Predicting how much a consumer is willing to pay for a bottle of wine: a preliminary study

Hugo Alonso[a,b,*], Teresa Candeias[a]

[a]Universidade Lusófona do Porto, Rua Augusto Rosa, n.º 24, 4000-098 Porto, Portugal
[b]Universidade de Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

Abstract

The wine industry is an important business sector, generating billions in annual revenue. In the last year, there were several lockdowns due to the COVID-19 pandemic and wine consumption at home has increased. This paper considers the problem of predicting how much a consumer is willing to pay for a bottle of wine to drink at home, in a regular occasion. As far as we know, this is the first study on the subject. The problem is treated as a classification task and several prediction models, based on artificial neural networks, support vector machines and decisions trees, are proposed and compared.

Keywords: Wine; Classification; Artificial neural networks; Support vector machines; Decision trees.

1. Introduction

The COVID-19 crisis is an event that has already shown large economic and social impacts worldwide. From an economic perspective, it has been an exogenous shock to local and international markets. According to the European Commission [1], the volume of wine consumption in the European Union (EU) would decrease 8% in 2020 comparing to the previous five years' average. Indeed, international trade and domestic sales have been affected by social distancing measures and self-isolation that led to the closure of restaurants, bars and clubs. In addition, the reduction in international travel and tourism activities affected wine consumption. Wine sales globally are being damaged by COVID19, and markets may take years to recover. This negative effect is greatest for premium sparkling wines, but all wine quality segments are affected [2]. Additionally, the pandemic situation has disrupted distribution channels [3]. The lockdown has changed consumption situations. Wittwer and Anderson [2] suggest that there was a huge reduction of opportunities to have a drink with friends and colleagues and a significant increase in self-consumption in Portugal. Thus, it is important understanding the consumer decision-making process. The consumption situation is pointed by [4] as one of the constructs that can affect the purchase intent and purchase decision, so personal preference and

* Corresponding author. Tel.: +351 222 073 230.
E-mail address: hugo.alonso@ulp.pt

1877-0509

purchase intent are not, themselves, fully signals of purchasing behavior. According to [5], situation is as a set of particular factors occurring at a time and place that are not connected with the personality, desires and capacities of the individual, or with the attributes of the product or service and have a systematic influence on the individual's behavior. Several studies suggest that wine purchase and consumption are significantly affected by the purchasing and consumption situation [6-9]. They also suggest that, since behavior depends on the consumption situation, they behave differently when placed in different contexts and at different times. As wine is a multi-attribute product that can only be evaluated during consumption, then the ability of consumers to assess prior quality to purchase is strongly asymmetric and consumers will rely on extrinsic cues to measure wine quality [10-11]. Thus, price has an important role in quality perception when there are few cues available, when the product cannot be evaluated or when the perceived risk of making a wrong choice is high [12-13]. Wine value for money enables the reasons and magnitude of the decision to purchase to be established, by measuring the gap between different prices ranges (minimum/ maximum) depending on levels of perceived quality that consumers associate with it, setting up a relevant signal of potential demand [14]. Likewise, Hall and Lockshin [6] found evidence that there is association between price and the occasion situation. In the same context, Orth [15] studied situational aspects related to the consumption of wine in the United States. The author found evidence that brand's choice and the benefits required in a wine differ in three situations: self-consumption, hosting friends or as a gift. Stöckl [16] argues that price is significant when consumer buys wine for own consumption but has little influence when wine is to be offered. Concern about social benefits is more often mentioned by younger consumers, who are more likely to have distress about judgment. In this regard, Stöckl [16] states that there are several influences that interfere with the purchasing decision process. However, their recognition, the weight of such influences can differ from high to null, depending on the consumption situation/occasion [16]. While a low price, for instance, may play an important role when consumers buy wine for consumption, it has little effect when wine is to be offered [16].

The problem here considered is to predict how much a consumer is willing to pay for a bottle of wine to drink at home, in a regular occasion. As far as we know, this is the first study on the subject. Given information about a consumer, such as his/her age and income, we are interested in predicting how much he/she is willing to spend in a bottle: less than EUR 2.99; between EUR 3 and 4.99; between EUR 5 and 9.99; EUR 10 or more. Since these intervals can be viewed as classes, the prediction problem can be treated as a classification task. The prediction models we propose are classifiers based on artificial neural networks, support vector machines and decisions trees [17]. In addition to the usual classification approach, where we develop classifiers without taking into account the existence of an order relation between the classes, we consider an ordinal classification approach, where we develop classifiers that tackle the classes' natural ordering. In the ordinal approach, we consider the so-called unimodal model [18] and a modification of Frank and Hall's method [19], proposed in [20].

The remainder of this paper is organized as follows. The next section describes the data used in this study. The classification methods and models are presented in Section 3. Finally, the results are shown in Section 4 and the conclusions and future work in Section 5.

2. Data

The data used in this work were collected during a two-month period (January and February 2020), using a questionnaire created with Google Forms and made available through social media channels. The population was limited to wine buyers of legal drinking age (+18) purchasing wine. To make sure that the respondents were a suitable target group for wine consumption, they were first asked if they are wine consumers and how often they drink wine. Table 1 describes the variables registered and their role in our prediction problem. Recall that we are interested in predicting how much a consumer is willing to pay for a bottle of wine to drink at home, in a regular occasion, *i.e.*, a bottle price, which is therefore the output or response variable, given information about his/her gender, age and all other input or predictor variables mentioned in the table. Considering that factors which influence consumer behavior are cultural, social, personal and psychological [21], one of the goals of this study is to study the effect of personal-demographic factors on price decisions. Spawton [22] argues that wine has been generally perceived as a feminine beverage. Richie [23] argues that even though wine buying is considered a traditional male role, more females buy wine. Other studies have studied the gender effect of wine behavior consumption [24-26]. According to Atkin, Nowak and Garcia [27], gender plays a significant role in the wine information search as well as in subsequent buying behavior. The authors suggest that males and females show different patterns and this information should be taken into account when approaching customers. Besides gender, age segmentation in wine business is becoming more important as wine becomes a lifestyle beverage for all generations [28]. Nevertheless, wine image changes among age [29] because wine consumption increases with age and experience [30-31].

Table 1. Variables in the questionnaire and their role in the prediction problem.

Variable	Measurement level	Values	Role
Gender	Nominal	Male and Female (2 categories)	Input (Predictor)
Age	Ordinal	18-24, 25-34, 35-44, 45-64 and 65+ years (5 categories)	Input (Predictor)
Marital status	Nominal	Married / Civil union, Single, Divorced and Widowed (4 categories)	Input (Predictor)
Education level	Ordinal	Basic, Secondary, Bachelor, Master and Doctorate (5 categories)	Input (Predictor)
Region of residence	Nominal	North, Center and South of Portugal (3 categories)	Input (Predictor)
Income	Ordinal	EUR -649.99, 650-999.99, 1000-1999.99, 2000-2999.99, 3000-3999.99, 4000+ (6 categories)	Input (Predictor)
Wine knowledge	Ordinal	None, Poor, Acceptable, Good and Very good (5 categories)	Input (Predictor)
Consumption frequency	Ordinal	Never, Rarely, Sometimes, Very often and Always (5 categories)	Input (Predictor)
Bottle price	Ordinal	EUR -2.99, 3-4.99, 5-9.99, 10+ (4 categories)	Output (Response)

We gathered valid data for a sample of $n = 228$ individuals. Table 2 presents a summary of it. Figure 1 shows that the frequency distribution of the bottle price class variable is unbalanced: most people are willing to pay less and few people are willing to pay more for a bottle of wine. Note that some people said that they never drink, but they buy wine for others at home.

Table 2. Summary of the valid data collected (n = 228).

Variable	Mode (Frequency)	Minimum	Maximum	Median	Interquartile range
Gender	Male (51.32%)	---	---	---	---
Age	25-34 years (28.95%)	18-24 years	65+ years	35-44 years	2 categories
Marital status	Married / Civil union (50.88%)	---	---	---	---
Education level	Secondary (35.53%)	Basic	Doctorate	Secondary	1 category
Region of residence	North (50.88%)	---	---	---	---
Income	EUR 650-999.99 (31.58%)	EUR -649.99	EUR 4000+	EUR 650-999.99	2 categories
Wine knowledge	Acceptable (39.47%)	None	Very good	Acceptable	1 category
Consumption frequency	Rarely (26.32%)	Never	Always	Sometimes	2 categories
Bottle price	EUR 3-4.99 (41.67%)	EUR -2.99	EUR 10+	EUR 3-4.99	1 category

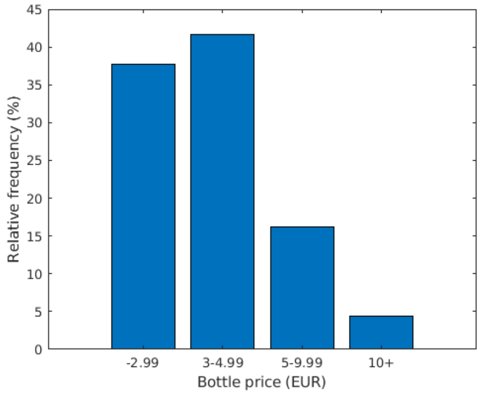

Fig. 1. Frequency distribution of the bottle price class variable.

As will be seen later on, two of the most important variables for bottle price prediction are the education level and the income of the consumer. Kendall's tau-b correlation between bottle price and education level is 0.49 and between bottle price and income is 0.22 (both are significant at the 1% level). Thus, the greater the education level or the income, the greater the consumer's willingness to pay for a bottle of wine tends to be.

In order to develop and assess the prediction models mentioned in the next section, we partitioned the data set into training and test subsets. The former was used to fit models and was assigned 2/3 of the cases, *i.e.*, 152, while the latter was used to test selected models and was assigned the remaining 1/3 of the cases, *i.e.*, 76. We did a stratified partitioning, keeping the percentage of cases in each bottle price class in the training and test sets the same as in the entire data set.

3. Models and methods

In this paper, three types of predictive models are considered: artificial neural networks, support vector machines and decisions trees [17]. Two advantages of decision trees are their interpretability and the ease with which they deal with qualitative predictive variables. Artificial neural networks and support vector machines are not as easily interpretable, but very often they have better generalization results. Details about these models are given in the previous reference. Applications in areas such as marketing, stock exchange and industrial engineering can be found in [32-34], for instance.

Next, we present a short description of the two ordinal supervised classification approaches considered in this work, namely the so-called unimodal model [18] and a modification of Frank and Hall's method [19], proposed in [20]. Information about the conventional approach to supervised classification, where the order relation between the classes is not taken into account, can be found in [17].

3.1 The unimodal model

The unimodal model is a machine learning paradigm intended for supervised classification problems where the classes are ordered. It was introduced in [18] and was recently considered, for instance, in [35-36]. The main idea behind this model is that the random variable class associated with a given query should follow a unimodal distribution, so that the order relation between the classes is respected. In this context, the output of a classifier where the *a posteriori* class probabilities are estimated is obliged to be unimodal, *i.e.*, to have only one local maximum. There are different ways to impose unimodality and in [18] the authors suggested two approaches. In the parametric approach, a unimodal discrete distribution, like the binomial and Poisson's, is assumed and its parameters are estimated by the classifier. In the non-parametric approach, no distribution is assumed and the classifier is trained so that its output becomes unimodal. In all practical experiments conducted by the authors, the parametric approach led to better results, in particular when the binomial distribution was considered. The superior performance achieved with this distribution was also justified in theoretical terms. For these reasons, our focus here is on the binomial model.

Furthermore, since the classifiers chosen by us are artificial neural networks, support vector machines and decisions trees, we refer hereafter to binomial networks, binomial support vector machines and binomial tress, respectively. For the sake of conciseness, next, we only present a detailed description of the binomial networks applied to our problem.

As mentioned before, given information about a consumer, we are interested in predicting how much he/she is willing to spend in a bottle: less than EUR 2.99; between EUR 3 and 4.99; between EUR 5 and 9.99; EUR 10 or more. Representing the information given about the consumer by $\boldsymbol{x}$ and the $K = 4$ bottle price classes, less than EUR 2.99, …, EUR 10 or more, by $C_1, \dots, C_K$, respectively, Bayes decision theory [17] suggests classifying the bottle price in the class maximizing the *a posteriori* probability $P(C_k|\boldsymbol{x})$. To that end, the *a posteriori* probabilities $P(C_1|\boldsymbol{x}), \dots, P(C_K|\boldsymbol{x})$ need to be estimated. In the binomial network, these probabilities are calculated from the binomial distribution $B(K-1, p)$. As this distribution takes values in the set $\{0,1,\dots,K-1\}$, we take value 0 to represent class C_1, 1 to C_2, and so on, until $K-1$ to C_K. Now, since K is known, the only unknown parameter is the probability of success p. Hence, we consider a network architecture as in Fig. 2 and train it to adjust all connection weights from layer 1 to layer 3. Note that the connections from layer 3 to layer 4 have a fixed weight equal to 1 and serve only to forward the value of p to the output layer of the network where the probabilities from the binomial distribution are calculated. For a given query $\boldsymbol{x}$, the output of layer 3 will be a single numerical value in $[0,1]$, denoted by $p_{\boldsymbol{x}}$. Then, the probabilities in layer 4 are calculated from the binomial distribution:

$$P(C_k|\boldsymbol{x}) = B_{k-1}(K-1, p_{\boldsymbol{x}}), \quad k = 1, \dots, K, \tag{1}$$

where

$$B_{k-1}(K-1, p_{\boldsymbol{x}}) = \frac{(K-1)!\, p_{\boldsymbol{x}}^{k-1}(1-p_{\boldsymbol{x}})^{K-k}}{(k-1)!\,(K-k)!}. \tag{2}$$

When $p_{\boldsymbol{x}}$ is in $\left[0, \frac{1}{K}\right[$, the highest *a posteriori* probability is $P(C_1|\boldsymbol{x})$, and, therefore, the predicted bottle price class is C_1. More generally, when $p_{\boldsymbol{x}}$ is in $\left[\frac{i-1}{K}, \frac{i}{K}\right[$, for some i in $\{1, \dots, K-1\}$, the highest *a posteriori* probability is $P(C_i|\boldsymbol{x})$, and, therefore, the predicted bottle price class is C_i. Hence, in order to train the network on a training set $T = \{(\boldsymbol{x}_n, C_{\boldsymbol{x}_n})\}_{n=1}^{N} \subset X \times \{C_k\}_{k=1}^{K}$, where X is the feature space, we replace C_k by the value of p corresponding to the midpoint of $\left[\frac{k-1}{K}, \frac{k}{K}\right[$, *i.e.*, $p_k = \frac{k-0.5}{K}$, and apply a suitable optimization algorithm, like the Marquardt method [37], to find connection weights that minimize the mean squared error

$$\frac{1}{N}\sum_{n=1}^{N}\left(p_{\boldsymbol{x}_n}^{target} - p_{\boldsymbol{x}_n}^{network}(\boldsymbol{w})\right), \tag{3}$$

where $p_{\boldsymbol{x}_n}^{target}$ is the value of p replacing $C_{\boldsymbol{x}_n}$ and $p_{\boldsymbol{x}_n}^{network}(\boldsymbol{w})$ is the output of layer 3 given the query $\boldsymbol{x}_n$ and having the network the weights $\boldsymbol{w}$.

3.2 Modified Frank and Hall's method

Frank and Hall's method was originally introduced in [19]. Just like the unimodal model approach previously presented, the method is intended for supervised classification problems where the classes are ordered. As before, suppose that the $K = 4$ bottle price classes are represented by $C_1, \dots, C_K$, respectively. Frank and Hall propose to use $K-1$ binary classifiers to address the K-class ordinal problem. In order to train the classifiers, such as artificial neural networks, support vector machines or decisions trees, $K-1$ datasets are derived from the original dataset. The i-th classifier is trained to discriminate $C_1, \dots, C_i$ from $C_{i+1}, \dots, C_K$. Given an unseen instance $\boldsymbol{x}$, *i.e.*, information about a new consumer, the *a posteriori* probabilities $P(C_1|\boldsymbol{x}), \dots, P(C_K|\boldsymbol{x})$ of the original K classes can be estimated by combining the outputs of the $K-1$ binary classifiers for that instance. As noticed in [20], the combination scheme suggested by Frank and Hall may lead to negative probabilities, but the problem can be overcome in the following manner: identifying the output p_i of the i-th classifier with the conditional probability $P(C_{\boldsymbol{x}} > C_i | C_{\boldsymbol{x}} > C_{i-1})$, the classes can be ranked according to the following formulas:

$$P(C_{\boldsymbol{x}} > C_1) = p_{\mathbf{1}} \qquad\qquad P(C_1|\boldsymbol{x}) = 1 - p_{\mathbf{1}} \tag{4}$$

$$P(C_{\boldsymbol{x}} > C_j) = p_j P(C_{\boldsymbol{x}} > C_{j-1}), \qquad P(C_j|\boldsymbol{x}) = (1 - p_j) P(C_{\boldsymbol{x}} > C_{j-1}), \quad j = 2, \ldots, K-1,$$

$$P(C_K|\boldsymbol{x}) = P(C_{\boldsymbol{x}} > C_{K-1}).$$

This is known as the modified Frank and Hall's method.

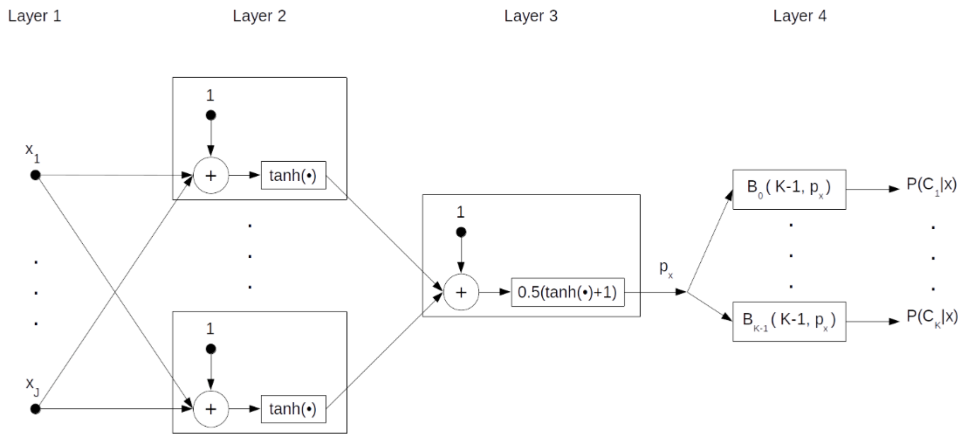

Fig. 2. Binomial network.

4. Results

All computer experiments were carried out using Matlab R2021a. We fitted artificial neural networks (NNs), support vector machines (SVMs) and decisions trees to the training data. The models' hyperparameters were chosen in order to obtain the best estimate of the prediction error, calculated by applying stratified 5-fold cross-validation to the training set [17]. In this way, we avoided underfitting and overfitting. This was done in the conventional approach to supervised classification, in the unimodal paradigm and in the modified Frank and Hall's method. The trained models were then applied to the test data. Their performance was measured by the classification accuracy or rate of correct classifications and also by a coefficient called r_{int}, which measures the association between the ordinal variables true class and predicted class [18], [38]. This association coefficient takes values in $[-1, 1]$: 1 when the two variables are identical and -1 when they are completely opposite. Remark that the classification accuracy alone is not completely adequate to measure the models' performance, because it is not true that every misclassification is equally costly. For instance, assume for a certain consumer that the true bottle price class is C_1, *i.e.*, less than EUR 2.99. Then, it is worse to have C_3 for predicted class, *i.e.*, between EUR 5 and 9.99, than C_2, *i.e.*, between EUR 3 and 4.99, since in the first case the predicted class is farther from the true class. The results of the trained models in the test set are shown in Table 3. It can be seen that the best performances were always achieved by SVMs and NNs, with a slight advantage of the former. The SVM in the conventional approach exhibited the highest classification accuracy, 64%. In turn, the binomial SVM and the binomial NN exhibited the highest r_{int}, 0,66. Thus, the SVM in the conventional approach made more correct classifications, but it was worse than the binomial SVM and the binomial NN when it misclassified. Considering all results, the binomial SVM is probably the best prediction model, with a good degree of agreement between true and predicted classes. We found that all classifiers have more difficulty in correctly predicting cases from higher classes, *i.e.*, in identifying consumers who are willing to pay more for a bottle of wine. This is somehow related to the fact that there are few of these cases in our dataset. Finally, we applied a sensitivity analysis, proposed in [39] and recently used, for instance, in [40-41], to measure the importance of the predictor variables in the various models and we found that the most important are the consumer's education level, consumption frequency, income and wine knowledge.

Table 3. Classification performance in the test set.

Performance measure	Conventional approach			Binomial model			Modified Frank and Hall's approach		
	Tree	SVM	NN	Tree	SVM	NN	Tree	SVM	NN
r_{int}	0,57	0,62	0,61	0,53	0,66	0,66	0,49	0,64	0,60
Classification accuracy	56%	64%	58%	55%	62%	61%	49%	62%	62%

5. Conclusions and future work

In this paper, we considered the problem of predicting how much a consumer is willing to pay for a bottle of wine to drink at home, in a regular occasion. Several prediction models, based on artificial neural networks, support vector machines and decisions trees, where proposed and compared. As far as we know, this is the first study on the subject and our good preliminary results encourage us to continue our research. Since our dataset is unbalanced, because most people are willing to pay less and few people are willing to pay more for a bottle of wine, in the future we plan to apply methods specifically designed for unbalanced data in order to try to further improve our results.

Acknowledgements

This work was partially supported by the Center for Research and Development in Mathematics and Applications (CIDMA) through the Portuguese Foundation for Science and Technology (FCT – Fundação para a Ciência e a Tecnologia), references UIDB/04106/2020 and UIDP/04106/2020.

References

[1] European Commission (2020). Short-term Outlook for EU Agricultural Products in 2020. Brussels: European Commission, DG Agriculture and Rural Development.

[2] Wittwer, G., and K. Anderson. (2021) "COVID-19 and global beverage markets: Implications for wine." *Journal of Wine Economics* **16 (2)**.

[3] Dubois, M., L. Agnoli,J. M. Cardebat, R. Compés, B. Faye, B. Frick, ... & K. Simon-Elorz. (2021) "Did wine consumption change during the COVID-19 lockdown in France, Italy, Spain, and Portugal?." *Journal of Wine Economics*, **16 (2)**: 131-168.

[4] Kotler, Philip. (2000) Marketing Management, NJ: Prentice-Hall.

[5] Belk (1974)

[6] Hall, J. and L. Lockshin. (1999) "Understanding Wine Purchasing. It's not the Consumer, It's the Occasion." *Australian & New Zealand Wine Industry* **14 (3)**: 69-78.

[7] Olsen, J. and L. Thach. (2001) ''Consumer behaviour and wine consumption: a conceptual framework.'' *Australian and New Zealand Wine Industry* **16 (6)**: 123-9.

[8] Olsen, J.E., K.J. Thompson and T.K. Clarke. (2003) ''Consumer self-confidence in wine purchases.'' *International Journal of Wine Marketing* **15 (3)**: 40-51.

[9] Ritchie, C. (2007) ''Beyond drinking: the roleof wine in the life of the UK consumer.'' *International Journal of Consumer Studies* **31 (5)**: 534-40.

[10] Salaun, Y. and K. Flores. (2001) "Information quality: meeting the needs of the consumer". *International Journal of Information Management* **21**: 21-37.

[11] Antonelli, G. (2004) (a cura di) *Marketing agroalimentare. Specificità e temi di analisi*. Milano: FrancoAngeli.

[12] Dodds, W.B. and K.B. Monroe. (1985) "The Effect of Branded & Price Information on Subjective Product Evaluations." *Advances in Consumer Research* **12**: 85-90.

[13] Spawton, T. (1991) "Marketing Planning for Wine". *European Journal of Marketing*, **25** (**3**): 6-48.

[14] Keown, C. and M. Casey. (1995) “Purchasing behaviour in the Northern Ireland Wine Market.” *British Food Journal* **97** (**1**): 17-20.

[15] Orth, Ulrich. (2005) "Consumer personality and other factors in situational brand choice variation." *Journal of Brand Management* **13** (**2**): 115-133.

[16] Stöckl, Albert. (2007) Der Flaschenverschluss zur Differenzierung im Marketing, in Der Winzer, 09/2007, Vienna.

[17] Hastie, T., Tibshirani, R., and Friedman, J. (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd edition), New York, USA: Springer-Verlag.

[18] Pinto da Costa, J. F., Alonso, H., and Cardoso, J. S. (2008) The unimodal model for the classification of ordinal data, *Neural Networks*, **21**: 78-91.

[19] Frank, E., and Hall, M. (2001) A simple approach to ordinal classification, in *Proceedings of the 12th European Conference on Machine Learning*, 1 (pp. 145-156).

[20] Cardoso, J. S., and Pinto da Costa, J. F. (2007) Learning to classify ordinal data: the data replication method, *Journal of Machine Learning Research*, **8**: 1393-1429.

[21] Kotler, P. and K. Keller. (2008) *Marketing Management*, 13th edition, Prentice Hall.

[22] Spawton, T. (2000) A taxonomy of wine consumer research in Australia 1990-2000, paper presented at the XXIVeme Congress of Wine and Vine, Paris.

[23] Ritchie, C. (2009) “The culture of wine buying in the UK off-trade.” *International Journal of Wine Business Research* **21** (**3**): 194 – 211.

[24] Schamberg, A. (2002) “Battle of the sexes plays out in wine.” *Milwaukee Journal Sentinel*, 18 August.

[25] Robinson, J. 2004. A woman’s place is in the cellar, San Francisco Chronicle, 8 July.

[26] Hoffman, C.A. (2004) “When consumers buy wine, what factors decide the final purchase?” *Australian and New Zealand Wine Industry Journal* **19** (**2**): 82-91.

[27] Atkin, T., L. Nowak and R. Garcia. (2007) “Women wine consumers: information search and retailing implications.” *International Journal of Wine Business Research* **19** (**4**): 327-339.

[28] Bruwer J., Saliba A. & Miller B. (2011) “Consumer behaviour and sensory preference differences: implications for wine product marketing.” *Journal of Consumer Marketing* **28** (**1**): 5–18.

[29] Olsen, J.E., L. Thach and L. Nowak. (2007) “Wine for my generation: exploring how US wine consumers are socialized to wine.” *Journal of Wine Research* **18** (**1**): 1-18

[30] Quester, P.G. and J. Smart. (1996) “Product involvement in consumer wine purchases: its demographic determinants and influence on choice attributes.” *International Journal of Wine Marketing* **8** (**3**): 37-56.

[31] Stanford, L., P. Bailey and L. Rowley. (2008) Market insights: Australia, Wine Industry Outlook Conference Supply and Demand Report. Adelaide: Australian Wine and Brandy Corporation.

[32] Pereira, P. J., Cortez, P., and Mendes, R. (2021) Multi-objective grammatical evolution of decision trees for mobile marketing user conversion prediction, *Expert Systems with Applications*, **168**: 1-12.

[33] Zhong, X., and Enke, D. (2019) Predicting the daily return direction of the stock market using hybrid machine learning algorithms, *Financial Innovation*, **5**: 1-20.

[34] Guven, I., and Simsir, F. (2020) Demand forecasting with color parameter in retail apparel industry using artificial neural networks (ANN) and support vector machines (SVM) methods, *Computers & Industrial Engineering*, **147**: 1-9.

[35] Vargas, V. M., Gutiérrez, P. A., and Hervás-Martínez, C. (2022) Unimodal regularisation based on beta distribution for deep ordinal regression, *Pattern Recognition*, **122**: 1-10.

[36] Krinitskly, M., Aleksandrova, M., Verezemskaya, P., Gulev, S., Sinitsyn, A., Kovaleva, N., and Gavrikov, A. (2021) On the generalization ability of data-driven models in the problem of total cloud cover retrieval, *Remote Sensing*, **13**: 1-28.

[37] Rao, S. S. (2019) *Engineering Optimization: Theory and Practice* (5th edition), New Jersey, USA: John Wiley & Sons, Inc.

[38] Pinto da Costa, J. F., Alonso, H., and Cardoso, J. S. (2014) Corrigendum to “The unimodal model for the classification of ordinal data” [Neural Netw. 21 (2008) 78-79], *Neural Networks*, **59**: 73-75.

[39] Kewley, R., Embrechts, M., and Breneman, C. (2000) Data strip mining for the virtual design of pharmaceuticals with neural networks, *IEEE Transactions on Neural Networks*, **11**: 668-679.

[40] Rodríguez-Hernández, C. F., Musso, M., Kyndt, E., and Cascallar, E. (2021) Artificial neural networks in academic performance prediction: systematic implementation and predictor evaluation, *Computers and Education: Artificial Intelligence*, **2**: 1-14.

[41] Yan, H., Fu, H., Li, Y., Lin, T.-H., Wang, G., Zheng, H., Jin, D., and Zhao, B. Y. (2021) On migratory behavior in video consumption, *IEEE Transactions on Network and Service Management*, **18**: 1775-1788.