248. Percentage wise cases of each month in the entire data set are shown in Table 1.

Table 1: *Percentage wise cases of each month in the entire data set*

	increase	%	unchanged	%	decrease	%
Jan-15	10	50%	1	5%	9	45%
Feb-15	11	55%	0	0%	9	45%
Mar-15	12	55%	0	0%	10	45%
Apr-15	10	48%	5	24%	6	29%

	increase	%	unchanged	%	decrease	%
Apr-16	8	38%	10	48%	3	14%
May-16	6	29%	5	24%	10	48%
Jun-16	7	32%	5	23%	10	45%
Jul-16	4	22%	10	56%	4	22%

May-15	9	47%	6	32%	4	21%
Jun-15	8	36%	4	18%	10	45%
Jul-15	7	32%	8	36%	7	32%
Aug-15	8	38%	6	29%	7	33%
Sep-15	8	40%	2	10%	10	50%
Oct-15	9	43%	2	10%	10	48%
Nov-15	12	57%	1	5%	8	38%
Dec-15	5	22%	12	52%	6	26%
Jan-16	8	40%	4	20%	8	40%
Feb-16	4	19%	7	33%	10	48%
Mar-16	3	14%	15	68%	4	18%
Aug-16	1	5%	17	77%	4	18%
Sep-16	8	44%	1	6%	9	50%
Oct-16	10	48%	0	0%	11	52%
Nov-16	11	50%	0	0%	11	50%
Dec-16	9	41%	1	5%	12	55%
Jan-17	11	50%	2	9%	9	41%
Feb-17	9	45%	2	10%	9	45%
Mar-17	10	43%	1	4%	12	52%
Apr-17	5	26%	4	21%	10	53%
May-17	8	38%	3	14%	10	48%
Jun-17	8	40%	6	30%	6	30%
Total	239	38%	140	22%	248	40%

Gold price, stock returns, exchange rates, oil price, lagged raisin price and social and political issue information are used as input variables. In the light of previous studies as Sujit and Kumar (2011), Samanta and Zadeh (2012), Mensi et al (2013) and Malliaris and Malliaris (2013), we hypothesize that gold price, stock returns, exchange rates and oil price which are most volatile variables and can be used as predictor. Besides, usability of information of social and political issues occurred in Turkey and social and political issues occurred in the World as input variables in prediction models for forecasting direction of raisin price index's movement is explored. For each predictor, gold price, stock returns, exchange rates and oil price, daily exchange ratios are compiled. Information about social and political issues, which are occurred in Turkey and World are constructed by authors by searching online dictionaries, newspaper etc. and scanning hard copy news, magazines and other related papers. This variable includes issues such as coup, extraordinary actions of government, terrorist attack resulting in deaths or election, which cause tension and uncertainty on social life and economy and impress them affluently. We designed two different social and political issues variables, one for Turkey and one for World. Hence, we can establish different models with or without global events to understand the effect of global and local events on prediction model separately. If an important trouble happens at time t, SPI variable t equals to 1. Else, value of SPI variable t is 0. Table 2 summarizes the predictors in depth. The direction of daily change in the raisin price index is categorized as "-1", "0" and "1". If the raisin price index at time t is higher than that at time t-1, direction t is "1". If the raisin price index at time t is lower than that at time t-1, direction t is "-1". Else which means there is no changing between the price at time t and t-1, direction t is "0".

Table 2: *Summary statistics for predictors*

Name of predictors	**Max**	**Min**	**Mean**	**Std. Dev.**
Crude oil price	61.43	26.21	46.86	6.89
Gold price	152.76	89.33	117.94	17.15
EUR/TRY	4.15	2.63	3.33	0.37
USD/TRY	3.88	2.28	3.03	0.38
BIST 100	100617.69	68567.89	81246.55	6821.54
Lagged Raisin Price	6.17	2.94	4.10	0.80
Social and political issues (Turkey)	1.00	0.00	-	-
Social and political issues (World)	1.00	0.00	-	-

Proposed Models

As mentioned before the aim of the study is predicting directional movement of raisin price with various input variables. According to research questions and motivation of the paper, we designed three different prediction models. Input variables are grouped by the research question to explore their effects on raisin price and prediction model. First model, as general model, includes all predictors as input variables. So, the aim on this model is estimating the direction of raisin price's movement in Turkey due to Crude oil price, Gold price, EUR/TRY, USD/TRY, BIST 100, Lagged Raisin Price, Social and political issues (Turkey), Social and political issues (World). Second model includes all variables except Social and political issues (World). Thereby, effectiveness of using the global events to predict raisin price movement in Turkey will be arisen. Third model does not include any social and political issues variable neither from Turkey nor from World. Through the model performance of prediction without social and political issues will be appeared. First model is built as general model to explore whole variables effects together. Second and third models will help to answer first research question because they provide prediction results with and without SPI variables. Also, third model presents the answer for second research question.

Table 3: *Prediction models and predictors matrix*

Predictors	Prediction Model1	Prediction Model2	Prediction Model3
Crude oil price	+	+	+
Gold price	+	+	+
EUR/TRY	+	+	+
USD/TRY	+	+	+
BIST 100	+	+	+
Lagged Raisin Price	+	+	+
Social and political issues (Turkey)	+	+	
Social and political issues (World)	+		

Artificial neural networks, support vector machines and decision tree methods are applied to proposed models that are shown in Table 3. Common aim of those methods is forecasting the output variable through the inputs and calculating prediction performances. Thereby while forecasting the direction movement of raisin price, prediction performances will be compared. Detailed information about application of methods and results are existing in "Experimental Results" section.

Predictive Methods

Artificial Neural Networks (ANN)

Thinking, creating and discovering are results of functioning of human brain. Human learns the ability of exporting information, creating or discovering new things by biological neural networks. This ability has become an inspiration to ANN. In this artificial network, inputs and outputs have designed as neurons. The neurons on input layer provide information to neurons in hidden layer, which serve as transfer function. Afterwards neurons that are produced knowledge compile on output layer. As shown in Figure 3, a three-layer feed forward neural network designed for our study. Inputs are USD/TRY, EUR/TRY, crude oil, gold price, stock return, lagged raisin price, Social & Political Issues in Turkey and worldwide issues. Political issues contain elections and military coup. Social issues include terrorist incidents and crisis. In the models, Social & Political Issues in Turkey and worldwide issues are defined as binary that if an unexpected issue occurs in relevant day, it will be 1, otherwise 0. Output layer contains neurons, which has categorical variables and represents direction of raisin price index. Those categories symbolize the up, unchanged or down movement of tomorrow's price prediction. Number of neurons in hidden layer has designated empirically.

Generating the output neurons with given input neurons is possible due to weights. Weights that to connect with neurons of neighboring layers are designated randomly as the initial values. Main parameters in ANN model are number of neurons in hidden layer (n), learning rate (lr) that determines the amount of change in weights, momentum constant (mc) which is rate of previous iteration's effect to the next and number of iterations (ep) (Kara, Boyacioglu, & Baykan, 2011). The levels of parameters that used in our model represented in Table 4. As is seen from Table 4 a total of 10 x 10 x 9 =900 treatments for ANN are carried out. Prediction accuracies are computed to compare parameters combination. Learning rate is varied in the interval of 0.1and 0.9 for the top three parameters combinations to determine the best combination.

Table 4: Lev*els of parameters that used in ANN model*

Parameters	Levels
Number of hidden layer neurons (n)	10,20,30,….,100
Epochs (ep)	100,200,…1000
Momentum constant (mc)	0.1,0.2,….,0.9
Learning rate	0.1,0.2,….,0.9

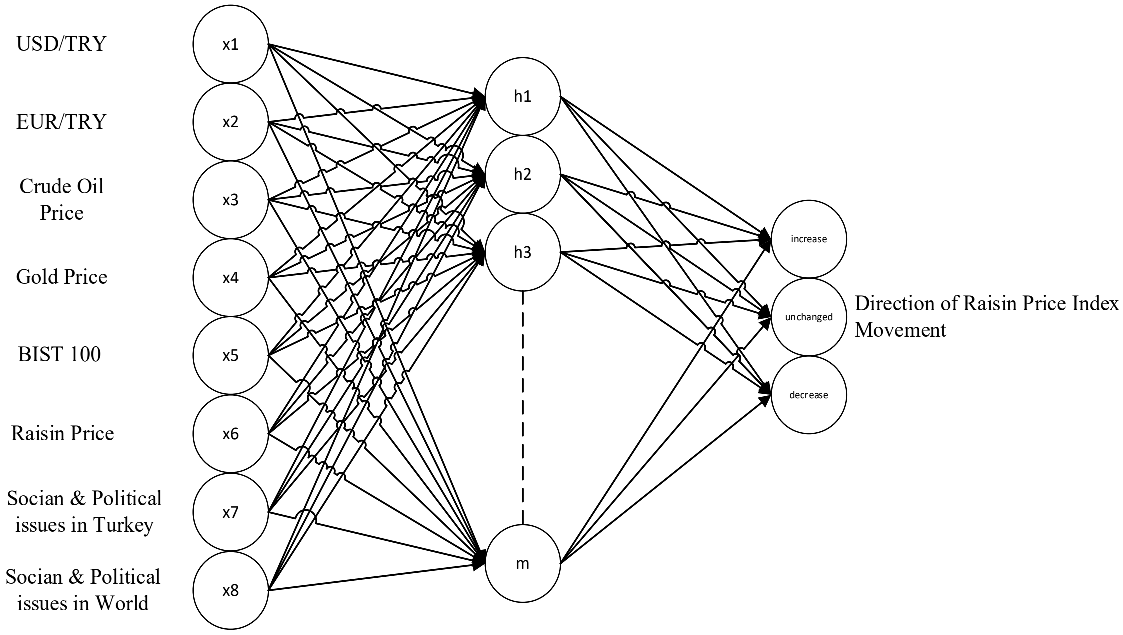

Figure 3: *Design of three-layer feed forward artificial neural network*

Support Vector Machines (SVM)

Cortes and Vapnik (1995) aimed to discovery optimal hyperplanes for linearly detachable classes in support vector machines concept. This concept asserts that optimal hyperplanes depend on support vectors, which are a small subset of training dataset (Dean, 2014). In other words, aim in algorithm is maximizing the gap between support vectors, which represents different classes. A three-class problem requires to have a set of inputs $x_i \in R^d$ (i=1, 2, …, N) with corresponding labels $y_i \in \{-1, 0, 1\}$ (i=1, 2, …, N). Here -1, 0 and 1 indicate the three classes.

Kernel function is used to lay out the input vectors $x_i \in R^d$ into a high dimensional feature space $\Phi(x_i) \in H$ and to establish an Optimal Separating Hyperplane (OSH) the following decision function is implemented by SVM (Scholkopf & Smola, 2001).

$$f(x) = sgn\left(\sum_{i=1}^{N} y_i \alpha_i \cdot K(x, xi) + b\right) \quad (1)$$

The coefficients are unknown parameters, which are derived from following mathematical model.

M a x i m i z e

$$\left(\sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_{jj} \cdot k_i k_j \cdot K(x_i, x_j)\right) \quad 2$$

Subject to $0 \leq a_i \leq C$ (3)

$$\sum_{i=1}^{N} \alpha_i y_i = 0_i = 1, 2, \ldots N \quad (4)$$

C, that appears in equation 3 is a parameter that manage the interchange of margin and misclassification error. Through several types of Kernel functions exist, polynomial and radial basis are most popular functions, which are applied in this study. In equation (5), *d* implies the degree of polynomial function and *y* in equation (6) is the constant of radial basis function.

$$K(x_i, x_j) = (x_i \cdot x_j + 1)^d \quad (5)$$

$$K(x_i, x_j) = exp(-\gamma ||x_i - x_j||^2) \quad (6)$$

The combinations, which are constituted by different Kernel functions, different parameters and different levels of those parameters, are used in our study. Detailed information about combinations is explained in Table 5. Predictive accuracies will be compared to identify the best combination.

Table 5: *Levels of parameters that used in SVM model*

Parameters	**Levels of Polynomial Function**	**Levels of Radial Basis Function**
Degree (d)	1,2,3,4	-
Gamma (γ)	0.1,0.2,….5	0.1,0.2,….5
Regularization Parameter (c)	1,10,100,100	1,10,100,100

Decision Tree: C4.5 Algorithm

Positive aspects as being fast, easy to use, simple and accurate configuration make decision trees popular in data mining (Duch, Setiono, & Zurad, 2004). Decision tree method is used to classify samples by arranging them down to tree from root to leaf nodes. Each internal node represent test for some attribute of the tree. The root is a node that has not entering edges. The other nodes have one entering edge precisely (Quinlan, 1993). In decision tree learning, samples are classified by starting from the root, down the tree to the leaves. Each leaf belongs to specified class, called target value (Mitchell, 1997).

C4.5, a later version of the ID3 algorithm, repeatedly visits each decision node, choosing the ideal split, until no more splits are existing. Information gain or entropy reduction thoughts are used in C4.5 algorithm to select the ideal split (Larose & Larose, 2015). Entropy of variable x can be defined as;

$$H(x) = -\sum{}^{j} p_j log_2(p_j) \qquad (7)$$

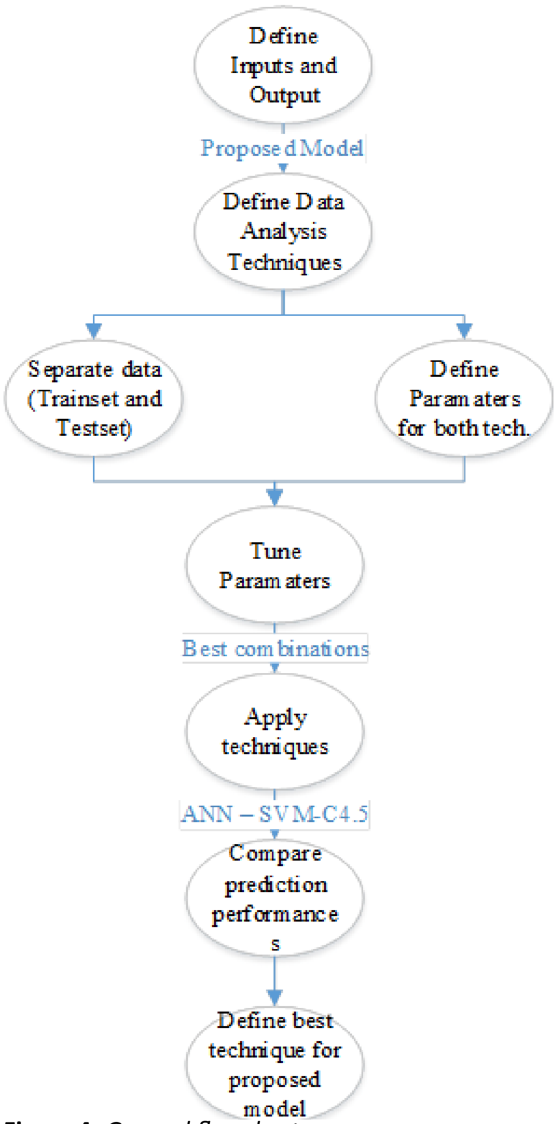

Figure 4: *General flowchart*

where p means probabilities. Besides information gain can be formulated as;

$$Gain(S)=H(T)-Hs(T) \qquad (8)$$

Gain (S) denotes the increase in information obtained by splitting the training data T according to this candidate split S. At each decision node, C4.5 selects the ideal split that has the highest information gain, gain(S).

Experimental Results

Various studies showed that improved data analysis techniques such as neural network and support vector machines are powerful and useful for forecasting applications like; financial, exchange rate, stock market and so on. These data analysis techniques are commenced to apply on commodity price index prediction models after 2005's. The studies especially about comparison of traditional techniques or linear forecasting techniques with new techniques have displayed the power of ANN, SVM and decision tree. Hence, in this study ANN, SVM and decision tree techniques used for the prediction model.

ANN, SVM and decision tree methods established with the inputs, which are stated in previous section to predict the direction of raisin price index. R, open library programming language, is utilized to forecast the direction. R has numerous packages for statistical computing and graphics. In the study, e1071 package for support vector machines model, RSNNS package for artificial neural networks model and C45 package for decision tree model are used. General application flowchart for models can be shown in Figure 4. In order to make more reliable predictions, a model evaluation method, called cross validation is used. Cross-validation focuses on examining predictive models by splitting the original sample into two data sets. Those are a training set to train the model, and a test set to examine it. Firstly, entire dataset is separated in two subsets randomly. Eighty percent of entire dataset represented as training set and remains are test set. The model learned from the training set predicts the output values for the data in test set. Thus, performance of model to estimate a data which is not seen before is proved. Then, parameters are tuned for each technique, in order to define which parameter combinations provide feasible prediction performance. Set parameter combinations for ANN, SVM and decision tree are revealed and results are obtained. Finally, the results acquired by proposed models are investigated due to the accuracies. Accuracy and f-measure are calculated to compare the performances. Accurate forecasting is in parallel with the minimized size of forecasted error (Stevenson, 2011). Accuracy and f measure for a perfect classifier must be equal to 1. Confusion matrix and accordingly; True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN) should be determined initially to designate accuracy and f-measure. Confusion matrix is shown in Table 6 and formulation of the accuracy and f-measure are represented in Eqs. (9)-(14).

$$Precision_{Positive} = TP / (TP + FP) \quad (9)$$

$$_{Negative} = TN / (TN + FN) \quad (10)$$

$$Recall_{Positive} = TP / (TP + FN) \quad (11)$$

$$_{Negative} = TN / (TN + FP) \quad (12)$$

$$Accuracy = (TP + TN) / (TP + FP + TN + FN) \quad (13)$$

$$F\text{-}measure = \frac{(2 \times Precision \times Recall)}{(Precision + Recall)} \quad (14)$$

Table 6: *Confusion Matrix*

		Predicted class	
		P	N
Actual Class	P	True Positives (TP)	False Negatives (FP)
	N	False Positives (FP)	True Negatives (TN)

Parameters for ANN and SVM are tuned in R via applying a grid search over provided parameter arrays. Aim of tuning is reaching the parameter combination, which provides the highest prediction performance. ANN is tuned for each prediction model and best parameters are determined. Those parameters and prediction performances of ANN for each method are as shown in Table 7. Results of ANN methods provide weights between the neurons. According to those weights in the first model, lagged raisin price variable has meaningful power. Another variable that has more effect on decreasing situation is social and political issues occurred in Turkey. Gold price is effective on unchanged situation and USD/TRY variable has powerful effect on increasing situation. In the second model, social and political issues occurred in Turkey has effect on decreasing situation, BIST 100 variable on unchanged situation and lagged raisin price variable on increasing situation. On the last model, lagged raisin price variable is powerful on decreasing situation, BIST 100 variable is important for unchanged situation and USD/TRY variable has effect on increasing situation.

Table 7: *Prediction performance of ANN for each model*

	Performance Measures	epochs : neurons : mc & lr= 0.1
Prediction Model 1		500 : 34 : 0.01
	Accuracy	0.822
	F-Measure	0.823
Prediction Model 2		500 : 33 : 0.01
	Accuracy	0.822
	F-Measure	0.823
Prediction Model 3		500 : 33 : 0.01
	Accuracy	0.755
	F-Measure	0.756

Same procedure is followed for SVM model. Due to accuracy and f measure parameters, best combination of depolynomial SVM. Therefore, performance of other prediction models and techniques compared with Radial Basis SVM model

Table 8: *Prediction performance of SVM for each model*

	Performance Measures	**Kernel: Polynomial**	**Kernel: RBS**
Prediction Model 1		C:10 gamma:0.2 d:3	C:100 gamma:0.6
	Accuracy	0.666	0.888
	F-Measure	0.667	0.889
Prediction Model 2		C:10 gamma:0.2 d:3	C:100 gamma:0.6
	Accuracy	0.681	0.851
	F-Measure	0.682	0.852
Prediction Model 3		C:10 gamma:0.2 d:3	C:100 gamma:0.6
	Accuracy	0.570	0.807
	F-Measure	0.571	0.808

To see the effects of global and local events on prediction model three different models established and analyzed by using three different methods. The accuracies of whole models and methods summarized on Table 10. It shows that all prediction models are significant and it means that first and second research questions are satisfied. Gold price, stock returns, exchange rates, crude oil price, lagged raisin price and social and political issue occurred in Turkey and global events can be used as input variables to predict direction of daily closing price of raisin. As an answer to third research question, comparison of prediction models indicates support vector machines method performs better on each prediction model while predicting direction of daily closing price of raisin. When examining the accuracies by prediction models and by methods, prediction model 1 with SVM has the best accuracy. Thereby, in order to make more powerful and accurate prediction on direction of raisin price, proposed inputs should be used together.

Table 10: *Prediction performance for each model and for each method*

	Support Vector Machines (RBS)	Decision Tree C4.5	Artificial Neural Networks
Prediction Model 1	0.888	0.837	0.822
Prediction Model 2	0.851	0.844	0.822
Prediction Model 3	0.807	0.780	0.755

Conclusions

Predicting the direction of movement of the agricultural commodity market index is important for the development of effective trading strategies. Farmers and traders are influenced by variation in agricultural commodity market index, thus they need efficient predictions before making an action. This study attempted to predict the direction of raisin price (one of the most trading agricultural product of commodity exchange in Turkey) movement along machine learning techniques. In order to expose the predictability of agricultural commodity price with machine learning methods, raisin price, which is compiled by Izmir Commodity Exchange is appointed. Three prediction models namely ANN, SVM and Decision Tree were established to predict and compare the prediction performance. Daily historical data is used for explanatory variables, which are gold price, stock returns, exchange rates, oil price and social and political issues occurred in Turkey and global events and responding variable.

Following research questions has developed within motivations of study;

1. Do the social & political issues effect raisin price in countries that have internal disturbance?

2. Can most volatile indicators such as gold price, stock returns, exchange rates and oil price help to forecast movement of raisin price?

3. Which one has the best forecasting performance as artificial neural networks, decision tree or support vector machine?

Experimental results are demonstrated important outcomes and responses for the research questions. Firstly, three models performed significant performance in predicting the direction of raisin price movement. The first prediction model, which has highest prediction performance, includes all inputs and that responses first and second research questions meantime. Furthermore, the outcomes of methods demonstrated that local social and political events have meaningful effect on prediction model and causes the decreasing movement. Another focus is discovering that machine-learning techniques such as ANN, SVM and Decision Trees can be utilized in prediction of raisin price movement is also proved. Due to third research question performances of ANN, SVM and decision tree models in predicting the direction of raisin price movement are compared. The accuracy performance of SVM (%88.8, %85.1, %80.7) is better than ANN model (%82.2, %82.2, %75.5) and decision tree (%83.7, %84.4, %78.0). Another outcome that comparisons reveal is third model has lowest accuracy in all algorithms. Which shows that, gold price, stock returns, exchange rates and oil price and daily exchange ratios have weak power while estimating the direction of raisin price movement and SPI variables have significant impacts for this study.

In this paper, raisin price index is designated as agricultural commodity price index. For future studies other agricultural commodity price indexes such as cotton or olive oil can be determined. As other future studies; instead of short term prediction, long term prediction can be thought", length of historical data can be extended. In addition, this study can provide advantages to applicability of other types of agricultural commodity price indexes and applicability of other countries agricultural commodity price indexes.

Kaynakça/References

Ahmed, S. (2008). Aggregate economic variables and stock markets in India. *International Research Journal of Finance and Economics, 14*, 141-164.

Azadeh, A., Moghaddam, M., & Khakzad, M. (2012). A flexible neural network-fuzzy mathematical programming algorithm for improvement of oil price estimation and forecasting. *Comput. Ind. Eng. 62(2): 421–430*.

Baffes, J., & Haniotis, T. (2016). What explains agricultural price movements? *Journal of Agricultural Economics , 67*(3), 706-721.

Chen, A.-S., Leung, M. T., & Daouk, H. (2003). Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index . *Computers & Operations Research, 30*(6), 901-924.

Chen, Y. C., Rogoff, K., & Rossi, B. (2010). Predicting agri-commodity prices: An asset pricing approach. *SSRN*.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning, 20*(3), 273-297.

Dean, J. (2014). *Big data, data mining, and machine learning: value creation for business leaders and practitioners.* New Jersey: John Wiley & Sons. Computational intelligence methods for rule-based data understanding. *Proceedings of the IEEE, 92*(5), 771-805.

Duch, W., Setiono, R., & Zurad, J. (2004). Computational intelligence methods for rule-based data understanding. Proceedings of the IEEE, 92(5), 771-805.

Gargano, A., & Timmermann, A. (2014). Forecasting commodity price indexes using macroeconomic and financial predictors. *International Journal of Forecasting, 30*(3), 825-843.

Groen, J., & Pesenti, P. A. (2011). Commodity prices, commodity currencies, and global economic developments . In T. Ito, & A. Rose, *Commodity prices and markets, East Asia Seminar on Economics* (Vol. 20, pp. 15–42). Chicago: University of Chicago.

Haidar, I., & Wolff, R. (2009). *Forecasting crude oil price.* Retrieved from http://www.usaee.org/usaee2011/ submissions/ OnlineProceedings/Forecasting%20 Crude%20Oil%20Price%20%28Revisited%29.pdf.

Hong, H., & Yogo, M. (2012). What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics, 105*(3), 473–490.

Hu, J. W.-S., Hu, Y.-C., & Lin, R. R.-W. (2012). Applying neural networks to prices prediction of crude oil futures. Mathematical Problems in Engineering.

Jammazi, R., & Aloui, C. (2012). Crude oil price forecasting: experimental evidence from wavelet decomposition and neural network modeling 34: 828–841. *Energy Econ., 34*(3), 828-841.

Jha, G. K., & Sinha, K. (2014). Time-delay neural networks for time series prediction: an application to the monthly wholesale price of oilseeds in India. *Neural Computing and Applications, 24*(3), 563-571.

Kara, Y., Boyacioglu, M., & Baykan, Ö. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert systems with Applications, 38*(5), 5311-5319.

Kohzadi, N., Bahman, M. S., & K, I. (1996). Comparison of artificial neural network and time series models for forecasting commodity prices. *Neurocomputing, 10*(2), -.

Larose, D., & Larose, C. (2015). *Data mining and predictive analytics.* New Jersey: John Wiley & Sons.

Li, G.-q., Xu, S.-w., & Li, Z.-m. (2010). Short-term price forecasting for agro-products using artificial neural networks. *Agric. Sci. Procedia, 1*, 278–287.

Malliaris, A., & Malliaris , M. (2013). Are oil, gold and the euro inter-related? Time series and neural network analysis. *Review of Quantitative Finance and Accounting, 40*(1), 1-14.

Mensi, W., Beljid, M., Boubaker, A., & Managi , S. (2013). Correlations and volatility spillovers across commodity and stock markets: Linking energies, food, and gold. *Economic Modelling, 32*, 15-22.

Mitchell, T. (1997). *Machine Learning.* New York: McGrawHill.

Parisi, A. (2008). Forecasting gold price changes: rolling and recursive neural network models. *Journal of Multinatl. Financ. Manag, 18*(5), 477–487.

Quinlan, J. (1993). *Programs for Machine Learning.* California.: Morgan Kaufmann.

Samanta, S., & Zadeh, A. (2012). Co-movements of oil, gold, the US dollar, and stocks. *Modern Economy, 3*(1), 111-117.

Scholkopf, B., & Smola, A. (2001). *Learning with kernels: support vector machines, regularization, optimization, and beyond.* MIT press.

Shannon, C. (1951). Prediction and entropy of printed English. *Bell system technical journal, 30*(1), 50-64.

Sujit, K. S., & Kumar, R. (2011). Study on dynamic relationship among gold price, oil price, exchange rate and stock market returns. *International Journal of Applied Business and Economic Research, 9*(2), 145-165.

United States Department of Agriculture. (2017, 10

10). Retrieved from Market and Trade Data: https://apps.fas.usda.gov/psdonline/app/index.html#/app/topCountriesByCommodity

Wang, Y., & Chueh, Y. (2013). Dynamic transmission effects between the interest rate, the US dollar, and gold and crude oil prices. *Economic Modelling, 30*, 792-798.

Yakut, E., & Gemici, E. (2017). LR, C5. 0, CART, DVM Yöntemlerini Kullanarak Hisse Senedi Getiri Sınıflandırma Tahmini Yapılması ve Kullanılan Yöntemlerin Karşılaştırılması: Türkiye'de BIST'de Bir Uygulama. *Ege Akademik Bakis*, 17(4), 461-479.

Ye, M., Zyren, J., & Shore, J. (2002). Forecasting crude oil spot price using OECD petroleum inventory levels. *International Advances in Economic Research, 8*(4), 324-333.

Zou, H., Xia, G., Yang, F., & Wang, H. (2007). An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting. *Neurocomputing, 70*(16-18), 2913–2923.