EXPERT BUSINESS MODEL DESIGN IN LIVESTOCK

Muammer Akcay, Eyyup Gulbandilar and Demet Bektas

Dumlupinar University, Faculty of Engineering, Department of Computer Engineering, Kutahya, Turkey

Abstract

Small scale businesses in the livestock sector, inefficiencies, and inability to use technology are faced with problems such as market-oriented production and inability to organize. In this study, for the development of livestock breeds in Turkey had revised the entire value chain, measures and incentives throughout the entire value chain is to help in the regulation. Software systems are established by considering data structure of small scale business by considering software systems of expert systems of data subsets in order to get better livestock sector. In this study, data provided by 90 people is used, 80 of these data for training and 10 of these data for testing were selected. The decision tree had been constructed by using ID3 algorithm with training data set. The software for developed algorithm by considering the decision tree algorithm was developed by using C# programming language. The developed software had been tested on ten samples. It was found correct for eight samples and incorrect for two samples.

Keywords: *Livestock, Meat Products, ID3 Algorithm, Entropy, Decision tree, C#*

1. INTRODUCTION

As all the world enough and livestock sector is vital to the balanced nutrition of the population increased in Turkey, producing high value-added industries as raw materials in many areas. The livestock sector also provides an important contribution to the national economy as well as due to host many sectors. Global livestock sector generates 40% of the agricultural gross domestic product and provides livelihoods for 1.3 billion people. Turkey has suitable properties and potential for the production of all kinds of animal products in terms of geographical features. However, animals should not be enough and sufficient quantities of cheap breeding race and failed to protect livestock as a result of the current situation cannot be made of quality fodder cultivation. Thus, there has been a reduction in the animal existence. Other challenges facing the development of livestock breeding in Turkey is the lack of material and high-quality livestock fodder deficit. Fluctuations in raw material prices in the global space industry non-competitive due to high input costs and instability is emerging as another important issue. Reduction in the supply of meat in the presence of animals has led to decline. Turkey meat consumption per person was difficult to meet domestic consumption despite being far behind the developed countries. This situation led to an increase in the price of registration, such as informality and unfair competition is an obstacle to the development of the sector, together with other issues. Due to large-scale modern farming businesses will change the situation overall in size and although the nature of bottlenecks in integrating established in recent years in our country is experiencing difficulties in coming to the desired level.

Traditional statistical techniques are frequently used in classification problems. But any who do not have and made a number of assumptions that are determined to be stronger than the classification performance by operating Data Mining interest in the technology is increasing every day. Livestock Expert Business Model Design data mining techniques in the study of decision trees technique introduced, the business model has been developed (Negnevitsky, 2002).

Data mining previously unknown patterns hidden in the data is a dynamic process that allows revealing and predicting the future by obtaining useful results from this information. Analyze the non-linear relations in recent years and the need for traditional statistical methods and assumptions, such as the data mining techniques are in great demand. Decision Tree method, to produce easily understood rules, rules can be visualized and one of the most commonly used technique by the data mining to have an appropriate structure features powerful algorithms to different variables (Dogan and Ozdamar,

2003). Nonlinear, multidimensional, noisy, complicated, imperfect solution to the lack of data and problems, especially a lack of mathematical models and algorithms used in the case of the data mining techniques which could be used for many scientific fields. The expert system to get better data subsets that have been submitted by considering engineering analysis method systematic studies and engineering disciplines based businesses with an easy-to-read on a small scale model of the software system has been developed in the livestock sector. This study will provide solutions to the problems faced by enterprises and system aimed at making an Expert Business Model Design will lead to better animal husbandry. Problems faced by businesses are to cause fluctuations in the rate of sales of seasonal changes, the lack of demand for meat products, the negative supply, depending on the climatic conditions reduced the increase in price as a result, income / expense statements, depending on the small business labor productivity cannot be obtained any profit.

In the next section, the main problem of businesses was determined by transferring scientific solutions to the general scale. Within the scope of evaluation technical information, data mining and allocation model through expert systems (allocation pattern) was used. Then, by using decision trees method with the ID3 algorithm, “Information Calculation'” and “Gain” formulas were calculated. ID3 algorithm has been added to interpreting the decision tree branching. In the last section, the requirement analysis is applied to Naive Bayes classifier and pattern recognition problem. An application is developed by C# programming language.

2. EXPERT SYSTEMS

Acting as experts in solving a specific problem programs called expert systems. Expert systems without the need for processing the data are described as a transition to processing information. Algorithms and inference mechanisms have been replaced in traditional programs and expert systems Expert systems are knowledge based systems to examine problems in a broader perspective and to add an artificial intelligence application that aims to imitate human intelligence in the solution. Expert system is a consultant program which aim to support expert knowledge and inference process (Allahverdi, 2002, Poyrazoglu et al., 1997). Expert systems are based on information compiled from real people in a particular area of expertise in the software are the ability to improve themselves by the time. Very different methods are used to imitate the behavior specialist. The most commonly used methods are representing information and knowledge engineering methods. Each expert system may not be learning behavior all the time. After all the expert system is developed, the same conclusion should be reached for real problems.

1. Knowledge base: contains the necessary information for problem solving. Information is kept a collection of rules in the rule-based expert system. Each rule expresses a relationship, an order, a strategy or a means of intuition. Rules are structure of IF - THEN (condition - action). If the first part of the rule is verified (It is said that the rules are excited), the second part is applied.

2. Database: contains facts that match the first part of the rules in the knowledge base.

3. Inference engine: makes the inference process (matches the first part of the rule with facts, then applies to the second part).

4. Explanation facilities: the system, how to reach a conclusion that access to that user (inference chain), representation / description feature.

Characteristics of Expert Systems: It is expected to work with performance close to a human expert in narrow and areas requiring expertise. Solutions are expected in the desired time interval. Heuristic rules can be used to narrow the search space. They can inference symbolic. Symbols can be used to represent different types of information. Expert people rarely make mistakes. Similarly, expert systems make fewer mistakes. The source of the error is expected from one of the components in the overall structure.

Uses of Expert System: Using the commonly used if need. It is used in case of decision-making situation is complex (a simple computer program can be used for simple cases). If a decision-making

logic can be converted to a rule hierarchy, it can be used. When the review process will takes a very long time, it is used in a large number of possible combinations of problems. When meaning to large data or information retrieval process is required, it is used. Expert systems are used for application proposals, classification, diagnosis, review, description; choose the solution, evaluation of the situation and focuses to estimate.

Roles in the Development of Expert System: Expert system model is shown in Figure 1. Expert systems are knowledge-based systems as shown in Figure 2. By studying the problem in a broader perspective, the solution is an application of artificial intelligence that aims to imitate human intelligence (Allahverdi, 2002). Expert systems as well as looking at a different view to the current system, used in a more systematic algorithm in terms of applicability and a different language (Visual Studio 2012 Application C #) with less steps, the easier and developed useful software.

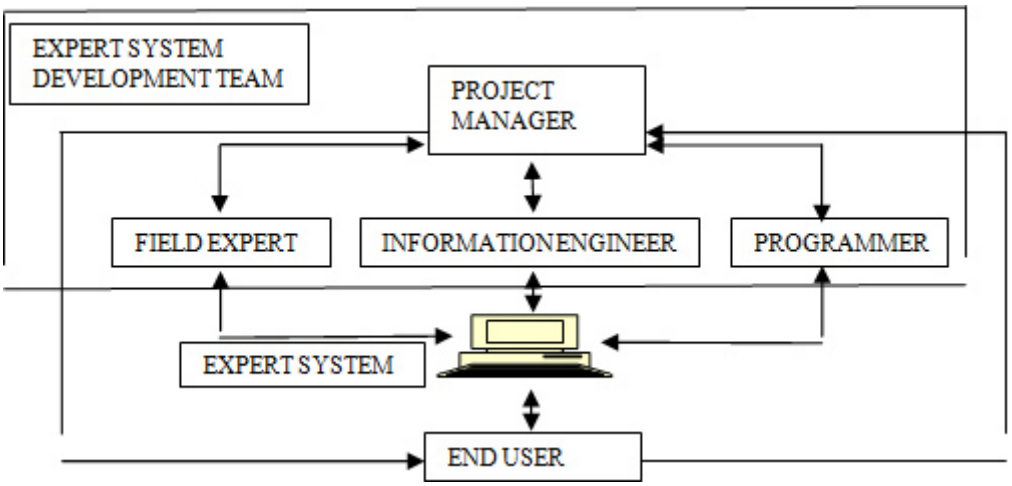
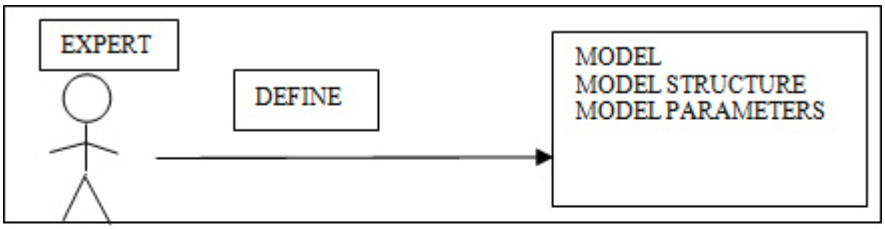
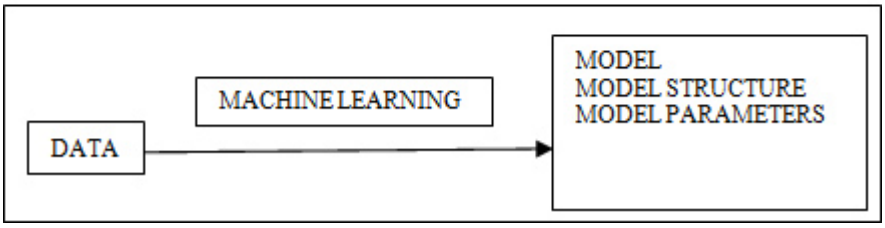


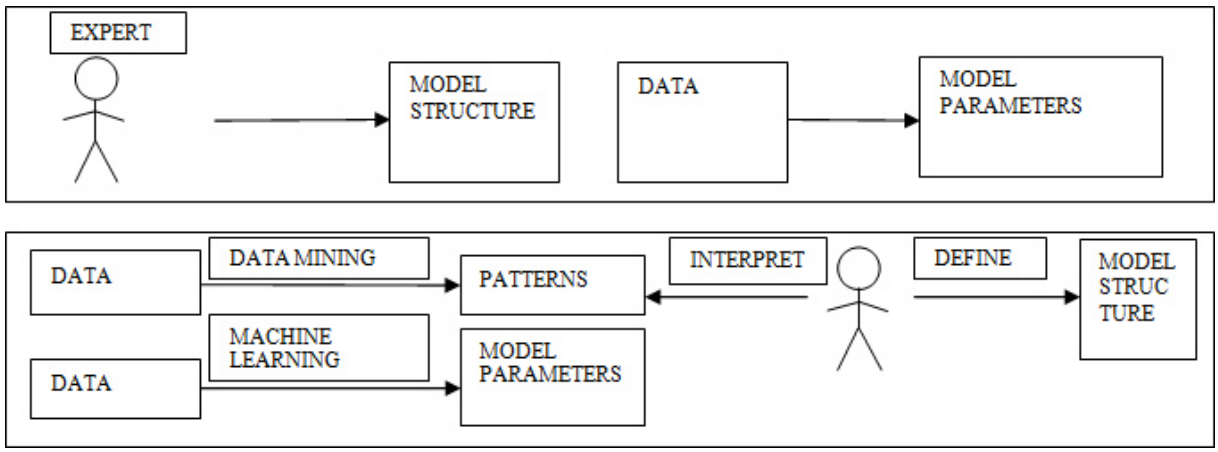
Figure 1. Expert system model.

Four approaches to construct a knowledge base:





(a)



(b)

Figure 2. (a) Construction of knowledge base, (b) Construction of knowledge base.

Benefits of Expert Systems: It provides an expertise where expertise is rare and expensive. It provides an expertise when and experts cannot be reached. It provides an expertise where experts are located in inaccessible places. The analysis will be used to shorten the time of a specialist. It is used to increase performance of less experienced staff. Expertise analysis can be done when experts cannot find the time to analyze (to analyze large amount or consider every possibility of data before making a decision). It is used to obtain and store information when reassignment, retirement or staff loss of expertise. It is used to guarantee to give expertise outcome which objective and consistent for each case (For example, to consistent for a company policy). The best experts are used for the most difficult problems or other important tasks.

3. ALGORITHM DESIGN

3.1 Implementation of Classification Method

Decision trees is a collection of rules to find meaningful statistically significant groups, give answers clearly by tree diagrams with easy to read, classify observations or estimate (Dogan and Ozdamar, 2003). Decision tree represents the division of data set in a sequential order to maximize differences on the dependent variable (Dormen, 2003). Inputs and outputs which are used in the decision tree algorithm, that help to classify according to certain variables are determined variables of the data. Decision tree algorithm discovers the input data variables with variable data structure for output data variables (Tan et al., 2006). Decision tree is an effective classification method. There are a large number of decision tree reasoning algorithms in the machine learning and applied statistics. These algorithms are supervised learning methods by using a form of a serial input-output set. A typical decision tree learning system uses top-down method which finds a solution in some part of the research space. This method guarantees to be able to find a simple tree. A decision tree contains nodes where variables are tested. Branches from a node to open outsides respond all the possible results of the test in the node.

3.2 General Structure of ID3 Algorithm

ID3 algorithm is one of the most frequently used algorithms in decision tree. ID3 algorithm is an example of Symbolic Learning and Rule Induction. It is also a supervised learner which means it looks at examples like a training data set to make its decisions. It was developed by J. Ross Quinlan in 1979. It is a decision tree that is based on mathematical calculations.

A decision tree algorithm classifies data by using its attributes. It is upside down. The tree has decision nodes and leaf nodes. The leaf node has homogenous data which means further classification is not

necessary. ID3 algorithm builds similar decision trees until all the leaf nodes are homogenous (http://www.scism.lsbu.ac.uk/inmandw/ir/irtopics/c1.doc, May, 2015).

Entropy refers to homogenous of the data. It ranges between 0 and 1. Data sets with entropy 1 means it is as random as it can get. A data set with entropy 1 means that it is homogenous. It has high entropy which means the data is random. Entropy is calculated by the formula:

$$H(S) = -\sum_{x \in X} p(x).\log_2 p(x) \quad \text{(Equation.1)}$$

where, S- The current (data) set for which entropy is being calculated (changes every iteration of the ID3 algorithm), X- Set of classes in S, p(x) - The proportion of the number of elements in class x to the number of elements in set S (http://www.scism.lsbu.ac.uk/inmandw/ir/irtopics/c1.doc, May, 2015, http://en.wikipedia.org/wiki/ID3_algorithm, May, 2015).

Information gain (IG (A)) is the measure of the difference in entropy from before to after the set S is split on an attribute A. In other words, how much uncertainty in S was reduced after splitting set S on attribute A.

$$IG(A,S) = H(S) - \sum_{x \in X} p(t).\log_2 H(t) \qquad \text{(Equation 2)}$$

Where H(S)- Entropy of set S. T - The subsets created from splitting set S by attribute A . p(t) - The proportion of the number of elements in t to the number of elements in set S. H(t) - Entropy of subset t. In ID3, information gain can be calculated (instead of entropy) for each remaining attribute. The attribute with the largest information gain is used to split the set S on this iteration. ID3 chooses the attribute with the highest gain to create nodes in the decision tree. If the resulting subsets do not have entropy zero or equal to zero then it chooses one of the remaining attribute to create further nodes until all the subsets are homogenous (http://en.wikipedia.org/wiki/ID3_algorithm, May, 2015).

3.3 Calculation According to the Survey Results

We are collected dataset in which 90 people at Dumlupinar University, Evliya Celebi Campus in Kutahya, Turkey, between 2014 and 2015 years. This dataset contains same information such as age (young, middle aged and old), selection criteria (price, appearance, taste, advertisement, presentation, transportation and date), animal type (cattle, sheep), meat product (entrecote, steak, tenderloin, grill, ground beef, flaked and chop) and presentation form (barbeque grill and cook at home) (as an output sample). We are randomly selected 80 people dataset for training and the remaining 10 dataset used for testing.

ID3 tree was requested for a dataset. First step, the entropy is calculated by using Equation 1 (Dormen, 2003, Tan et al., 2006, Altas and Gulpinar, 2012, https://weka.waikato.ac.nz). While the information gain is calculated, all of the data in the dataset is used. This will be called sampling. The calculation of this sampling is done for all the dataset.

After calculating the information value for each sample, gain can be calculated. Generally, information is fragmented and the process is done on the divided part (partition). Finally, gain can be calculated by using Equation 2. For each sample, these calculations are repeated and gain the highest (maximum) sample is determined as the root of the tree.

All features are controlled in each step. Normalized information gain (gain information) for each feature is calculated. The feature which gives the best information gain is carried as a decision in the decision tree. Then, the decision tree is built by creating a sub-list under the new decision node

By using ID3 algorithm, the calculation of the survey was carried out. The result is shown in Table 1.

Table 1. Decision point according to the meat product.

Property	Value (Bit)
Age	0.049
Sex	0.04
Selection Criteria	0.064
Animal Type	0
Meat Product	**0.192**
Service Type	0.0696

The maxim value in the Table 1 is 0.192 for meat products. The root of the decision tree is meat product.

Table 2. Decision point Value (Bit) according to meat product choices (Entrecote, Steak, Tenderloin, Grill, Ground Beef, Flaked, and Chop)

Property	Entrecote	Steak	Tenderloin	Grill	Ground Beef	Flaked	Chop
Age	0.0735	0.065	0.403	0.097	0.11	0.0946	0.173
Sex	0.311	0.002	0.225	0.025	0.106	0.072	0.009
Selection Criteria	**0.561**	**0.235**	**0.764**	**0.403**	**0.109**	**0.277**	**0.42**
Animal Type	0.059	0.002		0.081	0.006	0.0254	0.174
Service Type	0.109	0.155		0.093	0		0.023

The maxim value for each column in the Table 2 is selection criteria. The root of the decision tree is selection criteria.

The algorithm of the program follows by taking into account decision tree at the Annex

if meat product=entrecote and SC=1 or SC=4 then H
if meat product=entrecote and SC=2 or SC=3 or SC=5 or SC=6 then BG
if meat product=steak and SC=1 then H
if meat product=steak and SC=4 then BG
if meat product=steak and SC=2 or SC=3 or SC=5 and Age=O then H
if meat product=steak and SC=2 or SC=3 or SC=5 and Age=Y or Age=M then BG
if meat product=tenderloin and SC=2 then BG
if meat product= tenderloin and SC=2 or SC=3 or SC=5 then H
if meat product=grill and SC=3 then H
if meat product= grill and SC=1 or SC=2 or SC=4 or SC=6 then BG
if meat product= ground beef and SC=1 then BG
if meat product= ground beef and SC=2 or SC=3 or SC=6 and Age=Y then BG
if meat product= ground beef and SC=2 or SC=3 or SC=6 and Age=M or Age=O then H
if meat product= flaked and SC=5 then H
if meat product= flaked and SC=1 or SC=2 or SC=3 and Age=M or Age=O then H

if meat product= flaked and SC=1 or SC=2 or SC=3 and Age=Y and Animal type=Sheep then BG
if meat product= flaked and SC=1 or SC=2 or SC=3 and Age=Y and Animal type=Cattle then H
if meat product= chop and SC=1 or SC=2 or SC=3 then BG
if meat product= chop and SC=4 then H

The software was developed in C# programming language by using this algorithm. The developed software program interface is shown in Figure 3. 10 data sets for testing are applied for the developed software. It has been found correct result for eight samples and wrong results for two samples.

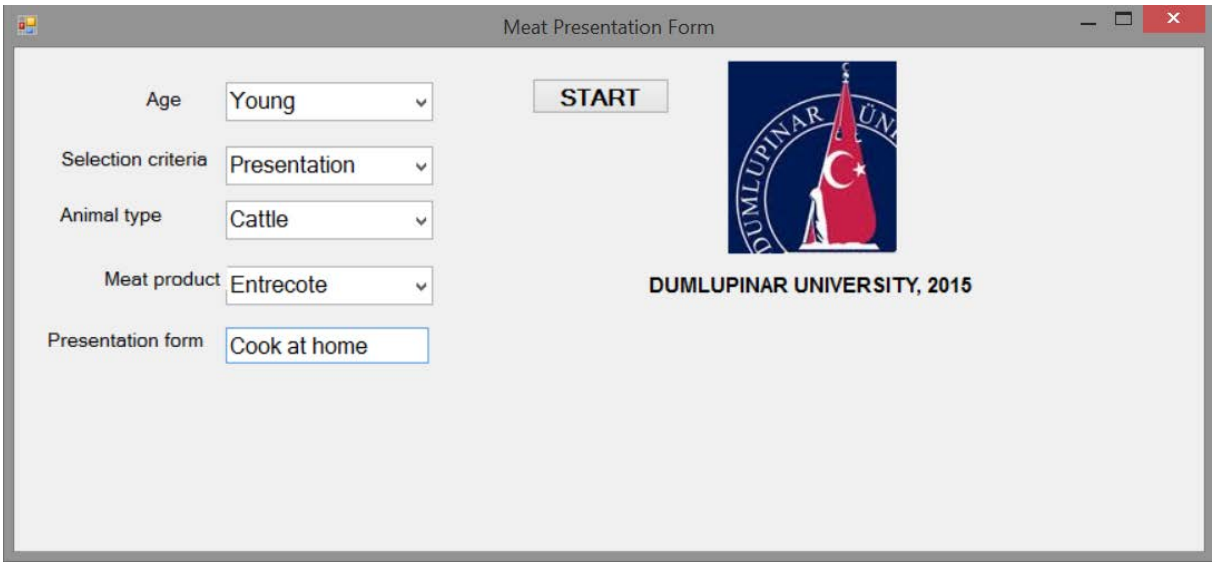


Figure 3. Interface of the developed software

4. CONCLUSIONS

The developed software is a prototype work to determine the type of meat consumption habits of the customers and it is worked on a limited number of data sets. Although a limited number of data set is used, it has very successful performance demonstrated on the test data set. In the future, the more successful results will be achieved by increasing the amount of data sets in the coming years.

REFERENCES

Negnevitsky, M., 2002, "Articial Intelligence: A Guide to Intelligent Systems", Pearson, 2002. http://www.pcai.com/web/ai_info/expert_systems.html.

Dogan, N., Ozdamar, K., 2003, "Chaid Analizi ve Aile Planlaması ile Bir Uygulama", T. Klin Tıp Bilimleri Dergisi, Cilt: 23, Eskisehir, pp.392-397.(In Turkish)

Allahverdi N., 2002, "Uzman Sistemler", Atlas Yayin, Istanbul. (In Turkish)

Poyrazoglu O., Gulesin M., Kayir Y., 1997, "Uzman Sistem Tekniklerine Dayali Alisilmamis imalat Yontemlerinin Tespiti", Mak. Tas. ve iml. Yon. Knf., Gazi Universitesi, sayfa 80-89. (In Turkish)

Dormen, D., 2003, "Bankacilik Sektorunde Müsteri İliski Yontemi: CRM Acisindan Veri Madenciligi Yontemi" Doktora tezi, istanbul pp.34. .(In Turkish)

Tan, P. N., Steinbach, M., Kumar V., 2006, "Introduction to Data Mining", Pearson Addison Wesley, Boston, pp.146.

http://www.scism.lsbu.ac.uk/inmandw/ir/irtopics/c1.doc (Last access date: May, 2015).

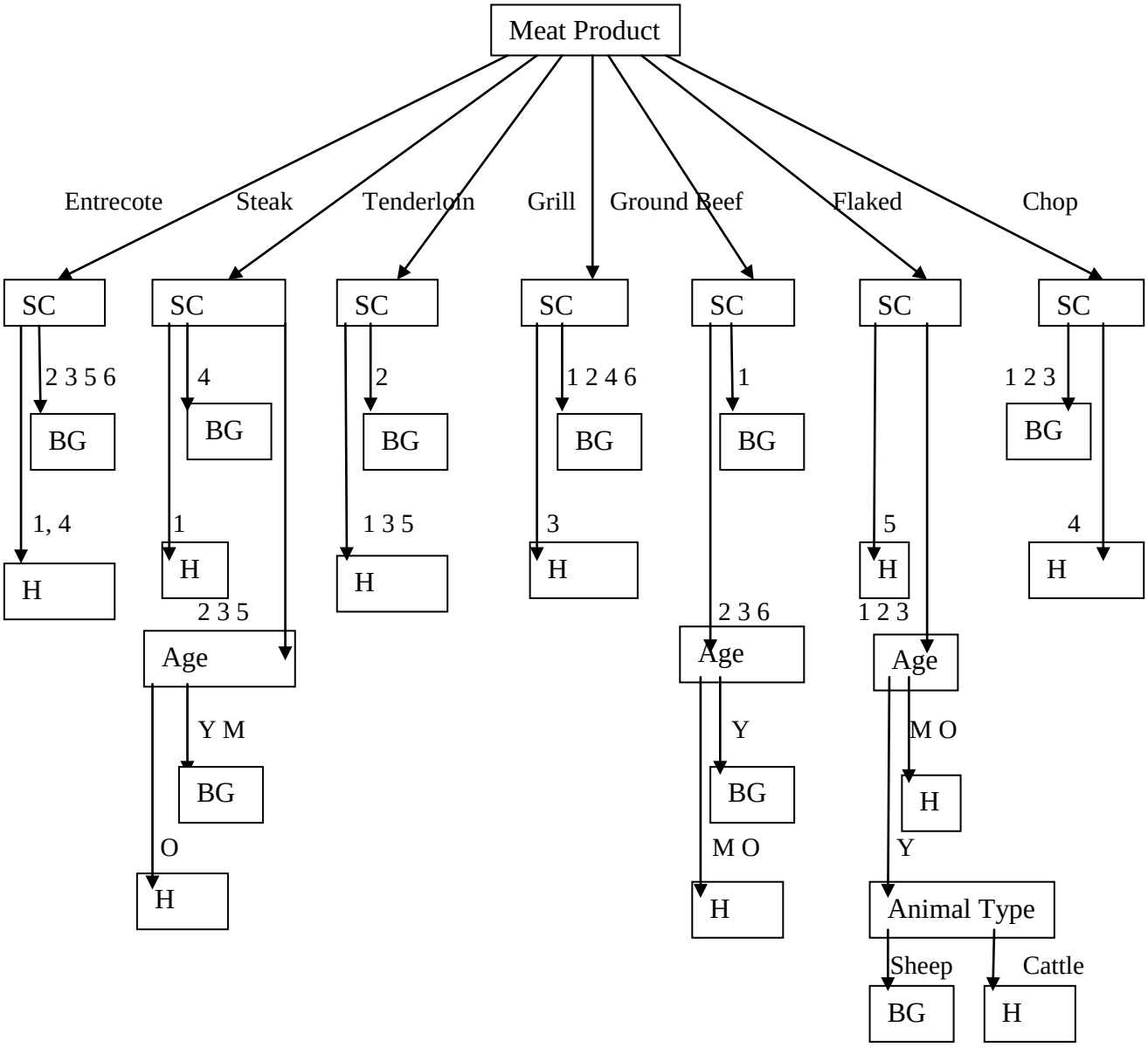
Altas, D., Gulpinar, V., 2012, "Karar Agaclari ve Yapay Sinir Aglarinin Siniflandirma Performanslarinin Karsilastirilmasi: Avrupa Birligi Ornegi", Trakya Universitesi, Sosyal Bilimler Dergisi, Volume 14, No 1, June 2012, pp. 1-22. (In Turkish)

https://weka.waikato.ac.nz (Last access date: 17.11.2014)

http://en.wikipedia.org/wiki/ID3_algorithm (Last access date: May, 2015)

Appendix: The complete decision tree

Properties are age (young, middle aged, old), selection criteria (price, appearance, taste, advertisement, presentation, transportation, date), animal type (cattle, sheep), meat product (entrecote, steak, tenderloin, grill, ground beef, flaked, chop) and output sample presentation form (barbeque grill and cook at home). Barbeque grill and cook at home are as an output.



SC: Selection Criteria
1: Price, 2: Appearance, 3: Taste, 4: Advertisement, 5: Presentation, 6: Transportation, 7: Date.
Y: Young, M: Middle Age, O: Old
BG: Barbeque Grill, H: Cook at home