A DECISION SUPPORT SYSTEM FOR DETERMINING THE SUITABLE FISH SPECIES TO FISH FARMS

İnci Elif HADIK[1*], Ukbe Üsame UÇAR[2], Mehmet ATAK[3], Selçuk Kürşat İŞLEYEN[4]

[1]Gazi Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Ankara
ORCID No : https://orcid.org/0000-0001-8769-1921
[2]Fırat Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Elazığ
ORCID No : https://orcid.org/0000-0002-9872-2890
[3]Gazi Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Ankara
ORCID No : https://orcid.org/0000-0002-4373-5192
[4]Gazi Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Ankara
ORCID No : https://orcid.org/0000-0002-0113-1083

Keywords

Aquaculture, Fish Farming, Data Mining, Classification and Regression Tree.

Abstract

Fishery industry is one of the main sources of income and most important subsectors for national economies. Nevertheless, natural fish sources have unfortunately diminished recently and not every fish species can be grown in every region due to some reasons. Moreover, many countries have limited resources to meet the need for fish. Therefore, aquaculture comes into prominence to eliminate these problems. Also, taking right decisions and selecting the right fish breeds become crucial for fish farming. In this context, the aim of this study is to develop a Decision Support System (DSS) with intent to enable decision makers to determine most suitable fish species according to the features of their farms easily. The system was built based on Classification and Regression Trees algorithm, one of the data mining techniques. Sixty-two breeds of fish and thirteen factors affecting their growth was studied to create a database. The results show that the suggested DSS functions successfully in terms of not only determining appropriate and profitable fish species but also using existing resources more efficiently. It is expected that foreign trade volume will be increased with the raising productivity and; hence, countries will create new business branches which will have reflections in employment figures in the long run.

ÇİFTLİKLERDE YETİŞTİRİLECEK UYGUN BALIK TÜRLERİNİN BELİRLENMESİNE YÖNELİK BİR KARAR DESTEK SİSTEMİ

Anahtar Kelimeler

Su Ürünleri Yetiştiriciliği, Balık Yetiştiriciliği, Veri Madenciliği, Sınıflandırma ve Regresyon Ağacı.

Öz

Balıkçılık endüstrisi ulusal ekonominin en önemli alt sektörlerinden ve temel gelir kaynaklarından biridir. Buna karşın, son zamanlarda doğal balık kaynakları ne yazık ki yok olmakta ve bazı nedenlerden dolayı her bölgede her balık türü yetiştirilememektedir. Ayrıca, birçok ülke balık ihtiyacını karşılamak için sınırlı kaynaklara sahiptir. Bu sebeple, su ürünleri yetiştiriciliği bu problemlerin çözülmesinde ön plana çıkmaktadır. Aynı zamanda, doğru balık türlerinin seçilmesi ve doğru kararların alınması balık yetiştiriciliği için büyük bir öneme sahiptir. Bu bağlamda, bu çalışmanın amacı karar vericilerin, ilgili çiftliğin özelliklerine en uygun balık türünün kolay bir şekilde belirlenmesini sağlayan bir Karar Destek Sistemi (KDS) geliştirmektir. İlgili sistem veri madenciliği tekniklerinden Sınıflandırma ve Regresyon Ağaçları Algoritması' na dayalı olarak kurulmuştur. 62 balık türü ve bu balıkların büyümesine etki eden 13 faktörden oluşan bir veri tabanı oluşturulmuştur. Sonuçlar, önerilen KDS yapısının yalnızca uygun ve karlı balık türlerinin belirlenmesinde değil aynı zamanda mevcut kaynakların daha verimli bir şekilde kullanması açısından da başarılı bir şekilde çalıştığını göstermiştir. Artan verimlilik ile birlikte dış ticaret hacminin artacağı ve dolayısıyla ülkelerin uzun vadede istihdam rakamlarına yansıyacak yeni iş kolları oluşturacağı beklenmektedir.

Araştırma Makalesi		Research Article	
Başvuru Tarihi	: 01.09.2020	Submission Date	: 01.09.2020
Kabul Tarihi	: 13.11.2020	Accepted Date	: 13.11.2020

*Sorumlu yazar; e-posta : incielifhadik@gazi.edu.tr

1. Introduction

In today's competitive world, countries have to operate in various sectors both at national and international level. Fishery is one of the most profitable fields and most important subsectors for national economies of countries that have rich water resources (FAO, 2016). In this regard, Turkey has a substantial potential for the sector which provide income and employment (Kaygisiz & Eken, 2018).

Fishing has been the main source of food since the existence of mankind. Fish is rich in protein and easily digestible in a short time. Its fat ratio is low while its meat is nutritious and tasty. Besides, it is very beneficial for children and adults because it contains many vitamins, minerals and phosphorus. Despite all these benefits, natural fish sources have unfortunately declined and some species of fish are threatened with extinction as a result of population growth, wrong, unconscious and excessive fishing, and inefficient use of water resources. The problem arising with the decrease in the number of natural fish can only be eliminated by aquaculture. Approximately 40% of the world's aquaculture production is carried out by fish farming. It is expected that the amount of fish obtained by aquaculture will be equal to fishing by 2030, and its importance will increase in the long run. (http://www.isub.org.tr).

Around the world, many countries have coastal areas or engage in fishing activities through their water and river beds. In this sense, Turkey has a great potential in fisheries in terms of its unique geological properties such as being a peninsula and having many water resources. The amount of fish obtained by aquaculture has grown up over the years while the quantity of the produced fish has gradually declined (https://www.tuik.gov.tr). Despite the decrease in hunting, the rate of fish export and the market for fish have increased substantially thanks to aquaculture. In Turkey, the amount of fish export shows a 20% increase compared to that of last year (https://www.tarim.gov.tr). Moreover, in direct proportion to the increase in the export rate, the import rate belonging to 2018 presents a 25% decrease in reference to last year. The foreign trade volume can be surged by the detection and farming of other fish species which have not been found in Turkey.

As opposed to the countries that have rich water resources, certain countries or areas do not have adequate water resources for fishing. Besides, every geography on Earth has different climatic characteristics and every fish species cannot be raised in every region. According to the Ministry of Economy fisheries sector report, many European Union member countries are highly dependent on import to meet their fish needs and they import 57% of their consumption from others (https://www.ekonomi.gov.tr). Foreign source dependency stemming from inadequate resources can be reduced by aquaculture. However, selecting the right fish breeds for aquaculture has great importance for productive use of available sources. Therefore, governments develop many policies and give substantial investment supports to make the best use of available sources and maintain fish populations (http://www.tkdk.gov.tr). In this context, proper policies in fisheries and aquaculture have upmost importance not only for more efficient use of water resources and preservation of fish populations but also for enhancement of domestic and foreign trade volume.

Incorrect decisions in fisheries often result in failure and cause great financial loss. Accordingly, decision support systems come into prominence as a systematic approach for the decision making process. These systems are quantitative tools which allow decision makers to analyze outcomes of their policies before implementation (Azadivar, Truong, & Jiao, 2009). Especially, data analysis is considered to be the most important phase for the system to function successfully (Xiaoshuan, Tao, Revell, & Zetian, 2005). Data Mining (DM) has gained ground for various data analysis tasks including forecasting, clustering and association in recent years.

DM is the process of analyzing large data sets to find unsuspected and unknown relationships and discovering useful information in novel ways that are both understandable and useful to analysts (David, Heikki, & Padhraic, 2001; Tan, Steinbach, & Kumar, 2006). The interpretation of huge amount of data and extracting useful information has become extremely challenging with developing data collection and storage technology. Despite the increase in the amount of data, the difficulty in obtaining meaningful information from large data sets is described as "data rich but information poor situation" by Han and Micheline (2006). DM applications have been applied a variety fields including finance (Dutta, Dutta, & Raahemi, 2017), banking (Aburrous, Hossain, Dahal, & Thabtah, 2010), customer relationship management (Bahari & Elayidom, 2015), manufacturing (Vazan, Janikova, Tanuska, Kebisek, & Cervenanska, 2017), marketing

(Wang, Ting, & Wu, 2013), healthcare and medicine (Delen, Fuller, McCann, & Ray, 2009; Daş ve Türkoğlu 2014), education (Angeli, Howard, Ma, Yang, & Kirschner, 2017), animal husbandry (Uçar, Balo, Eraslan, & Çetin, 2017) and so on.

Multiple studies using DM techniques take part in the literature which focused on fisheries and aquaculture. Various techniques including Artificial Neural Networks (ANNs), Support Vector Machine (SVM), Self-Organizing Mapping (SOM), Back-Propagation Classifier, and Decision Tree (DT) algorithms were applied for classification purposes, particularly in the studies of fish species identification (Cabreira, Tripode, & Madirolas, 2009; Buelens, Pauly, Williams, & Sale, 2009). Generally, fish species are classified based on their hydro acoustic data (Haralabous & Georgakarakos, 1996), width and height (Storbeck & Daan, 2001), texture (Larsen, Olafsdottir, & Ersbøll, 2009), size (Alsmadi, Omar, Noah, & Almarashdah, 2009; Alsmadi, Omar, Noah and Almarashdeh, 2010; Alsmadi, Omar, & Noah, 2011), color (Hu, Li, Duan, Han, Chen, & Si, 2012), shape (Ogunlana, Olabode, Oluwadare, & Iwasokun, 2015) etc.

Fisheries monitoring and management is another field that take advantage of DM techniques. Marzuki, Garello, Fablet, Kerbaol and Gaspar (2015) and Marzuki, Gaspar, Garello, Kerbaol and Fablet (2017) implemented SVM and Random Forest to identify illegal fishing activities, whereas Joo, Bertrand, Chaigneau and Niquen (2011) applied neural network to specify fishing set positions using vessel monitoring system (VMS) data. Similarly, Su and Chang (2008) used DBSCAB clustering algorithm with the aim of controlling over-fishing through VMS. Mendoza, García and Baro (2010) analyzed the impacts of fisheries management plans on the yield of a fish breed using classification tree algorithms. Gerami and Rabbaniha (2018) forecasted the status of a fish species, Anchovy Kilka, landings using time series model with the purpose of analyzing the feature catch status of Kilka and understanding its population decline dynamics. In relation to that, certain studies (Yuan, Li, & Chen, 2009; Hong-Chun, Ying, & Ying, 2009) pointed out that DM applications reveal better results than statistical techniques for fishery forecasting.

Tsai, Huang, Cheng, Shao and Chang (2017) investigated the relationship between fish communities and water quality using ANNs for upstream eco-hydrological system in northern Taiwan while Yang, Cai and Herricks (2008) developed an approach using Genetic Programming to identify hydrologic indicators related to fish community. Gutiérrez-Estrada, Yánez, Pulido-Calvoa, Silva, Plaza and Bórquez (2009) also used ANNs along with Multiple Linear Regressions techniques to forecast a fish species, sardine, landings considering their relationships with environmental factors such sea surface temperature and sea level. Sucipto, Kusrini and Taufiq (2016) applied C4.5 algorithm to understand relationship history among fish diseases using data about the symptom history of fish.

In addition to aforementioned studies, a literature review study by Suryanarayana, Braibanti, Rao, Ramam, Sudarsan and Rao (2008) shows that DM techniques were used to solve different problems in the field of fisheries such as the acoustic classification of fish (Anderson and Horne, 2007), detection of presence-absence (Mastrorillo, Lek, Dauba, and Belaud, 1997) and the migration of fish species (Kim, 2003) and then the prediction of fish abundance in a certain region (Zhou, 2003), clustering of fish communities (Reyjol, Fischer, Lek, Rosch, and Eckmann 2005; Park, Grenouillet, Esperance, and Lek, 2006), forecasting the distribution of caught fish species (Joy and Death, 2004; Gaertner and Dreyfus-Leon, 2004), and so on.

Xiaoshuan et al. (2005) provided a forecasting support system, Aquatic Products Price Forecasting Support System (APPFSS), to decrease aquatic price variability in Chinese fishery industry. The system contained nine modules including user interface, data collection, data analysis, data forecasting, database, model base and knowledge base modules. The data forecasting module was generated with ANN and such statistical techniques as moving average and linear regression. Similarly, Halide, Stigebrandt, Rehbein and McKinnon (2009) developed a Decision Support System (DSS) called CADS_TOOL to cage aquaculture managers. They applied Analytical Hierarchy Process (AHP) with four main criteria and twelve sub criteria for site classification and site selection tasks while used MOM (Modelling-On Growing-Monitoring) method to determine the capacity and economic value of a fish farm at a given site. Azadivar et al. (2009) submitted a DSS with the aim of enabling fishery managers to determine the amount of fish species caught in each sub-area in the Northeastern US so that sustainability of fish stocks objectives is satisfied and the value of landings is maximized. They preferred operations research and systems science approach to model their problem. In a similar manner, Sun, Li, Xiao, and Yang (2009) proposed a DSS including operation research and system science approach to fisheries management application. While the optimization approach determined the optimal fishing effort allocation in terms of time, location and amount of catch, simulation function evaluates the performance of an area management plan based on a group of criteria. Kvyatkovskaya, Kosmacheva, Sibikina, Galimova, Rudenko and Barabanova (2017) developed a DSS to address risk management in fishery industry. As distinct from abovementioned studies, they proposed a mathematical model for choosing an optimal strategy of the producer behavior of in the field of fisheries management and a fuzzy model of information risk analysis for processing information in information systems.

Overfishing is another area where the DSS is used. Overfishing is a global fisheries problem because many of the fish stock of the fisheries have already declined to below a tolerable level (Sholahuddin, Ramadhan, and Supriatna, (2015). In this context, Supriatna, Ramadhan, and Husniah (2015) proposed an integrated DSS for supporting fisheries management decision making in fisheries industries. They used Multiple Linear Regression (MLR) with Ordinary Least Square (OLS) to estimate the values of the growth parameters of harvested population and Maximum Sustainable Yield (MSY), known as the best harvest of the fishery. Similarly, Sholahuddin et al. (2015) proposed a DSS to deal with the fisheries problem by calculating the value of MSY. To estimate growth parameter, they applied Artificial Neural Network with Linear Perceptron method (ANN-LP) instead of MLR-OLS. They claimed that effectiveness of the proposed ANN-LP is as good as the MLR-OLS in estimating both the growth parameters and the MSY. Supriatna Sholahuddin, Ramadhan, and Husniah (2016) presented a DSS, SOFish ver 1.2, for estimating the best harvest in a fishery industry. They employed MLR-OLS method but different mathematical models, such as Verhulst, Gompertz and Richards models, were preferred to forecast the growth parameter. Thanks to their systems, decision makers, harvesters, find the value of maximum sustainable yield and harvest level recommendation easily.

To sum up, most of the reviewed articles applied various DM techniques to solve different fishery problems whereas a few studies dealt with DSS to satisfy the need in this field and aquaculture. Moreover, there are only a few studies benefiting from the advantage of DM techniques in DSS development phase. In particular, a study on DSS that helps fish farmers to identify which fish species could be grown in their fish farms was not found in the literature. In this study, we developed a Decision Support System using a Data Mining technique, Classification and Regression Trees (C&RT) algorithm, to fill the gap in the fishery literature and meet the systematic decision-making mechanism the need of the farmers. C&RT algorithm was applied to determine suitable fish breeds for the feature of their farms so as to increase their market volume and profitability. In this sense, the aim of this study is to develop a novel DSS with intend to enabling decision makers to easily determine pertinent fish species for their fish farms. In this context, this study has great importance in terms of offering a new insight into academic researchers. Moreover, this application assists professionals to increase the profitability by determining the most suitable fish species.

The rest of this paper consists of four sections and organized as follows. Section 2 introduces DSS design and C&RT algorithm. Section 3 presents the implementation of proposed DSS. Section 4 discusses the DSS and the results of case study. Finally, section 5 furnishes the concluding remarks of the research.

2. Decision Support System for Fish Selection

In this study, we declare that we comply with scientific and ethical principles. A decision support system is developed to enable decision makers to determine easily most suitable fish breeds in keeping with the features of their fish farm. The system is designed so as to present the most productive and profitable fish species. In order to achieve this goal, the fish breeds are tried to be classified through associating the characteristics of fish species with geographic and climatic factors.

2.1 Decision Support Systems (DSS)

In some cases, decision makers can rely on their experiences to make a quality decision or they do not feel necessary for looking at any additional information. In particular, decision makers who at tactical and strategic level face difficult situations beyond the ability to fully synthesize complex factors and they need a systematic decision-making mechanism. In these cases, decision support systems come to the help of users. A Decision Support System (DSS) is an interactive system that provides easy access to decision models and data in order to support the user in semi-structural and non-structural decision-making (Yıldız, Dağdeviren, and Çetinyokuş, 2008). In general, basic components of a DSS are as Figure 1.

The DSS database consists of a combination of past and current data from many applications or groups. The DSS software system covers the software tools used for data analysis. This system consists of a variety of online analytical processing tools, data mining tools, or a combination of mathematical and analytical models. Online analytical process tool and data mining tool are used for data analysis. User-interactive interface allows decision-makers to access DSS. The user leads the decision support system through the interface (Yıldız, Dağdeviren, and Çetinyokuş, 2008).

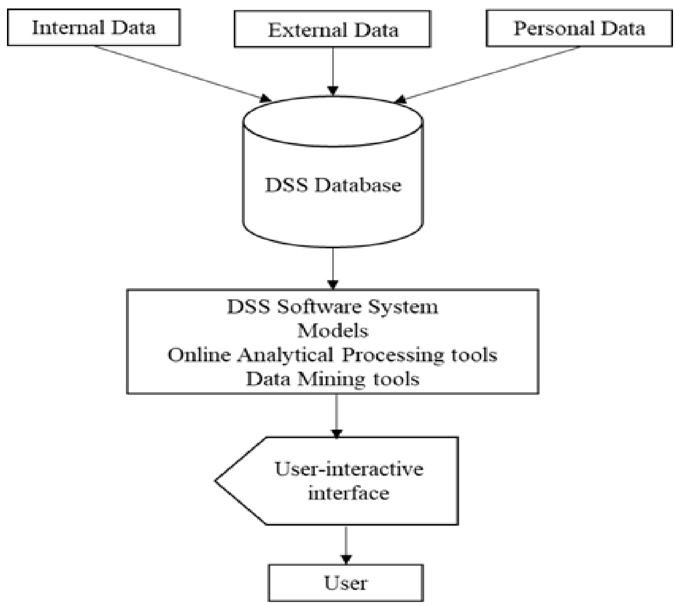

Figure 1. Basic components of a decision support system (Yıldız et al., 2008)

2.2 Data Mining Tools

Data Mining (DM) in general terms is defined as finding out the previously unknown and usable information that remain hidden in large amount of data. The aim of DM is to examine large amount of collected, meaningless data through a number of statistical methods, to render them as meaningful and useful information, and to benefit from them by getting them utilized in related institutional and management support systems.

The method is also a multidisciplinary field combining mathematical methods, statistical methods, computer technologies and database systems (Emel & Taşkın, 2005). The DM method is basically composed of six steps stated in Figure 2 (Larose, 2004; Albayrak & Yılmaz, 2009).

Through the indicated steps, it is aimed to obtain significant results from the DM. The DM models have different functions and can be used for various purposes. These functions are grouped under three headings such as classification and regression models, clustering models, and association rules in the literature (Ngai, Xiu, & Chau, 2009). Classification is one of the data mining methods that estimates data values by moving from known results in a data set consisting of different data. (Bhardwaj & Pal, 2011). There are many models developed for the classification process in the literature. Decision trees, artificial neural networks, bayes classifiers, k-nearest neighbor method, support vector machines, genetic algorithms are some of them. (Ngai et al., 2009).

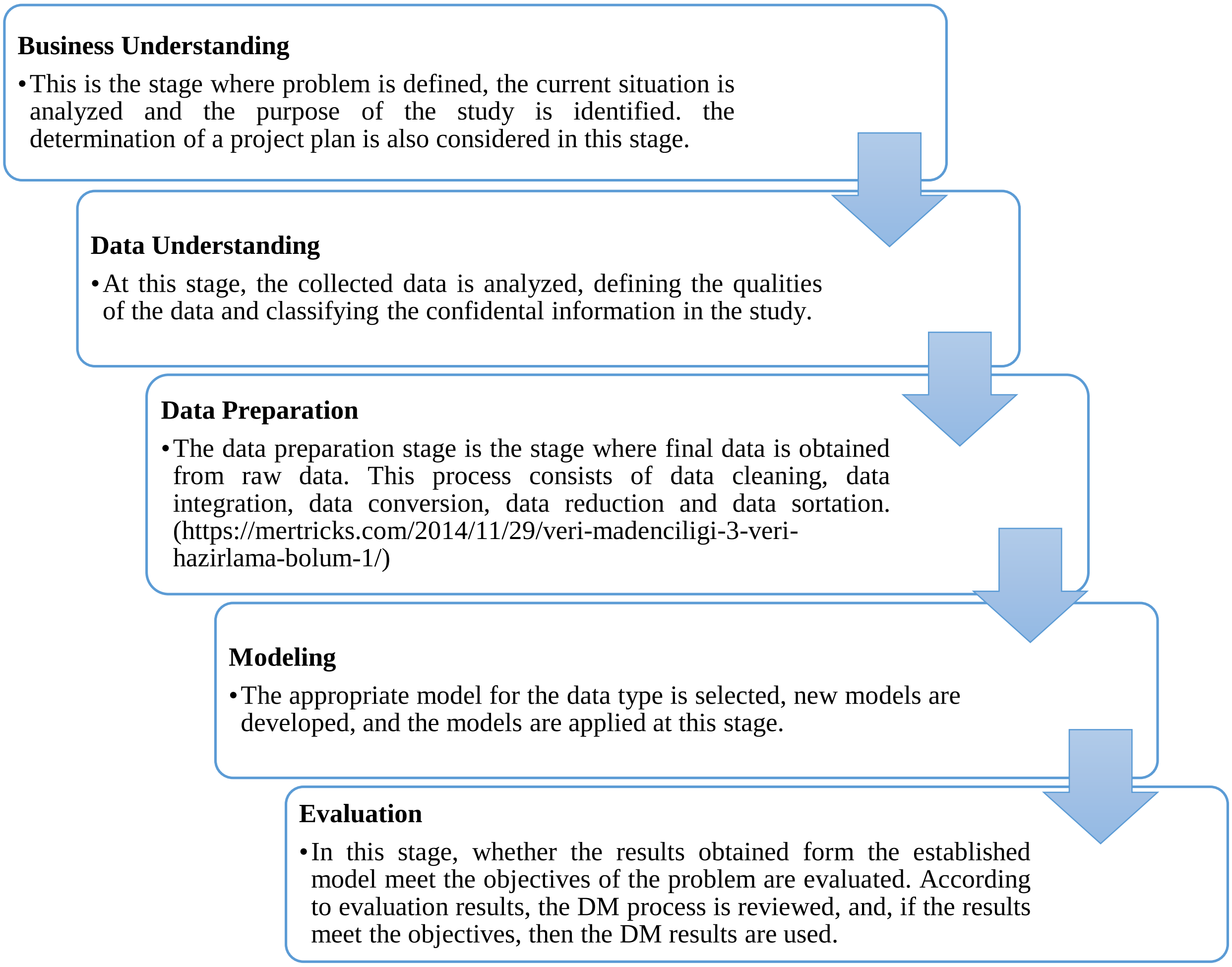

Figure 2. DM Process (Larose, 2004; Albayrak & Yılmaz, 2009).

First introduced in the 1960's, Decision Trees (DTs) are one of the most frequently used DM classification methods; it starts from a root node and then makes prediction by branching the data. (Song & Ying, 2015) Since the image resulting from the branching looks like a tree, it is called the DT.

In the literature, there are many DT algorithms. The Classification and Regression Trees (C&RT), is one of the DT algorithms, can deal with both classification and regression problems. A C&RT tree is a binary decision tree that begins with root node that contains the whole learning sample and is constructed by splitting the node into two child nodes once and again (Breiman, Friedman, Olshen, and Stone, 1984). In the method, the data are processed as raw, and no grouping operation is needed (Wu, Kumar, Quinlan, Ghosh, Yang, Motoda, ... & Zhou, 2008).

3. Decision Support System Implementation for Turkey

The developed system comprises of a database consisting 1948 observations, data analysis tool, and a user-friendly interface. The database contains 62 different fish breeds that exist in Turkey. There are many fish species in the world and in Turkey, and there are many factors having impact on the breeding of the fish species. In this study, 13 different factors affecting fish breeding were considered. The factors were determined based on expert and academician opinion. Some of these factors take nominal values while some of them take numerical values. These factors and their values are shown in the Table 1.

Table 1
The criteria taken into account for fish farming

Attribute	Value
The habitat in which fish live	river, lake, sea, ocean
The characteristics of body of water that fish live	fresh, salty, bitter
The temperature of the water in which the fish live	warm, mild, cold
Nutrition type	octopus, fish pups, plants, insects, jellyfish, other fish, bottom animals, lobster, mackerel, shellfish, shrimp, small fish, worms, larvae, mussels, squid, invertebrates, plankton, mosquitoes, dew worm, underwater animals, water plants, bonito, crab, artificial bait, algae, mollusks
The size of the fish	Numerical
The season in which the fish propagate	spring, summer, autumn, winter
The temperature of the water in which the fish propagate	Numerical
The depth of the water that fish live	Numerical
Bottom of the water that the fish propagate	moss, vegetation, gravel, mud, rocky, sandy, mussel shells
Is it a predator fish, or not?	yes-no
The taste of the fish	not tasty, a little delicious, delicious, very delicious
Economic value of the fish	low, medium, high, very high
Market volume	low, medium, high, very high

As a data analysis tool, data mining approach was used to analyze the database and find appropriate fish breeds. In this study, C&RT algorithm, one of the DT algorithms, was preferred because it is able to consider both numerical, i.e. the size of the fish, and nominal data, i.e. the habitat in which fish live, types as input and predictive variable. Also, it is not only both predictive and descriptive but also capable of classify large amount of data (Albayrak & Yılmaz, 2009).

The algorithm proceeds by dual splitting in a recursive way. Each root node represents a single input variable and a split point on that variable. After splitting process, every branched variable is named as leaf. The leaf nodes of the tree contain an output variable which is used to make a prediction for next splitting. Creating a C&RT model includes selecting input variables and split points on those variables until a suitable tree is constructed. The selection of which input variable to use and the specific split or cut-point is chosen using a greedy algorithm to minimize a cost function. The greedy approach is used to divide the space called recursive binary splitting. This is a numerical procedure where all the values are lined up and different split points are tried and tested using a cost function. The split with the best cost (lowest cost because we minimize cost) is selected. All input variables and all possible split points are evaluated and chosen in a greedy manner, e.g. the very best split point is chosen each time (Breiman et al., 1984).

The Gini Index was used as a branching criterion and the branches proceed with continuous division without any stopping criterion. In other words, the

Index was preferred to provide an indication of how much "pure" the leaf nodes were.

Basically, the Gini Index is determined by deducting the sum of squared of probabilities of each class from one, mathematically, it can be expressed as (Breiman et al., 1984):

$$Gini\ Index = 1 - \sum_{i=1}^{n} (P_i)^2 \quad (1)$$

Where Pi denotes the probability of an element being classified for a distinct class. While designing the decision tree, the features possessing the least value of the Gini Index would get preferred.

If the branching stops and a new division does not occur, a pruning operation is performed from the end node to the root node. After every pruning, the algorithm tries to determine the most successful DT. (Çalış, Kayapınar, & Çetinyokuş, 2014). The developed tree can be stored to file as a graph or a set of rules.

Lastly, user-interactive interface was developed in C# so that the user can easily guide the DSS. The solution mechanism is shown in Figure 3 and the developed interface is presented in Figure 4.

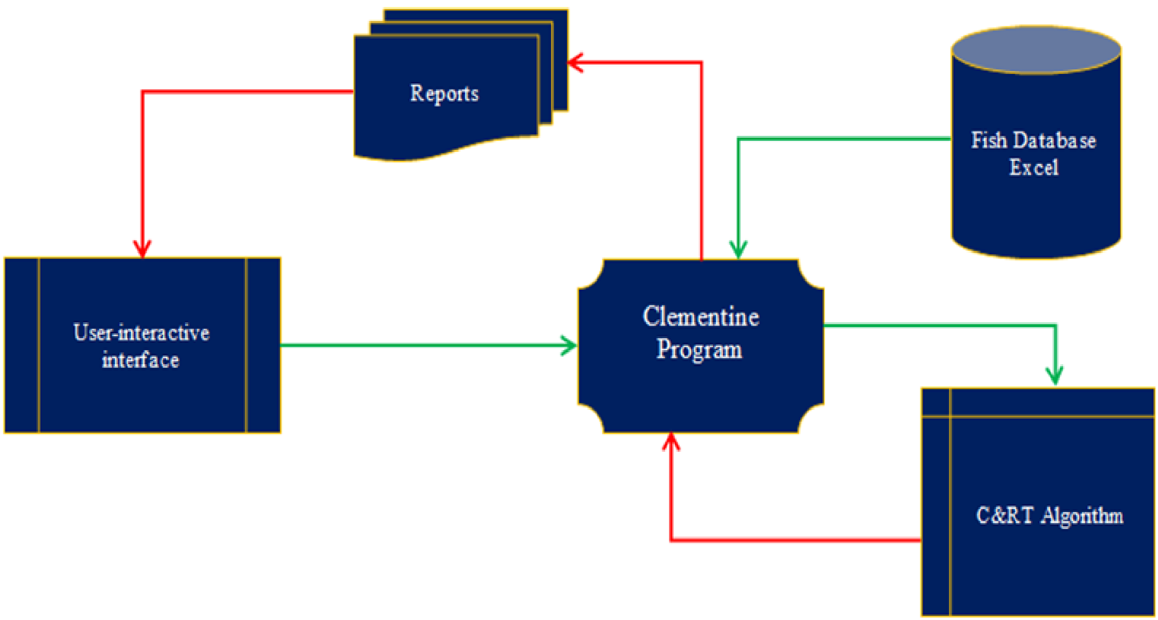

Figure 3. The workflow of the DSS

As seen in Figure 3, the solution mechanism begins with the execution of the interface by the user. The interface activates the Clementine Program in the background. The program retrieves data files from the excel program and performs the classification process using the C&RT algorithm. The classification results are reported by the Clementine program and transferred to the interface of the program. The interface of the program outputs the most suitable fish species based on the related parameters.

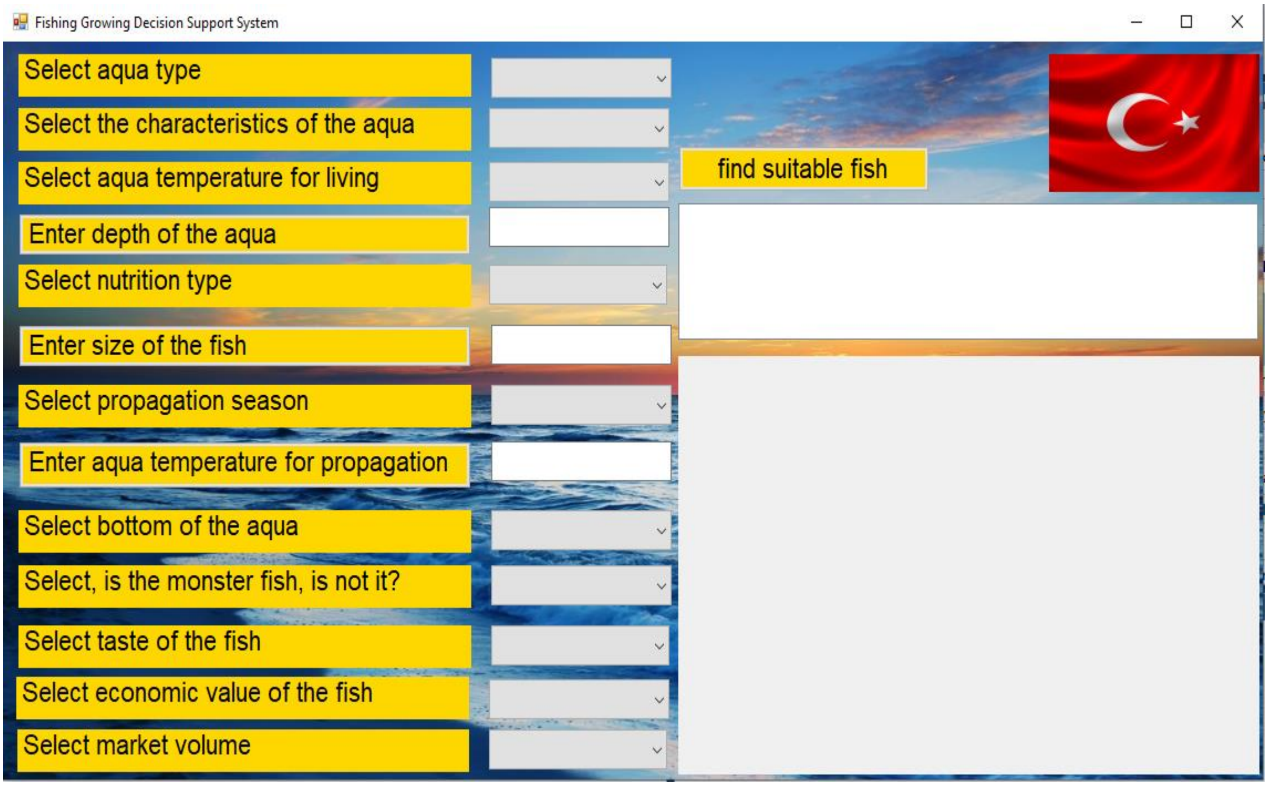

Figure 4. Developed user-interactive interface for fish farming

When the parameters in Figure 4 are selected and "find suitable fish" key is pressed, the system runs and presents the appropriate fish species with its photo.

4. Result and Discussion

In accordance with the indicated data and parameters, the classification operation was conducted using C&RT algorithm in SPSS Clementine 12.0 program on computer with 4 GB RAM and 2,3 GHz processor. As a result of the analysis, the accuracy rate was found as %79, 89. The accuracy rate indicates that the conducted classification is at a good level. A large DT was obtained as a result of the tree analysis; however, it is not possible to give the entire tree in the study. Thus, the first part of this tree is shown in Figure 5.

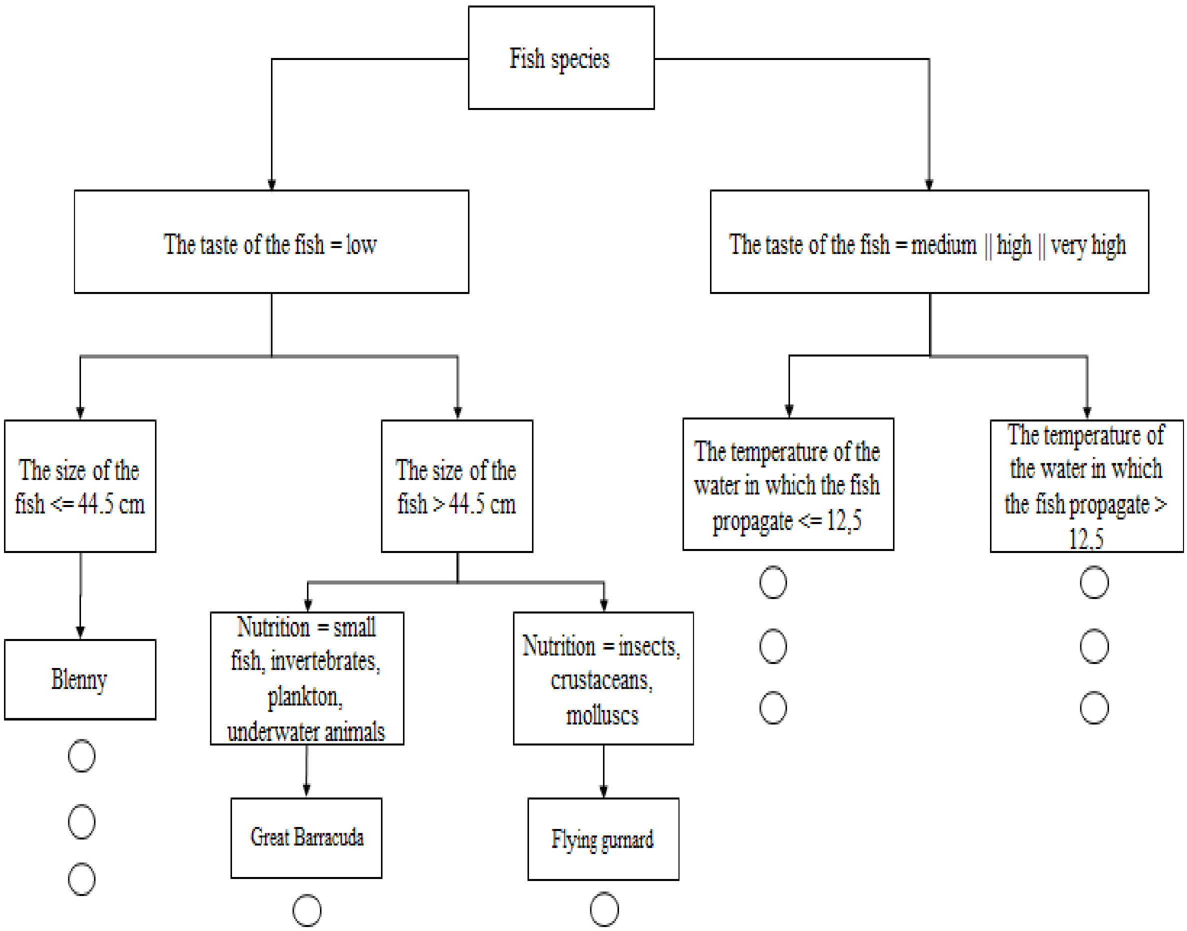

Figure 5. DT graph obtained by C&RT algorithm

Looking at Figure 5, the first branching was made according to "the taste of the fish". Reason for the selection of this factor is that it has the least Gini index value. Each branch is split to another factor taking into account the Gini index. Also, classification results as set of rules are illustrated in Table 2. The values in this table pave the way for people who will make fish farming, and offer appropriate fish species by evaluating the related farming area on the basis of the determined criteria.

For example, if the farmer wants to breed a fish species with a moderate and above-moderate taste value, the related water temperature is 12.5 °C in spring or summer months, the depth of the water is below 37.5 meters, if he / she wants the market status of the to-be-raised fish to be medium or high and the fish size to be between 9 and 15 cm. Table 2

Therefore, it is understood from the result that the most suitable fish breed for the region is picarel (spicara maena). The result is obtained when the parameters mentioned in the example are entered in the program (Figure 6).

According to the Figure 6, the appropriate fish species for the region and the photograph of the fish are shown in the interactive-interface. The interface has a very flexible usage and is open to development.

Information concerning the classification of fish

Condition	Fish species
The taste of the fish= low → fish size < 44,5	Blenny
The taste of the fish = low → fish size > 44,5 → Nutrition type = small fish, invertebrates, plankton, underwater animals	Great Barracuda
The taste of the fish = low → fish size > 44,5 → Nutrition type = insects, crustaceans, molluscus	Flying Gurnard
The taste of the fish = high and above → The temperature of the water in which the fish propagate <= 12,5 → Economic value of the fish = very high	Rainbow Trout
The taste of the fish= medium and above → The temperature of the water in which the fish propagate > 12,5 → The season in which the fish propagate = spring or summer → The depth of the water that fish live <= 60 m → Market situation = low or medium → Nutrition type = plankton, underwater animals, water plants	Spicara Smaris
The taste of the fish = medium and above → The temperature of the water in which the fish propagate > 12,5 → The season in which the fish are propagation = spring or summer → The depth of the water that fish live <= 60 m → Market situation = low or medium → Nutrition type = octopus vb. → Bottom of the water that the fish propagate = vegetation, gravel, mud, rocky, sandy, mussel shells → Nutrition type = mussel	Sarpa salpa
The taste of the fish = medium and above → The temperature of the water in which the fish propagate > 12,5 → The season in which the fish are propagation = spring or summer → The depth of the water that fish live <= 60 m → Market volume = medium or high → The depth of the water that fish live <= 37,5 m → fish size >9 cm & fish size <15,5 cm	Spicara maena

Although there are many fish species in the world; only 62 of them were included, which can be considered as the limitation of this study. Therefore, the obtained DT can produce results consistent with the aforementioned factors and the species included in this study. Without using any algorithms, it is almost impossible to take this many factors and fish species into account and determine the appropriate fish breeds in line with the indicated purposes. The study aims to enable decision makers to determine easily most suitable fish species for their features of fish farms. With the aid of the developed DSS, selection of fish breed for aquaculture bases upon a scientific foundation. Heightened awareness about fish farming and a contribution to the conservation of water resources are the expected results of the study in the long-run.

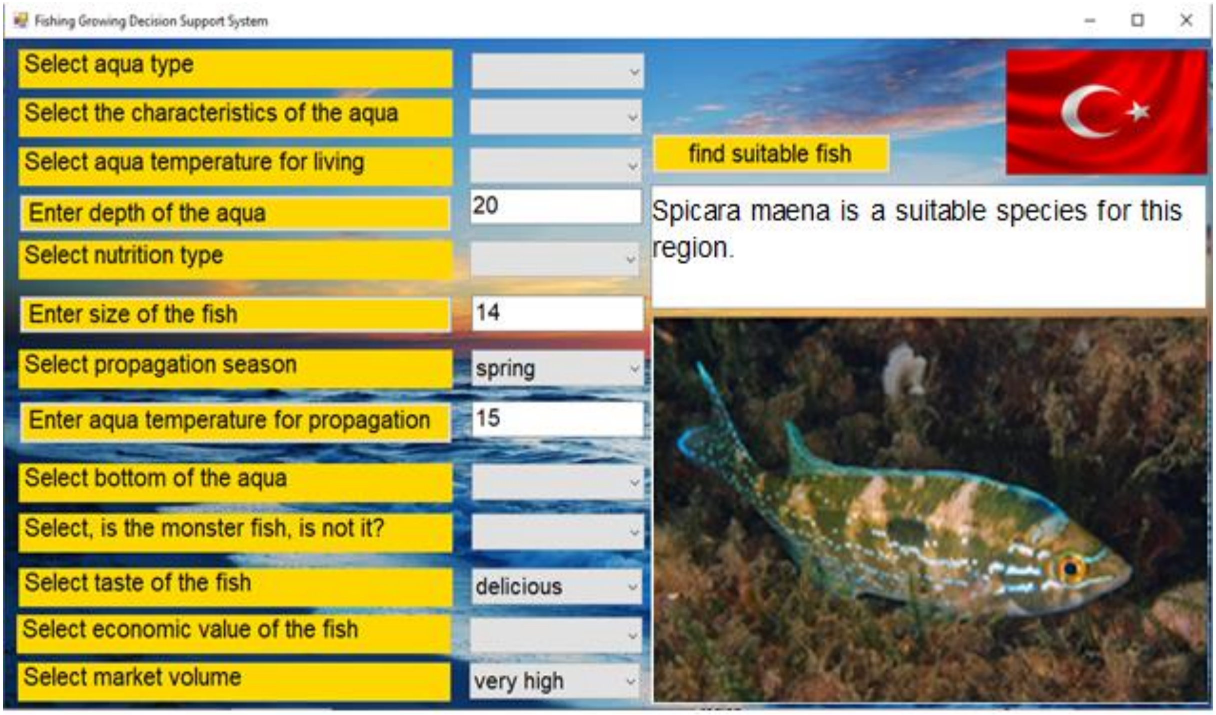

Figure 6. The result interface with recommendation box and fish photo

5. Conclusion

Fishing is an important source of livelihood and crucial subsector for many countries. Countries and fish farmers make serious investments in this sector; however, wrong investments cause great time and financial loss. Also, because every geography on Earth has different climatic characteristics and every fish species cannot be grown everywhere, decision making and selecting the right fish breed in aquaculture is a very difficult and time-consuming task without using any methodology. In this study, a decision support system was developed taking the features of fishing ground and the relevant geography into account to assist investors. Classification and Regression Trees algorithm, one of the data mining techniques, was used as a data analysis tool. Also, user-interactive interface was created so as to users benefit from the provided system easily. The study aims to enable decision makers to determine most suitable fish species for their features of fish farms easily.

The results show that the suggested system functions successfully with respect to enabling the fish farmers to identify appropriate species which can be bred in their farms. Thanks to the suggested system, fish farmers will not only increase their profitability by determining the most suitable fish breed, but also use existing water resources more efficiently. It is expected that foreign trade volume will be increased with the raising productivity and; hence, countries will create new business branches which will have reflections in employment figures in the long run. In this regard, the study exemplifies from the point of allowing governments and fish farmers to give right and more science-based decisions in a short time.

In future, it is thought that decision support systems can be developed by creating databases taking different fish species and factors into consideration. It is expected that classifications higher accuracy rates will be done using other DM and hybrid methods.

Contribution of Researchers

All authors were contributed equally to the article. The definition of problem, determination of solution methodology, application study and analysis of the results were realized by all authors simultaneously.

Confilict of Interest

The authors declared that there is no confilict of interest.

References

Aburrous, M., Hossain, M. A., Dahal, K., & Thabtah, F. (2010). Intelligent phishing detection system for e-banking using fuzzy data mining. *Expert Systems with Applications*, 37(12), 7913-7921. Doi: https://doi.org/10.1016/j.eswa.2010.04.044

Albayrak, A. S., & Yılmaz, Ş. K. (2009). Veri madenciliği: karar ağacı algoritmaları ve İMKB verileri üzerine bir uygulama [Data mining: decision tree algorithms and an application on ISE data]. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 14(1), 31-52. Retrieved from https://asalbayrak.files.wordpress.com/2011/12/a02.pdf

Alsmadi, M. K., Omar, K. B., & Noah, S. A. (2011). Fish classification based on robust features extraction from color signature using back-propagation classifier. *Journal of Computer Science*, 7(1), 52-58. Doi: https://doi.org/10.3844/jcssp.2011.52.58

Alsmadi, M. K. S., Omar, K. B., Noah, S. A., & Almarashdah, I. (2009). Fish recognition based on the combination between robust feature selection, image segmentation and geometrical parameter techniques using artificial neural network and decision tree. *Journal of Computer Science and Information Security*, 6(2), 215-221. Retrieved from https://arxiv.org/ftp/arxiv/papers/0912/0912.0986.pdf

Alsmadi, M. K., Omar, K. B., Noah, S. A., & Almarashdeh, I. (2010). Fish recognition based on robust features extraction from size and shape measurements using neural network. *Journal of Computer Science*, 6(10), 1059-1065. Doi: https://doi.org/10.3844/jcssp.2010.1088.1094

Angeli, C., Howard, S. K., Ma, J., Yang, J., & Kirschner, P. A. (2017). Data mining in educational technology classroom research: can it make a contribution?. *Computers & Education*, 113, 226-242. Doi: https://doi.org/10.1016/j.compedu.2017.05.021

Azadivar, F., Truong, T., & Jiao, Y. (2009). A decision support system for fisheries management using operations research and systems science approach. *Expert Systems with Applications*, 36, 2971–2978. Doi: https://doi.org/10.1016/j.eswa.2008.01.080

Bahari, T. F., & Elayidom, M. S. (2015). An efficient crm-data mining framework for the prediction of customer behaviour. *Procedia Computer Science*, 46, 725-731. Doi: https://doi.org/10.1016/j.procs.2015.02.136

Bhardwaj, B. K., & Pal, S. (2011). Data mining: A prediction for performance improvement using classification. *International Journal of Computer Science and Information Security*, 9(4), 136-140. Retrieved from: https://arxiv.org/ftp/arxiv/papers/1201/1201.3418.pdf

Breiman, L., Friedman, J.H., Olshen, R., & Stone, C.J. (1984). *Classification and regression tree*. Pacific California: Wadsworth & Brooks/Cole Advanced Books & Software.

Buelens, B., Pauly, T., Williams, R., & Sale, A. (2009). Kernel methods for the detection and classification of fish schools in single-beam and multibeam acoustic data. *ICES Journal of Marine Science*, 66(6), 1130-1135. Doi: https://doi.org/10.1093/icesjms/fsp004

Cabreira, A. G., Tripode, M., & Madirolas, A. (2009). Artificial neural networks for fish-species identification. *ICES Journal of Marine Science*, 66(6), 1119-1129. Doi: https://doi.org/10.1093/icesjms/fsp009

Çalış, A., Kayapınar, S., & Çetinyokuş, T. (2014). Veri madenciliğinde karar ağacı algoritmaları ile bilgisayar ve internet güvenliği üzerine bir uygulama [An application on computer and internet security with decision tree algorithms in data mining]. *Journal Of Industrial Engineering (Turkish Chamber Of Mechanical Engineers)*, 25(3-4), 2-19. Retrieved from: https://www.mmo.org.tr/sites/default/files/ab225c69b017f13_ek.pdf

Daş, B., & Türkoğlu, İ. (2014). *DNA dizilimlerinin sınıflandırılmasında karar ağacı algoritmalarının karşılaştırılması*. Elektrik–Elektronik–Bilgisayar ve Biyomedikal Mühendisliği Sempozyumu (ELECO 2014), 381-383, Bursa. Retrieved from: https://www.emo.org.tr/ekler/91ac323bf2a22ad_ek.pdf

David, H., Heikki, M., & Padhraic, S. (2001). *Principles of data mining*. Cambridge: The MIT Press.

Delen, D., Fuller, C., McCann, C., & Ray, D. (2009). Analysis of healthcare coverage: a data mining approach. *Expert Systems with Applications,* 36(2), 995-1003. Doi: https://doi.org/10.1016/j.eswa.2007.10.041

Dutta, I., Dutta, S., & Raahemi, B. (2017). Detecting financial restatements using data mining techniques. *Expert Systems with Applications,* 90, 374-393. Doi: https://doi.org/10.1016/j.eswa.2017.08.030

Emel, G. G., & Taşkın, Ç. (2005). Veri madenciliğinde karar ağaçları ve bir satış analizi uygulaması [Decısıon trees in data mining and a sales analysis application]. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi,* 6(2), 221-239. Retrieved from: https://dergipark.org.tr/tr/download/article-file/113054

FAO 2016. Assessing water availability and economic, social and nutritional contributions from inland capture fisheries and aquaculture. *FAO fisheries and aquaculture technical paper. FAO*, 101, Rome. Retrieved from: http://www.fao.org/3/a-i5878e.pdf

Gaertner, D., & Dreyfus-Leon, M. (2004). Analysis of non-linear relationships between catch per unit effort and abundance in a tuna purse-seine fishery simulated with artificial neural networks. *ICES J. Mar. Sci.,* 61 (5), 812– 820. Retrieved from: https://academic.oup.com/icesjms/article/61/5/812/867451

Gerami, M., H., & Rabbaniha, M. (2018). Forecasting the anchovy kilka fishery in the caspian sea using a time series approach. *Turkish Journal of Fisheries and Aquatic Sciences,* 18(11), 1288-1292. Doi: https://doi.org/10.4194/1303-2712-v18_11_05

Gutiérrez-Estrada, J. C., Yáñez, E., Pulido-Calvo, I., Silva, C., Plaza, F., & Bórquez, C. (2009). Pacific sardine (Sardinops sagax, Jenyns 1842) landings prediction. A neural network ecosystemic approach. *Fisheries Research*, 100(2), 116-125. Doi: https://doi.org/10.1016/j.fishres.2009.06.014

Halide, H., Stigebrandt, A., Rehbein, M., & McKinnon, A.D. (2009). Developing A decision support system for sustainable cage aquaculture. *Environmental Modelling & Software*, 24, 694–702. Doi: https://doi.org/10.1016/j.envsoft.2008.10.013

Han, J., & Micheline, K. (2006). *Data mining: concepts and techniques (second edition)*. San Francisco, ABD: Morgan Kaufmann Publishers.

Haralabous, J., & Georgakarakos, S. (1996). Artificial neural networks as a tool for species identification of fish schools. *ICES Journal of Marine Science,* 53, 173–180. Doi: https://doi.org/10.1006/jmsc.1996.0019

Hong-Chun, Y., Ying, L., & Ying, C. (2009, August). *Fishery knowledge discovery based on svm and fuzzy rule extraction*. In Computer Science and Information Technology, 2009. ICCSIT 2009, 2nd IEEE International Conference on IEEE, 167-171, Beijing. Retrieved from: https://www.infona.pl/resource/bwmeta1.element.ieee-art-000005234379

Hu, J., Li, D., Duan, Q., Han, Y., Chen, G., & Si, X. (2012). Fish species classification by color, texture and multi-class support vector machine using computer vision. *Computers and Electronics In Agriculture*, 88, 133-140. Retrieved from https://www.sciencedirect.com/science/article/pii/S0168169912001937

Joo, R., Bertrand, S., Chaigneau, A., & Niquen, M. (2011). Optimization of an artificial neural network for identifying fishing set positions from vms data: an example from the peruvian anchovy purse seine fishery. *Ecological Modelling,* 222(4), 1048-1059. Doi: https://doi.org/10.1016/j.ecolmodel.2010.08.039

Joy, M.K., & Death, R.G., (2004). Predictive modelling and spatial mapping of freshwater fish and decapod assemblages using gis and neural networks*. Freshwater Biology*, 49(8), 1036–1052. Doi: https://doi.org/10.1111/j.1365-2427.2004.01248.x

Kaygisiz, F., & Eken, M. (2018). A research on determination of fish marketing margins in istanbul province of turkey. *Turkish Journal of Fisheries and Aquatic Sciences*, 18, 801-807. Retrieved from http://www.trjfas.org/uploads/pdf_1223.pdf

Kim, Y.-H. (2003). New modelling of complex fish migration by application of chaos theory and neural network. *Journal of Fish Biology*. 63 (1), 234. Doi: https://doi.org/10.1111/j.1095-8649.2003.0216s.x

Kvyatkovskaya I.Y., Kosmacheva I., Sibikina I., Galimova L., Rudenko M., & Barabanova E.A. (2017). *Modular structure of data processing in* *automated systems of risk management in the fisheries industry*. In: Kravets A., Shcherbakov M., Kultsova M., Groumpos P. (eds) Creativity in Intelligent Technologies and Data Science (CIT&DS 2017). Communications in Computer and Information Science, 754, Volgograd, Russia: Springer, Cham. Doi: https://doi.org/10.1007/978-3-319-65551-2_21

Larose, D. T. (2004). *Discovering knowledge in data: an introduction to data mining*. New Jersey, USA: John and Wiley Sons Incorporated.

Larsen, R., Olafsdottir, H., & Ersbøll, B. K. (2009). *Shape and texture based classification of fish species.* In Scandinavian Conference on Image Analysis, 745-749, Springer, Berlin, Heidelberg. Retrieved from: https://link.springer.com/chapter/10.1007/978-3-642-02230-2_76

Marzuki, M. I., Garello, R., Fablet, R., Kerbaol, V., & Gaspar, P. (2015). Fishing gear recognition from vms data to identify illegal fishing activities in indonesia. *In OCEANS 2015-Genova* IEEE, 1-5. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7271551

Marzuki, M. I., Gaspar, P., Garello, R., Kerbaol, V., & Fablet, R. (2017). Fishing gear identification from vessel-monitoring-system-based fishing vessel trajectories. *IEEE Journal of Oceanic Engineering*, 99, 1-11. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7990476

Mastrorillo, S., Lek, S., Dauba, F., & Belaud, A. (1997). The use of artificial neural networks to predict the presence of small-bodied fish in a river. *Freshwater Biol.*, 38, 237–246. Doi: https://doi.org/10.1046/j.1365-2427.1997.00209.x

Mendoza, M., García, T., & Baro, J. (2010). Using classification trees to study the effects of fisheries management plans on the yield of merluccius merluccius (linnaeus, 1758) in the alboran sea (western mediterranean). *Fisheries Research*, 102(1-2), 191-198. Doi: https://doi.org/10.1016/j.fishres.2009.11.012

Ngai, E. W., Xiu, L., & Chau, D. C. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, 36(2), 2592-2602. Doi: https://doi.org/10.1016/j.eswa.2008.02.021

Ogunlana, S. O., Olabode, O., Oluwadare, S. A. A., & Iwasokun, G. B. (2015). Fish classification using support vector machine. *African Journal of Computing & ICT*, 8(2), 75-82. Retrieved from https://www.semanticscholar.org/paper/Fish-Classification-Using-Support-Vector-Machine-Ogunlana-Olabode/42596f3236d03f407abc04ed344264eadd671839

Park, Y., Grenouillet, G., Esperance, B., & Lek, S. (2006). Stream fish assemblages and basin land cover in a river network. *Sci. Total Environ.*, 365 (1–3), 140–153. Retrieved from https://www.sciencedirect.com/science/article/pii/S004896970600180X

Reyjol, Y., Fischer, P., Lek, S., Rosch, R., & Eckmann, R. (2005). Studying the spatiotemporal variation of the littoral fish community in a large prealpine lake, using self-organizing mapping. *Can. J. Fish. Aquat. Sci.*, 62 (10), 2294–2302. Retrieved from https://kops.uni-konstanz.de/bitstream/handle/123456789/7240/Reyjol_etal_CJFAS_2005.pdf?sequence=1&isAllowed=y

Sholahuddin, A., Ramadhan, A., & Supriatna, A. (2015). The application of ann-linear perceptron in the development of dss for a fishery industry. *Procedia Computer Science*. 72. 67-77. Doi: https://doi.org/10.1016/j.procs.2015.12.106

Song, Y. Y., & Ying, L. U. (2015). Decision tree methods: applications for classification and prediction. *Shanghai archives of psychiatry*, 27(2), 130-135. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4466856/

Storbeck, F., & Daan, B. (2001). Fish species recognition using computer cision and a neural network. *Fisheries Research*, 51(1), 11-15. Doi: https://doi.org/10.1016/S0165-7836(00)00254-X

Su, Y. Y., & Chang, S. J. (2008). Spatial cluster detection for the fishing vessel monitoring systems. *In OCEANS 2008-MTS/IEEE Techno-Ocean*, 1-4, Kobe. IEEE. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4531048

Su Ürünleri Raporu. (2014). Retrieved from http://www.isub.org.tr/assets/rapor_suurunlerivekulturbalikciligiileilgilirevize_3eylul2014.pdf

Sucipto, S., Kusrini, K., & Taufiq, E. L. (2016). *Classification method of multi-class on C4.5* *algorithm for fish diseases*. 2nd International Conference on Science in Information Technology (ICSITech) IEEE, 5-9, Balikpapan, Indonesia. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7852598

Sun, L., Li, S., Xiao, H., & Yang, D. (2009). Application of operation research and system science approach to fisheries management. *IFIP International Federation for Information Processing*, 294. 1359-1368. Retrieved from https://link.springer.com/chapter/10.1007/978-1-4419-0211-5_64

Supriatna, A., Ramadhan, A., & Husniah, H. (2015). A decision support system for estimating growth parameters of commercial fish stock in fisheries industries. *Procedia Computer Science*, 59 ,331-339. Retrieved from https://core.ac.uk/reader/82472635

Supriatna, A., Sholahuddin, A., Ramadhan, A., & Husniah, H. (2016). *SOFish ver. 1.2 - A decision support system for fishery managers in managing complex fish stocks*. IOP Conference Series: Earth and Environmental Science, 31. Retrieved from https://iopscience.iop.org/article/10.1088/1755-1315/31/1/012005/pdf

Suryanarayana, I., Braibanti, A., Rao, R., S., Ramam, V., A., Sudarsan, D., & Rao, G., N. (2008). Neural networks in fisheries research. *fisheries research*, 92, 115–139. Doi: https://doi.org/10.1016/j.fishres.2008.01.012

Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining*. England: First Edition, Pearson.

Tarım ve Orman Bakanlığı su ürünleri istatistikleri. (2019). Retrieved from https://www.tarim.gov.tr/sgb/Belgeler/SagMenuVeriler/BSGM.pdf

Tsai, W. P., Huang, S. P., Cheng, S. T., Shao, K. T., & Chang, F. J. (2017). A data-mining framework for exploring the multi-relation between fish species and water quality through self-organizing map. *Science of the Total Environment*, 579, 474-483. Doi: https://doi.org/10.1016/j.scitotenv.2016.11.071

Uçar, U. U., Balo, F., Eraslan, G., & Çetin, B. (2017). İdeal hayvan yetiştiriciliği için veri madenciliğine dayalı bir kds çalışması [A KDS study based on data mining for ideal animal breeding]. *Gaziosmanpaşa Bilimsel Araştırma Dergisi, 6(Özel Sayı (ISMSIT2017))*, 133-141. Retrieved from https://dergipark.org.tr/tr/download/article-file/399649

Vazan, P., Janikova, D., Tanuska, P., Kebisek, M., & Cervenanska, Z. (2017). Using data mining methods for manufacturing process control. *IFAC-PapersOnLine*, 50(1), 6178-6183. Doi: https://doi.org/10.1016/j.ifacol.2017.08.986

Wang, K. Y., Ting, I. H., & Wu, H. J. (2013). Discovering interest groups for marketing in virtual communities: an integrated approach. *Journal of Business Research*, 66(9), 1360-1366. Doi: https://doi.org/10.1016/j.jbusres.2012.02.037

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., & Zhou, Z. H. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1), 1-37. Retrieved from https://link.springer.com/article/10.1007/s10115-007-0114-2

Xiaoshuan, Z., Tao, H., Revell, B., & Zetian, F. (2005). A forecasting support system for aquatic products price in china. *Expert Systems with Applications*, 28, 119–126. Doi: https://doi.org/10.1016/j.eswa.2004.08.012

Yang, Y. C. E., Cai, X. & Herricks, E. E. (2008). Identification of hydrologic indicators related to fish diversity and abundance: a data mining approach for fish community analysis. *Water Resources Research*, 44(4). Doi: https://doi.org/10.1029/2006WR005764

Yıldız, Y., Dağdeviren, M., & Çetinyokuş, T. (2008). İşgören performansının değerlendirilmesi için bir karar destek sistemi ve uygulaması [A decision support system to evaluate employee performance and its application]. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 23(1), 239-248. Retrieved from https://dergipark.org.tr/en/download/article-file/75552

Yuan, H., Li, Y., & Chen, Y. (2009). *A novel knowledge discovery model for fishery forecasting*. In Fuzzy Systems and Knowledge Discovery, FSKD'09, Sixth International Conference on IEEE, 29-33, Tianjin, China. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5358886

Zhou, S. (2003). Application of artificial neural networks for forecasting salmon escapement. *North American Journal of Fisheries Management*, 23 (1), 48–59. Retrieved from https://www.tandfonline.com/doi/abs/10.1577/1548-8675(2003)023%3C0048%3AAOANNF%3E2.0.CO%3B2