Food Quality and Preference 51 (2016) 1–7

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Short Communication

Visual attention accompanying food decision process: An alternative approach to choose the best models

Attila Gere [a], Lukas Danner [b], Nino de Antoni [b], Sándor Kovács [d], Klaus Dürrschmid [c], László Sipos [a,*]

[a] Sensory Laboratory, Corvinus University of Budapest, Villányi út 29-43, H-1118 Budapest, Hungary
[b] Department of Food Science, University of Applied Sciences Wiener Neustadt, Campus Wieselburg, Zeiselgraben 4, A-3250 Wieselburg, Austria
[c] Department of Food Science and Technology, University of Natural Resources and Life Sciences, Muthgasse 18, A-1190 Vienna, Austria
[d] Department of Research Methods and Statistics, University of Debrecen, Böszörményi út 138, H-4032 Debrecen, Hungary

ARTICLE INFO

Article history:
Received 2 July 2015
Received in revised form 21 December 2015
Accepted 20 January 2016
Available online 21 January 2016

Keywords:
Gazing behavior and choice
Eye-tracking of food products
Comparison of prediction models
Model ranking
Sum of ranking differences

ABSTRACT

Visual attention plays an active role in food choice. During eye-tracking, several gazing behavior parameters are measured along with the consumer's choice. In this study, a Tobii T60 eye-tracker was used to record the gazing behavior of 59 participants during multi-alternative choice tasks (4AFC) in which pictures of six different food product groups (apples, salads, instant soups, sausages, soft drinks and beers) were presented. The aims of this study were (1) to investigate the relationship between gazing parameters and choice (2) to create prediction models based on gazing data and (3) identify the best model. The applied thirteen statistical models showed strong relationships between gazing behavior and choice and gave accurate predictions for choice. Sum of ranking differences method was used to rank the prediction models based on ten performance indicators. Iterative Dichotomiser 3 algorithm, Quinlan's C4.5 decision tree algorithm and *k*-Nearest Neighbor's algorithm showed the best performances in the cases of the separate product groups. After merging the data sets, Iterative Dichotomiser 3 algorithm showed clearly the best performance to describe the relationship between visual attention and food choice.

1. Introduction

In a purchasing situation, the first sensory contact with food is mostly through the eyes (Wadhera & Capaldi-Phillips, 2014). Expectations and associations are elicited by visual factors and therefore, analysis of visual attention has shown to be able to contribute to the research upon consumers' decision making process.

Eye-tracking is widely used to evaluate food packages (e.g. Rebollar, Lidón, Martín, and Puebla (2015)) and nutrition labels (e.g. Graham, Orquin, and Visschers (2012)) where participants are often asked to look at the packaging and to evaluate it.

Since food choice is an outstanding relevant parameter of nutritional behavior a promising research field of eye-tracking is the analysis of the relationships between gazing behavior and food choice.

A recent study investigating the influence of the first fixation on consumer choice (van der Laan, Hooge, de Ridder, Viergever, & Smeets, 2015) found that the location of the first fixation had no positive correlation with consumers' choice. Therefore, the authors stated that catching the first gaze of the consumer might be unnecessary. However, they used a binary choice situation and the effect of first fixation could be more pronounced in more complex situations. Modeling the connection between visual attention and decision gives a more detailed picture of the role of the different eye-tracking parameters in food choice.

Supervised pattern recognition techniques are widely used in food analysis to complete discrimination and prediction tasks (Berrueta, Alonso-Salces, & Héberger, 2007). These techniques can be grouped into logic-based (e.g. decision trees), perceptron-based (e.g. artificial neural networks (ANN)), statistical learning (e.g. Linear Discriminant Analysis (LDA)) or distance-based ones (e.g. *k*-Nearest Neighbor (*k*NN)) (Kotsiantis, 2007). The steps of decision trees are usually univariate since they use splits based on a single feature at each internal node, hence they cannot perform well with problems that require diagonal partitioning. Decision trees and other learning algorithms have been compared and the study showed that Quinlan's C4.5 algorithm (C4.5) has a very good combination of error rate and speed (Lim, Loh, & Shih, 2000). Perceptron-based methods depend on three main aspects, input and activation functions, network architecture and the weight of each input connection. Neural networks require much longer training time than decision tree algorithms but most striking disadvantage of ANNs is their lack of ability to reason about

* Corresponding author.
E-mail address: sipos.laszlo@etk.szie.hu (L. Sipos).

http://dx.doi.org/10.1016/j.foodqual.2016.01.009
0950-3293/

their output in a way that can be effectively communicated. On the other hand, they perform well when multicollinearity is present (Kotsiantis, 2007). LDA is a well-known statistical learning algorithm which works with continuous observations. One of the most straightforward distance-based learning algorithms is the *k-Nearest Neighbor* (kNN) one, which is based on the principle that close proximity means similar properties (Cover & Hart, 2006). Following the findings of van der Laan et al. (2015) one would expect that statistical learning methods would perform well, when modeling the connection between visual attention and food decision.

In our approach we focus on the practical application of choice prediction from visual attention data, which is with other words the measurement of visual attention accompanying the decision process. The aims of this study were therefore (1) to investigate, which gazing parameters have a strong relationship with choice and (2) to create and rank models based on their performance and choose the best performing one. By doing so, consumer's choice is predicted from gazing behavior data using statistical models. The models were ranked regarding their prediction performance to identify superior models.

2. Materials and methods

2.1. Eye-tracking experiment

A multi-alternative forced choice paradigm (4AFC) without time limit was used. Seven choice sets were presented consisting of pictures each of four product alternatives. The first choice set was used as a warm-up to familiarize the participants with the procedure; hence, it was not included in the data analysis. The remaining six choice sets represented different food product categories including apple, salad, instant soup, sausage, soft drink and beer. For each choice task, pictures of four alternatives of the corresponding product group, with comparable familiarity and liking ratings, based on a pilot study with 40 students (equal gender and age distribution as for the main study), were presented as *stimuli*. The participants had to choose the product that appealed most to them without time limit (Fig. 1).

A Tobii T60 eye-tracker (60 Hz) and Tobii Studio software (version 3.0.5, Tobii Technology AB, Sweden) were used to present the stimuli and to analyze the gazing behavior of the 59 volunteered participants (29 male and 30 female aged between 18 and 28) during the choice task. Following six eye-tracking parameters were measured: (1) time to first fixation: time elapsed between the appearance of a picture and the user first fixating his/her gaze within an area of interest. (2) First fixation duration: time a user gazes at his/her first fixation point. (3) Fixation duration: length of a fixation (in seconds). (4) Fixation count: number of fixations on a product. (5) Dwell duration: time elapsed between the user's first fixation on a product and the next fixation outside the product (in seconds). The total dwell duration (sum of all dwell durations on an alternative) was used during a choice task for statistical analyses. (6) Dwell count: number of "visits" to an area of interest (AOI). The experiment took place under controlled environment (illumination, temperature etc.) in the sensory laboratory of the Department of Food Science and Technology at the University of Natural Resources and Life Sciences in Vienna.

The study was performed in accordance with the ethical guidelines for scientific research of the University of Natural Resources and Life Sciences and University of Applied Sciences Wiener Neustadt. Before the test, all participants were informed about the procedure and that their gazing behavior would be recorded during the task. All participants gave written informed consent concerning the use of the eye-tracking and questionnaire data for further analysis. Additionally, they were informed that they could withdraw themselves and their data from the study without giving an explanation at any time. All participants agreed to these conditions and participated without receiving a reward for their participation.

2.2. Variable selection

In the first step of the data analysis, Relief-F and Fisher filtering feature selection (approach applied to define a subset of relevant variables for use in model construction) was applied to identify the proper variables describing the relationship between the dependent (consumers' choice as categorical variable) and independent (eye-tracking data as continuous and frequency variables) variables. During Relief-F feature selection, the data set contains cases of attributes belonging to several classes (chosen product alternatives). The iterative algorithm is repeated many times assigning weights to the attributes ranging from 0 to 1. The more important an attribute is in the classification, the larger the weight of that attribute becomes. Relief-F starts with zero weights and a random case is selected with all its attributes from a given class. Then the algorithm determines the *k*-Nearest Neighbors per class by using Euclidean distance at each iteration and attribute weights are updated according to the distance from the nearest neighbors from all classes (Kononenko, Šimec, & Robnik-Šikonja, 1997). The weight of a given attribute is decreased by the squared differences from the attributes in nearby cases of the same class and is increased by the squared differences from the attributes in nearby cases of the other classes. The major advantage of the method is that it is highly noise-tolerant and robust to interactions.

Fisher's feature selection algorithm calculates the ratio of "between class variance" to the "within class variance" similarly to F-statistic used in analysis of variance. The score for the *i*th attribute S_i is calculated based on Eq. (1) (Tang, Alelyani, & Liu, 2015).

$$S_i = \frac{\sum_{k=1}^{K} n_j \cdot (\mu_{ij} - \mu_i)^2}{\sum_{k=1}^{K} n_j \cdot \rho_{ij}^2} \tag{1}$$

where μ_{ij} and ρ_{ij} are the mean and the variance of the *i*th attribute in the *j*th class, n_j is the number of cases in the *j*th class, μ_i is the mean of the *i*th attribute. After computing the scores for all attributes, the algorithm ranks the attributes and the best *m* can be selected. As Relief-F is more robust to attribute interactions than Fisher score, we calculate the top 10 attributes regarding both algorithms and then select the common attributes to form consensus

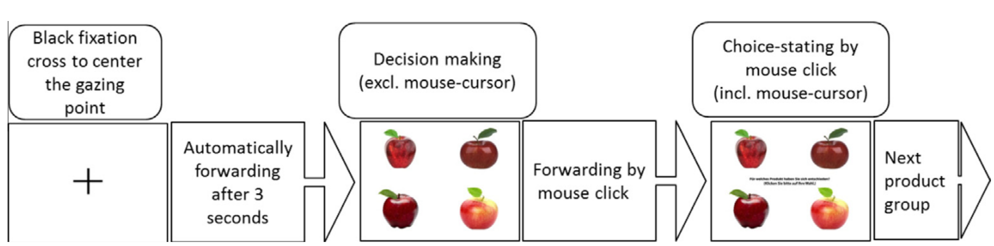

Fig. 1. Experimental layout of the apple choice set.

Table 1
Summary of the models.

	Computing time	Inputs	Major advantages	Major disadvantages	Accuracy in general	Explanation ability, transparency of classification	Type
KNN	Large	Continuous	Simplicity, easy implementation	Sensitive to noise, higher dimensions	High	Good	Instance-based
ID3	Medium	Continuous/discrete	Recursive procedure	Unable to handle missing and noisy data	Medium	Excellent	Logic-based
CSMC4	Medium	Continuous/discrete	Handles data with skewed distribution	Misclassification costs are often known	Medium	Excellent	Logic-based
C4.5	Large	Continuous/discrete	Handles high dimensions, missing and noisy data	Large tree	Medium	Excellent	Logic-based
CSCRT	Medium	Continuous/discrete	Handles data with skewed distribution	Misclassification costs are often known	Low	Excellent	Logic-based
RND	Medium	Continuous/discrete	Provides lower error rate	The rule set can be to large	High	Excellent	Logic-based
PLS-DA	Low	Continuous/binary	Robust to missing data, deals with multicollinearity	A single observation can be classified into more category	High	Good	Statistical learning
LDA	Low	Continuous	Provides maximal separability	Assumes Gaussian distribution, may overfit the data	High	Good	Statistical learning
MLP	High	Continuous	Fault tolerant	Difficult to interpret, no guaranteed solution	High	Average	Perceptron-based
NBC	Medium	Continuous/discrete	Deals well with large data, tolerates noisy data	Doesn't have min. error rate	Medium	Excellent	Statistical learning
RBF	High	Continuous	Provides incremental learning	Problem of selecting the appropriate number of basis functions	Low	Average	Perceptron-based
PNN	Large	Continuous	Easy to implement	Require an exhausting search to obtain optimal solution	Medium	Good	Instance-based
MLR	Medium	Continuous	Provides probability outcome, test interactions	Numerical problems (0 cell counts, collinearity of variables)	High	Good	Statistical learning

Abbreviations: KNN – *k*-Nearest Neighbor's algorithm, ID3 – Iterative Dichotomiser 3 algorithm, CSMC4 – Cost-sensitive Decision Tree algorithm, C4.5 – Quinlan's C4.5 decision tree algorithm, CSCTR – Cost-sensitive Classification Tree, RND – Random Trees, PLS-DA – Partial Least Squares Discriminant Analysis, LDA – Linear Discriminant Analysis, MLP – Multilayer Perceptron Neural Network, NBC – Naïve Bayes with Continuous variables, RBF – Radial Basis Function Neural Network, PNN – Prototype Nearest Neighbor, MLR – Multinomial Logistic Regression.

models. Both methods were run on the following variables: time to first fixation, first fixation duration, fixation duration, fixation count, dwell duration and dwell count.

2.3. Prediction models

In order to create balanced data sets (balanced choice frequencies) for the prediction models, bootstrapping was applied on each product within a product group. This resulted in a matrix with 4000 rows, 1000 rows corresponding to each chosen product alternative, in one product group. In the next step, thirteen models (*k*-Nearest Neighbor's algorithm (KNN), Iterative Dichotomiser 3 algorithm (ID3), Cost-sensitive Decision Tree algorithm (CSMC4), Quinlan's C4.5 decision tree algorithm (C4.5), Cost-sensitive Classification Tree (CSCTR), Random Trees (RND), Partial Least Squares Discriminant Analysis (PLS-DA), Linear Discriminant Analysis (LDA), Multilayer Perceptron Neural Network (MLP), Naïve Bayes with Continuous variables (NBC), Radial Basis Function Neural Network (RBF), Prototype Nearest Neighbor (PNN) and Multinomial Logistic Regression (MLR)) were used to predict the consumers' choice from the bootstrapped eye-tracking data. The classification models should (1) handle categorical outcomes, (2) be available to everybody and (3) have the same indicators. An introduction of the applied models is shown by Table 1. For a detailed discussion of the models, see Bhavsar and Ganatra (2012), Kotsiantis, Zaharakis, and Pintelas (2006) and Kotsiantis (2007).

Values of error rate, cross-validation results (minimum, maximum and average), prediction accuracy of each product in the group (upper right, upper left, bottom right and bottom left), bootstrapped error rate and results of leave-one-out cross-validation were computed to compare the performance of the models and to choose the superior one (the one having the best values of the parameters). The models task was to predict the choice based on gazing parameters for each product group as accurately as possible. All computations of the models were done using Tanagra version 1.4.50 (Rakotomalala, 2005).

2.4. The sum of ranking differences (SRD) procedure

SRD is a quick, simple and general technique suitable to compare methods or statistical models as well as to rank them based on their similarities and/or differences (Héberger, 2010). It is easy to use and the final result is a unique ranking (and grouping) validated by correct statistical tests. The SRD method has been applied in several fields (e.g., for column selection in chromatography (Héberger, 2010), for sensory panel testing (Kollár-Hunek & Héberger, 2013; Sipos et al., 2011)). Recently, the method was combined with analysis of variance (Héberger et al., 2014) and was used to evaluate proficiency tests along with principal component and cluster analysis (Škrbić, Héberger, & Durišić-Mladenović, 2013).

First, a theoretical best model (as a reference or benchmark) is defined. In our case, the theoretical best model has minimum error rates, cross-validation minimum, maximum and average values, bootstrapped error rate and leave-one-out validation values. Furthermore, its prediction accuracy for the four product alternatives is one (1). Then, rank numbers are ordered to the objects (rows or elements) of this theoretical model. Similarly, rank numbers are ordered to each model. This enables to calculate the absolute val-

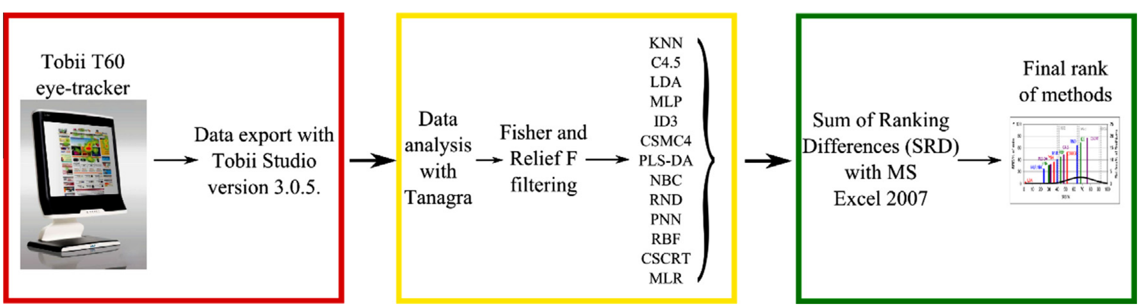

Fig. 2. Workflow of the data analysis. The three steps of data collecting and data analysis are grouped in boxes.

Table 2
Accuracy of the applied models and total dwell duration are expressed in percentages.

	Apple	Beer	Soft drink	Salad	Instant soup	Sausage	All products
Total dwell duration	*69.49*	*61.01*	*52.54*	*59.32*	*59.32*	*69.49*	*61.86*
KNN	99.75	99.75	99.50	98.25	99.75	99.75	99.75
ID3	97.25	97.00	98.75	97.25	97.00	97.50	99.63
CSMC4	95.00	96.50	91.00	78.50	91.25	93.50	84.58
C4.5	99.75	99.75	99.50	98.50	99.75	99.75	99.75
CSCRT	98.75	96.00	95.75	96.00	87.50	96.50	79.67
RND	99.75	99.75	98.00	95.25	99.75	99.75	99.58
PLS-DA	61.25	65.25	50.25	60.50	56.25	59.00	51.25
LDA	60.75	62.50	50.75	58.50	55.75	58.75	52.42
MLP	64.00	76.00	62.50	65.25	66.25	73.25	55.46
NBC	58.75	69.25	51.25	60.75	56.75	66.50	50.79
RBF	63.75	70.75	51.50	58.50	56.00	67.50	50.83
PNN	61.00	61.25	51.50	57.25	56.50	57.00	50.96
MLR	62.50	65.50	50.50	60.25	56.00	60.00	52.04

Abbreviations: KNN – *k*-Nearest Neighbor's algorithm, ID3 – Iterative Dichotomiser 3 algorithm, C4.5 – Quinlan's C4.5 decision tree algorithm, RND – Random Trees, CSCTR – Cost-sensitive Classification Tree, CSMC4 – Cost-sensitive Decision Tree algorithm, MLP – Multilayer Perceptron Neural Network, LDA – Linear Discriminant Analysis, MLR – Multinomial Logistic Regression, PLS-DA – Partial Least Squares Discriminant Analysis, PNN – Prototype Nearest neighbor, NBC – Naïve Bayes with Continuous variables and RBF – Radial Basis Function Neural Network.

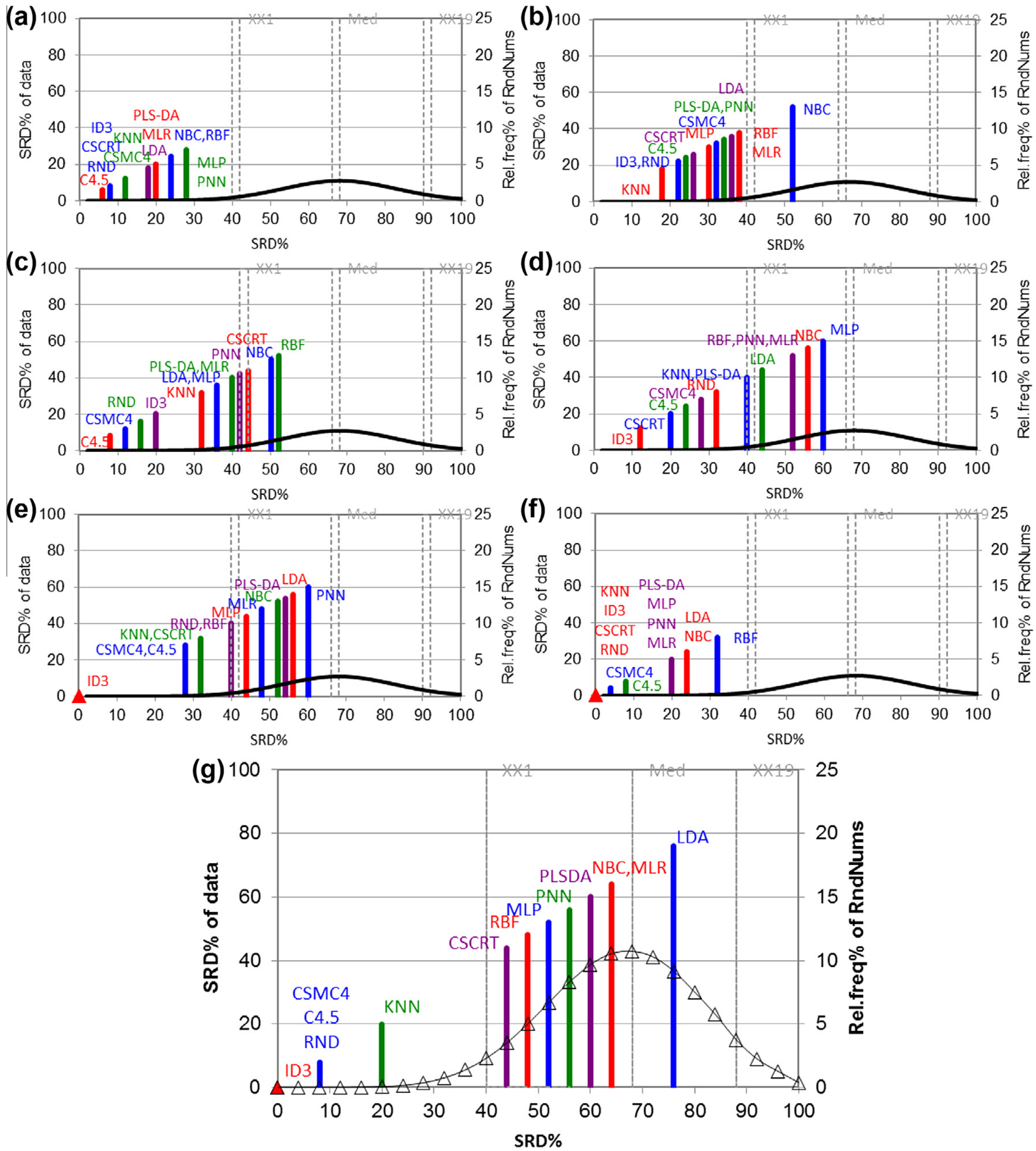

Fig. 3. The scaled SRD values of the models based on the performance indices determined by sum of ranking differences. The best possible values of the indices (Read) were used as reference (benchmark) column. Scaled SRD values are plotted on *x*-axis and left *y*-axis, right *y*-axis shows the relative frequencies (black curve). Probability levels 5% (XX1), Median (Med), and 95% (XX19) are also given. If a model crosses the Gauss-curve (XX1) say at $p = 0.10$ then, the method ranks the variable as random with a 10% chance. (a) – apple, (b) – salad, (c) – instant soup, (d) – sausage, (e) – soft drink, (f) – beer, (g) – all products.

ues of rank differences (SRDs) according to one of the models for each object. When the rank of the theoretical model and one other model is the same, the SRD value will be 0. Next, the SRDs are calculated for each of the models (thirteen times, hence thirteen models were included). With the comparison of the obtained SRD's, the methods can be easily compared. Models that deviate from the ideal one least are ranked better. Or, in other words, the lower the SRD of a method is, the better its performance is (i.e. a method with the smallest SRD is closer to the theoretical model than other models having larger SRDs). Generally, SRD requires standardization or rescaling due to different units and/or outliers, but in this study, all performance indicators were expressed in percentages; hence no standardization was required.

Sum of ranking differences method was calculated with Microsoft Office Excel 2007 macro (retrieved from: http://aki.ttk.mta.hu/srd).

The workflow of the applied three steps of data analysis is summarized in Fig. 2.

3. Results and discussion

After feature selection, the following variables were kept in the analysis: apple (total dwell duration, fixation count, fixation duration), salad (dwell count, fixation duration), instant soup (total dwell duration, dwell count), sausage (first fixation duration, dwell count), soft drink (dwell count, total dwell duration) and beer (first

Table 3
Results of sum of ranking differences method. The models are ranked according to their performances across all evaluated product groups.

Rank	Apple	Salad	Instant Soup	Sausage	Soft Drink	Beer	All products
1	**C4.5**	**KNN**	**C4.5**	**ID3**	**ID3**	**KNN**	ID3
2	**ID3**	**ID3**	CSMC4	CSCRT	CSMC4	**ID3**	CSMC4
3	CSCRT	RND	RND	**C4.5**	**C4.5**	CSCRT	C45
4	RND	**C4.5**	**ID3**	CSMC4	**KNN**	RND	RND
5	**KNN**	CSCRT	**KNN**	RND	CSCRT	CSMC4	KNN
6	CSMC4	MLP	LDA	**KNN**	RND	**C4.5**	CSCRT
7	LDA	CSMC4	MLP	PLS-DA	RBF	PLS-DA	RBF
8	PLS-DA	PLS-DA	PLS-DA	LDA	MLP	MLP	MLP
9	MLR	PNN	MLR	RBF	MLR	PNN	PNN
10	NBC	LDA	PNN	PNN	NBC	MLR	PLS-DA
11	RBF	RBF	CSCRT	MLR	PLS-DA	LDA	NBC
12	MLP	MLR	NBC	NBC	LDA	NBC	MLR
13	PNN	NBC	RBF	MLP	PNN	RBF	LDA

C4.5, ID3 and KNN are highlighted with underlined bold. Grey indicates non-significant models (which fall after XX1 (probability level 5%)). Models in the boxes have the same SRD values.

fixation duration, total dwell duration). These variables were found as most important and were included into the models of the different product groups.

The prediction rates of total dwell durations were computed by simply comparing the total dwell duration and chosen product alternatives and the obtained values were between 52.54% (soft drink) and 69.49% (apple and sausage) (Table 2). Comparing accuracy to the results of the total dwell duration revealed that the statistical models give more accurate predictions while using solely fixation/total dwell durations and counts. KNN, ID3, CSMC4, C4.5, CSCRT and RND produced high accuracy, which would indicate that their performance was similarly good. The other seven models showed weak performance. Their prediction accuracy was between 50% and 76% depending on the model-product combination. For the thirteen models, 10 performance indicators (see Section 2.3) were calculated and the best one should be chosen. To define the best models, sum of ranking differences (SRD) method was applied.

Fig. 3 shows characteristic groupings. A more detailed picture can be seen because the above mentioned six models (which have the highest accuracies) can be separated easily. Based on the SRD results of apple and instant soup, C4.5 was the closest to the zero point (SRD = 0), which indicates the theoretically best model (see Table 3). All other models had higher SRD values, which mean that C4.5 gave the best prediction results. Similarly good results were found for ID3 when analyzing sausages and soft drinks. KNN proved to be the best when predicting the choice from salad data. For the beer products, KNN, ID3, CSCRT and RND all had equally low SRD values, which indicates that there was no difference in their performance.

Further evaluation of Table 3 revealed that C4.5, ID3 and KNN had always low ranks but a superior modeling algorithm cannot be defined. Furthermore, RND, CSCRT and CSMC4 were significant in all six product groups. In the cases of apples and beers, all the models gave significant ranks because their SRD values were lower than the 5% probability levels (marked by XX1). The highest number of insignificant models was found in the cases of sausages and soft drinks, which indicates that the performance indicators of the models were inconsistent.

In the next step of the data analysis the different stimuli types were disregarded and all the six product groups were analyzed together. The last column of Table 2 shows the accuracy of the models run on all products. The accuracy of total dwell duration is 61.8% which was outperformed by KNN, ID3, CSMC4, C4.5, CSCRT and RND, the other models had weaker performance. In this data set, the two groups of models are more clearly separated.

After Fisher filtering and Relief-F feature selection, total dwell duration, fixation count and fixation duration were the three obtained variables. All the chosen variables were significant in the models. The SRD analysis of the performance indicators revealed that the lowest SRD values were found for ID3. It means that ID3 has the lowest error rates, cross-validation minimum, maximum and average values, bootstrapped error rate and leave-one-out validation values. Furthermore, its prediction accuracy for the four product alternatives is the highest among the tested models. On the second rank, three models were found, CSMC4, C4.5 and RND. The third rank was assigned to the last significant model, which was KNN. All the other models fell after XX1 (5% probability level), which means that their ranking cannot be distinguished from random ranking (Fig. 3g).

4. Conclusions

A close relation was found between gazing behavior and choice by the applied models, which supports the conclusion of the review by Orquin and Mueller Loose (2013).

Fisher filtering and Relief-F feature selection identified predictor variables, which were all significant in the prediction models. The results showed that it is valuable to include other performance indicators than accuracy. The 10 performance indicators and the 13 applied models were successfully compared and ranked using sum of ranking differences method.

C4.5, ID3, KNN, CSCRT and RND were among the first five best models in most of the cases when analyzing the products independently. After merging the products' data sets together, the following order was found: ID3 as first, CSMC4, C4.5 and RND as second and KNN as third best one. These findings support the conclusions of (Atalay, Bodur, & Rasolofoarison, 2012; Chandon, Hutchinson, Bradlow, & Young, 2009; van der Laan et al., 2015), who suggest that choice-related gazing patterns are similar for many types of stimuli and instructions. Decision tree algorithms showed better performance, which could be due to their logic-based system. The predictor variables seem to have more logic than nonlinear, linear or instance-based connection with the chosen product.

Additionally, the non-significant models (which application is not recommended to the given problem) were also identified (Linear Discriminant Analysis, Multinomial Logistic Regression, Naïve Bayes with Continuous variables).

A close relationship between gazing behavior and food choice was observed; furthermore, the food choice can be accurately predicted by gazing parameters using decision tree algorithms, preferably with ID3 models. The workflow, proposed in this study, is well-suitable to similar practical eye-tracking problems.

Acknowledgments

The authors are grateful to those who participated in this study. This work was supported by the research project of the National Development Agency TÁMOP 4.2.1.B-09/1/KMR-2010-005 and Campus Hungary TÁMOP4.2.4B/1-B1/1R/15128.

References

Atalay, A. S., Bodur, H. O., & Rasolofoarison, D. (2012). Shining in the center: Central gaze cascade effect on product choice. *Journal of Consumer Research, 39*(4), 848–866. http://dx.doi.org/10.1086/665984.

Berrueta, L. A., Alonso-Salces, R. M., & Héberger, K. (2007). Supervised pattern recognition in food analysis. *Journal of Chromatography A, 1158*(1–2), 196–214. http://dx.doi.org/10.1016/j.chroma.2007.05.024.

Bhavsar, H., & Ganatra, A. (2012). A comparative study of training algorithms for supervised machine learning. *International Journal of Soft Computing and Engineering, 2*(4), 74–81.

Chandon, P., Hutchinson, J. W., Bradlow, E. T., & Young, S. H. (2009). Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *Journal of Marketing, 73*(6), 1–17. http://dx.doi.org/10.1509/jmkg.73.6.1.

Cover, T., & Hart, P. (2006). Nearest neighbor pattern classification. *EEE Transactions on Information Theory, 13*(1), 21–27. http://dx.doi.org/10.1109/TIT.1967.1053964.

Graham, D. J., Orquin, J. L., & Visschers, V. H. M. (2012). Eye tracking and nutrition label use: A review of the literature and recommendations for label enhancement. *Food Policy, 37*(4), 378–382. http://dx.doi.org/10.1016/j.foodpol.2012.03.004.

Héberger, K. (2010). Sum of ranking differences compares methods or models fairly. *TrAC Trends in Analytical Chemistry, 29*(1), 101–109. http://dx.doi.org/10.1016/j.trac.2009.09.009.

Héberger, K., Kolarević, S., Kračun-Kolarević, M., Sunjog, K., Gačić, Z., Kljajić, Z., et al. (2014). Evaluation of single-cell gel electrophoresis data: Combination of variance analysis with sum of ranking differences. *Mutation Research. Genetic Toxicology and Environmental Mutagenesis, 771*, 15–22. http://dx.doi.org/10.1016/j.mrgentox.2014.04.028.

Kollár-Hunek, K., & Héberger, K. (2013). Method and model comparison by sum of ranking differences in cases of repeated observations (ties). *Chemometrics and Intelligent Laboratory Systems, 127*, 139–146. http://dx.doi.org/10.1016/j.chemolab.2013.06.007.

Kononenko, I., Šimec, E., & Robnik-Šikonja, M. (1997). Overcoming the myopia of inductive learning algorithms with RELIEFF. *Applied Intelligence, 7*(1), 39–55. http://dx.doi.org/10.1023/A:1008280620621.

Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. *Informatica, 31*, 249–268.

Kotsiantis, S. B., Zaharakis, I. D., & Pintelas, P. E. (2006). Machine learning: A review of classification and combining techniques. *Artificial Intelligence Review, 26*(3), 159–190. http://dx.doi.org/10.1007/s10462-007-9052-3.

Lim, T. S., Loh, W. Y., & Shih, Y. S. (2000). Comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. *Machine Learning, 40*(3), 203–228. http://dx.doi.org/10.1023/A:1007608224229.

Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica, 144*(1), 190–206. http://dx.doi.org/10.1016/j.actpsy.2013.06.003.

Rakotomalala, R. (2005). TANAGRA: Un logiciel gratuit pour l ' enseignement et la recherche. In Proceedings of EGC'2005 (Vol. 2, pp. 697–702). Retrieved from <http://eric.univ-lyon2.fr/~ricco/tanagra/fichiers/le_logiciel_tanagra_egc_2005.pdf>.

Rebollar, R., Lidón, I., Martín, J., & Puebla, M. (2015). The identification of viewing patterns of chocolate snack packages using eye-tracking techniques. *Food Quality and Preference, 39*, 251–258. http://dx.doi.org/10.1016/j.foodqual.2014.08.002.

Sipos, L., Kovács, Z., Szöllősi, D., Kókai, Z., Dalmadi, I., & Fekete, A. (2011). Comparison of novel sensory panel performance evaluation techniques with e-nose analysis integration. *Journal of Chemometrics, 25*(5), 275–286. http://dx.doi.org/10.1002/cem.1391.

Škrbić, B., Héberger, K., & Durišić-Mladenović, N. (2013). Comparison of multianalyte proficiency test results by sum of ranking differences, principal component analysis, and hierarchical cluster analysis. *Analytical and Bioanalytical Chemistry, 405*(25), 8363–8375. http://dx.doi.org/10.1007/s00216-013-7206-5.

Tang, J., Alelyani, S., & Liu, H. (2015). Feature selection for classification: A review. In C. Aggarwal (Ed.), *Data classification: Algorithms and Applications* (pp. 44). CRC Press.

van der Laan, L. N., Hooge, I. T. C., de Ridder, D. T. D., Viergever, M. A., & Smeets, P. A. M. (2015). Do you like what you see? The role of first fixation and total fixation duration in consumer choice. *Food Quality and Preference, 39*, 46–55. http://dx.doi.org/10.1016/j.foodqual.2014.06.015.

Wadhera, D., & Capaldi-Phillips, E. D. (2014). A review of visual cues associated with food on food acceptance and consumption. *Eating behaviors* (Vol. 15). http://dx.doi.org/10.1016/j.eatbeh.2013.11.003.