Fuzzy Approach to Purchase Intent Modeling Based on User Tracking For E-commerce Recommenders

Fuzzy Approach to Purchase Intent Modeling Based on User Tracking For E-commerce Recommenders

Piotr Sulikowski *
Faculty of Comp. Science and IT
West Pomeranian University of Technology
Szczecin, Poland
psulikowski@wi.zut.edu.pl
*Corresponding author

Tomasz Zdziebko
Faculty of Economics, Finance and Management
University of Szczecin
Szczecin, Poland
tomasz.zdziebko@usz.edu.pl

Omar Hussain
School of Business
University of New South Wales, Canberra
Canberra, Australia
o.hussain@adfa.edu.au

Anna Wilbik
Department of Data Science and Knowledge Engineering
Maastricht University
Maastricht, The Netherlands
a.wilbik@maastrichtuniversity.nl

2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) | 978-1-6654-4407-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/FUZZ45933.2021.9494585

***Abstract*—Recommender systems play a vital role in e-commerce by presenting personalized product suggestions, reducing habituation and leading to transactions in an environment with limited human touch. Data used for learning how to select optimal recommendation content, including mouse tracking data, are often imprecise in nature. In this paper, we present a fuzzy approach to model purchase intent based on tracking user interaction with a browser via mouse and keyboard. It appreciates data uncertainty and provides insights into e-commerce customer behavior and the development of shops online. The developed fuzzy rule-based systems had a good accuracy and low interpretability, and results show that to generalize possible purchase intent with fuzzy rules, it is good to begin with looking at such behavioral features as distance of mouse movement, distance of vertical page scrolling, number of mouse clicks and time of user activity on website in relation to page content length. In future work, we intend to look at more features reflecting product parameters and transactions to enhance the modeling results on a larger scale.**

***Keywords*— recommender system, human-computer interaction, mouse tracking, e-commerce, artificial intelligence, fuzzy systems**

I. Introduction

The question that many companies face is how to build a successful e-store. Online shop content presentation plays an important role in the success of any e-commerce enterprise. Being able to emphasize the recommendations of the right products, in which a user may be interested, could improve customer satisfaction and sales. Bringing more intelligence to e-commerce sites by identifying product interest and rules governing purchase intent may help act proactively and provide a user with personalized product content. In this paper we aim to provide e-commerce owners with the first steps towards getting appropriate insights regarding product recommendations for customers using fuzzy approach.

Recommender systems are used in a number of areas such as feeds in social media platforms, content in video playlists, and products in e-commerce. They suggest certain items according to user behavior and preferences in order to help them discover relevant objects with the potential to meet their needs [1]. In online stores, recommenders may be considered as the solutions stepping in for salespeople and providing the personal touch through offering a more personalized online shopping experience. E-customers are often overwhelmed in making the right selection. Identifying products according to the customer purchase intent is important in providing relevant recommendations that may alleviate that overwhelm and reduce habituation. User demographic, behavior and transactional data allow to build user models and profiles which reflect preferences and needs, and an important data stream comes from interactions with websites observed with techniques such as events tracking or eye tracking [2-9].

Recommender systems are usually based on collaborative or content-based filtering [10], and a number of machine learning methods can be used in recommender frameworks, such as deep learning [11-13]. Interest in a certain product leading to possible purchase can be determined by asking the user explicitly or by observing the user's interaction implicitly and inference. Explicit questioning may disturb natural behavior as it exerts and additional burden on the user [14-16] so it is mainly useful for gathering learning data. Unobtrusive implicit observations are better suited for monitoring subjects [17,18].

In case of implicit user tracking based on mouse moves and keyboard stroke events, commonly used implicit feedback indicators concern data such as mouse-move distance, scroll distance and certain key inputs. Studies are not consistent whether such indicators are always good and positive benchmarks of user interest and purchase intent [19,20].

Customer purchase intention is often affected by the attitude toward the purchase of the product driven mainly by its relevance to personal interests and the goal of the purchase, while also being affected by the external subjective norms regarding the purchase such as online feedbacks of a reputation system, in which all consumers can share opinions. Therefore, apart from easy-to-gather user-website interaction data describing user behavior on the website, there is other data which could well enhance purchase intent modeling, yet a lot of that data is not readily available or measurable, such as social influence [21].

Data obtained from mouse tracking or gaze tracking are usually rough in nature. The areas where users interact are not very precisely reflected in mouse tracking, less so in eye

This work was supported by the National Science Centre, Poland, Grant No. 2017/27/B/HS4/01216.

978-1-6654-4407-1/21/$31.00 ©2021

tracking. Fixation times of multiple subjects can be seen as granules, rather than precise times in milliseconds, to allow more natural interpretation of the results. Therefore, a fuzzy inference approach [22,23] seems natural for analyzing and interpreting tracking data, for the purpose of recommender optimization. It seems a good question whether this approach would allow for acceptable modelling of purchase intent based mainly on user-website interaction data.

Several proposals and implementations are found in literature to model intent of purchase using fuzzy logic. Most of them refer to the shopping process of certain product groups and factors influencing purchase intention, e.g. a study on two-wheelers sold in traditional shops [24], which includes a survey of fuzzy logic application in customer behavior modeling, a study on detecting intentions to purchase in B2C websites [25], focusing on technology, shopping, and product characteristics [26], or a study on customer involvement for selected electronic goods [27].

We have not come across studies testing the performance of fuzzy logic in purchase intent modeling based only on human-website interaction data and website characteristics.

The major contributions of this paper are: fuzzy rule-based systems (FRBS) for modeling user purchase intention, which can be used to improve the performance of recommender systems in e-commerce stores. The new approach is demonstrated through the use of implicit mouse and keyboard tracking with a browser extension and fuzzy modeling to discover relations between user behavior and purchase intent. This has the potential to model user purchase intent, and make better use of recommendation engines in e-commerce.

The rest of the article is structured as follows: General methodology and related work are presented in Section II. The structure of the experiments and empirical results are provided in Section III. Results are presented in Section IV, and conclusions in Section V.

II. Methodology

The main objective of this study is to present a fuzzy approach to product interest modelling to improve the effectiveness of e-commerce recommendation engines based on user activity data observed thanks to monitoring their interactions with website interfaces.

A popular technique for implicit monitoring of user behavior on websites is based on document object model (DOM) events model, which allows registering events resulting from interacting with the website on client side. The technique is called interaction logs and uses back-end languages such as Java, .NET and front-end AJAX technology and needs to be built into the e-commerce website by their developers. Since in the study we examined major e-commerce sites in Poland, such solution was not feasible for practical reasons, e.g. getting the owners' agreements, changing their website, sharing the data. Instead, we decided to develop a browser extension (add-on) – e-commerce customer preference monitor (ECPM) that allows for implicit monitoring of users' interactions with any existing website. With this method, it is possible to observe user behavior unobtrusively without any extra attention from them [28].

ECPM was built to collect online behavior data [2]. It was implemented as an add-on for the Mozilla FireFox browser, with the core code detachable and available for external use. ECPM gathers data on physical attributes of visited product pages and user interactions. It covers many parameters connected with the recommender interfaces such as time spent inside recommendation zones (mouse inside the zones) or the number of characters within recommended products section. Other research shows mouse position is quite strongly correlated with the eye gaze [29]. Times of mouse cursor being certain areas of interest (AOIs) such as the recommended products section can be related to other measures such as time spent on product page in general and page length etc. In this way relative measures were calculated.

Our aim is to provide e-commerce designers with insights on what to look at while developing websites with personalized recommendations when only behavioral data are available. Since the relations between the website content and product interest are not instinctively crisp, and mouse tracking data are rough in nature (e.g. a user may keep their cursor nearby the area they are looking at), we decided to use fuzzy theory in order to model those relations mathematically.

Fuzzy methodologies have two crucial features: the handling of uncertainty by describing knowledge with fuzzy sets and rules, and a good interpretability due to simple linguistic rules. In the presented fuzzy approach to modeling purchase intent based on DOM events data we will follow fuzzy modeling steps described below.

In this paper we develop two fuzzy rule-based systems, one used to explain user behavior, based on the Mamdani fuzzy inference (called from now on the Mamdani model), and the other employing the Takagi-Sugeno inference to test the prediction capabilities (the Sugeno Model).

For the Mamdani model, first, an induction of fuzzy partitions from data is necessary. This step allows defining meaningful linguistic variables and memberships of fuzzy sets that will be used later in rules generation. To ensure high interpretability, opting for strong fuzzy partitions is recommended [30]. Then rules from fuzzy partitions data are induced. Several induction methods are possible for rule induction with partitions previously defined, e.g. fast prototyping algorithm (FPA), Wang and Mendel (WM), and fuzzy decision trees (FDT) [31]. The knowledge base in the form of rules could be further optimized in order to increase accuracy while keeping high interpretability previously achieved.

Finally, the obtained knowledge base is evaluated using parameters such as accuracy and interpretability. While accuracy measures the closeness between the model and real data, interpretability is the measure of understating a system by analyzing the complexity knowledge base (the simpler the system, the more interpretable). Accuracy and interpretability have a trade-off relation [32].

To perform fuzzy calculations, we used the GUAJE software (Generating Understandable and Accurate Fuzzy Rule-based Systems in a Java Environment) [33-35] in Expert Mode. GUAJE is a user-friendly tool useful to generate and refine fuzzy knowledge bases related to specific problems. It implements the fuzzy modeling methodology called HILK (Highly Interpretable Linguistic Knowledge).

For the Takagi-Sugeno fuzzy inference system we use a different approach. Namely, first we divide data into a training and testing sets using stratified 10-fold cross validation. Next, we use fuzzy c-means clustering to derive the initial rules. Next, the parameters of the rules were tuned with ANFIS [36]. The Sugeno model has been created using the Matlab environment.

III. Experimental results

A. Implicit event tracking with ECPM add-on

The ECPM tool was adjusted for five major Polish online stores: Merlin.pl, Agito.pl, Electro.pl, Empik.com and Morele.net. From the perspective of assortment at the time of the study Agito.pl and Merlin.pl were horizontal shops, Empik.com was an online bookstore, whereas Electro.pl and Morele.net focused on electronic goods.

All participants in the study were non-profit volunteers who were active web users from Poland. There were 85 subjects aged between 19 and 33, all holding at least a high school degree. A limitation of the study is that the group of subjects was not randomly selected and did not constitute a representative subset of the whole population.

The task given to participants was to browse for interesting products and rate them. After leaving a product page a rating form was displayed, where a user could express the level of their product interest and purchase intention.

The subjects rated 1396 products and all customer interactions with websites were monitored with ECPM. User activity was monitored at the most granular level. Every DOM-fired event connected with mouse and keyboard was registered together with detailed information about the source element and its position in the structure of a web page. Beside this HTML code of every visited product page was collected. These data were used to generate parameters related to user behavior, which were classified into four groups: describing product page attributes (six parameters), describing user interactions times (in ms, eight parameters), describing user behavior (18 parameters), and finally relative parameters describing user behavior (10 parameters). On finishing browsing a product, participants were asked to rate their interest and purchase intent and declare possible earlier familiarity with the product. All extracted features are shown in Table I.

TABLE I. EXTRACTED FEATURES.

Parameter	Description
interest	Product interest and purchase intent rating
familiarity	Declared earlier familiarity with the product
Parameters describing product page attributes	
document_length	Number of characters within all texts on the page
desc_length	Number of characters within product description
review_length	Number of characters within product reviews
recommend_length	Number of characters within other recommended products section
image_number	Number of product images
page_height	Page height in pixels
Parameters describing user interactions times (in ms)	
page_time	Time between page load and page unload
tab_activ_time	Time while tab containing particular page is active
user_activ_time	Time while user is actively interacting with page (generating keyboard or mouse events)
prod_desc_time	Time while mouse pointer is positioned over product description
prod_recommend_time	Time while mouse pointer is positioned over recommended products section
prod_review_time	Time while mouse pointer is positioned over product reviews
prod_image_time	Time while mouse pointer is positioned over product images
prod_other_time	Time while mouse pointer is positioned in other sections of the webpage
Parameters describing user behavior	
mouse_distance	Total distance (in pixels) of mouse pointer movement
horizontal_scroll	Total distance (in pixels) of horizontal scroll of page content
vertical_scroll	Total distance (in pixels) of vertical scroll of page content
mouse_clicks	Total number of mouse clicks regardless which mouse key is pressed
lb_mouse_clicks	Total number of left mouse button clicks
rb_mouse_clicks	Total number of right mouse button clicks
mb_mouse_clicks	Total number of middle mouse button clicks
copycut_action	Total number of copy/cut actions performed via keyboard shortcut
select_action	Total number of page content selection actions
select_text_size	Total number of selected characters
keydown_single	Total number of key single pressing events
keydown_repeatable	Total number of key repeatable pressing events
find_action	Total number of find actions performed via keyboard shortcut
print_action	Total number of print actions performed via keyboard shortcut
bookmark_action	Total number of bookmarking actions performed via keyboard shortcut
save_action	Total number of save actions performed via keyboard shortcut
resize_action	Total number of page resizing actions performed via keyboard shortcut
search_referral	Boolean value indicating whether search result page was source of visit on product page
Relative parameters describing user behavior	
rel_page_time	page_time / document_length
rel_user_activ_time	user_activ_time / document_length
rel_tab_activ_time	tab_activ_time / document_length
rel_prod_desc_time	prod_desc_time / desc_length
rel_prod_recommend_time	prod_recommend_time / recommendation_length
rel_prod_review_time	prod_review_time / review_length
rel_prod_image_time	prod_image_time / image_number
rel_mouse_distance	mouse_distance / page_height
rel_vertical_scroll	vertical_scroll / page_height
rel_horizontal_scroll	horizontal_scroll / page_width

About 50% of the participants rated up to 13 browsed products, while the upper quartile rated over 20 products. Higher purchase intent ratings (interest variable) were provided more often than lower ones. The ratings 1 (lack or very low intent), 2, 3, 4 and 5 (very high intent) collected 130, 180, 325, 346 and 415 votes, respectively. An initial finding was that the users gave higher intent ratings to products which they were familiar with before participation in the study.

All the parameters beside binary variable *familiarity* were expressed on a numeric scale and were ranging from fractional values to several thousand. Building a model for generalizing user interest based on behavioral features expressed in a numerical way can result in a precise model. At the same time, establishing a sharp division between behavioral features which in nature are vague cannot guarantee that the discovered model will be easily graspable and usable in practice. By converting numerical features into fuzzy, rough concept features expressing user-website interactions e.g. total distance (in pixels) of vertical scroll of page content can be expressed in a more meaningful, interpretable way (low, average, large, very large, etc.).

B. Results of generalizing purchase intent based on fuzzy modeling

In order to build a fuzzy rule-based model for classifying purchase intent based on the observed interaction data we performed the following steps.

In the first step, data have been cleaned. Data collected from users least engaged in the study users, and thus the least reliable, have been removed. Only events collected from the participants who rated at least twenty products have been used since it was felt that users who rated a higher number of products better simulated the real behavior of customers. The resulting set consisted of 768 rows. Interest mark variable (in short: mark) reflects the level of purchase intent transformed onto a fuzzy scale consisting of two sets: low (interest 1 or 2) and high (interest 3, 4 or 5). This number of sets has been selected due to the requirements of many recommendation algorithms, which take user rating as input commonly on a binary scale.

As the next stage of the cleaning process 154 rows (roughly 20 percent) containing significant outlier values for input features have been removed and, as a result, the final set consisted of 614 cases. As the method for detecting outliers, isolation forest has been used [37]. The algorithm of isolation forest separates observed cases by randomly selecting both a feature and a split value between the maximum and the minimum values for a given feature. The number of splits required to separate a sample is equivalent to the path length from the root node to the terminating node because recursive partitioning can be represented as tree structure. The length of this path, averaged over a forest of all random trees, is a measure of normality and serves as a decision function. By using random partitioning, received paths for anomalies are much shorter. Thus, when a forest of random trees collectively produce shorter path lengths for particular samples, they are highly likely to be anomalies. The amount of contamination of the data set, i.e. the proportion of outliers in the data set, is used when fitting to define the threshold on the scores of the samples.

In the next step of pre-processing, data feature selection has been performed in order to reduce the number of features to ten, being more adequate for fuzzy rules generation, which can be interpreted by humans. Feature selection was based on the popular C4.5 algorithm introduced by Quinlan [38]. From two features relating to the same measure, original and relative, we have included one with higher information gain.

As a result, ten most informative features have been selected for fuzzy modelling procedure ordered with decreasing feature importance (measured with information gain ratio): rel_user_activ_time, rel_vertical_scroll, prod_review_time, mouse_distance, review_length, tab_activ_time, mouse_clicks, page_height, rel_prod_image_time, prod_recommend_time.

C. Inducing fuzzy partitions for input features for the Mamdani model

In the first step every input variable has been fuzzified with automatic induction of partitions. A great variety of validity indices have been proposed for fuzzy clustering, yet there is no universal index able to recognize the optimal number of clusters for various data sets [39-41]. We have considered three well established methods for inducing partitions from data: regular partitions, hierarchical fuzzy partitioning (HFP) and k-means. For all variables numerical distance was used for inducing partitions with all methods.

The first method creates standardized uniform strong fuzzy partitions. HFP is an ascending procedure, where at each step, for each given variable, two fuzzy sets are merged. The HFP algorithm discovers the high concentrated data areas by the agglomerative hierarchical clustering method, analyzes and merges the data areas, and then uses the evaluation function to find the optimum clustering scheme [42]. K-means method for inducing partitions uses a well-known clustering method of the same name. The center of the cluster defines the position where membership function is equal to one, while the other centers accordingly limit the support of that fuzzy set.

Decision which method (HFP or k-means) to use for a given input variable and how many number of fuzzy sets (clusters) to generate is one of the most difficult tasks in fuzzy sets induction. We have executed clustering algorithms several times with different numbers of clusters, ranging from one to nine, and then selected optimal shapes and numbers of clusters based on best index value of three metrics: Partition Coefficient, Partition Entropy and Chen Index, described in short below.

The Partition Coefficient (PC) proposed by Dunn [43] measures how close the fuzzy solution is to the corresponding hard solution. The hard solution is constructed by classifying each object into the cluster which has the largest membership. According to the formula for Dunn's PC (1), where k is the number of clusters, u_{ik} is the membership degree of data in cluster, and n is the number of observations, the coefficient ranges from $1/k$ to 1. Its value is 1/k when all memberships are equal to $1/k$. Its value is 1 when for each object the value of one membership is unity and the rest are zero.

$$PC = \frac{\sum_{k=1}^{n} \sum_{i=1}^{c} u_{ik}^2}{n} \quad (1)$$

The Partition Entropy (PE) proposed by Bezdek [44] is a simply index based on fuzzy membership values of fuzzy partitions and its formula is defined as (2).

$$PE = -\frac{1}{n}\{\sum_{k=1}^{n}\sum_{i=1}^{c}[u_{ik} log_a(u_{ik})]\} \quad (2)$$

The Chen index (Ch) [45] is defined as follows (3) and is an effective fuzzy partition measure as cluster validity criterion associated with fuzzy c-means algorithm.

$$Ch = \frac{1}{n}\sum_{k=1}^{n} max_i u_{ik} - \frac{2}{c(c-1)}\sum_{i=1}^{c-1}\sum_{j=i+1}^{c}\frac{1}{n}\sum_{k=1}^{n} min(u_{ik}, u_{jk}) \quad (3)$$

The results of partition induction for every input variable obtained with GUAJE are presented in Table II, where PN is the partitions number. For the majority of variables k-means partitioning method was finally applied with three partitions. Histograms and fuzzy partitions of three most important variables are presented in Figures 1-3.

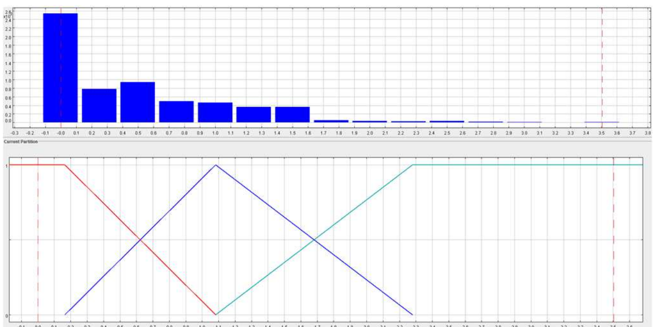

Fig. 1. Histogram and fuzzy subsets for input variable rel_verticall_scroll.

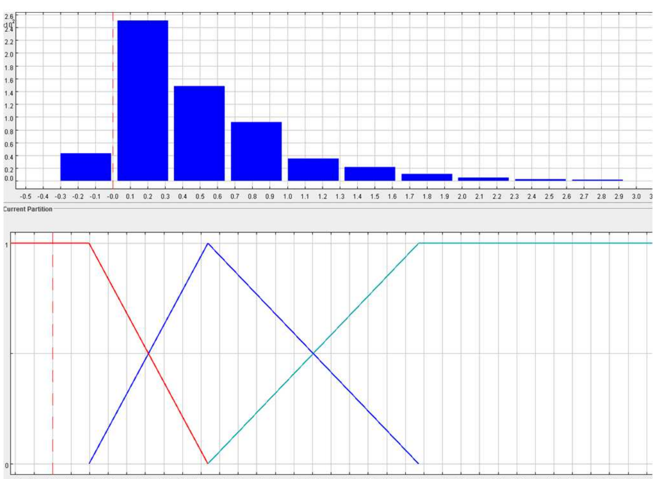

Fig. 2. Histogram and fuzzy subsets for input variable rel_user_activ_time.

TABLE II. VARIABLES FUZZY PARTITIONS WITH VALUES OF EVALUATION INDICES.

Variable	Partition induce method	PN	PC	PE	Ch
rel_user_activ_time	k-means	3	.799	.303	.802
rel_vertical_scroll	k-means	3	.815	.276	.818
prod_review_time	k-means	3	.981	.031	.982
mouse_distance	k-means	3	.807	.292	.810
review_length	k-means	3	.932	.113	.939
tab_activ_time	k-means	3	.801	.301	.806
mouse_clicks	k-means	5	.849	.234	.891
page_height	k-means	3	.835	.249	.754
rel_prod_image_time	regular	2	.965	.064	.956
prod_recommend_time	HFP	3	.950	.081	.954

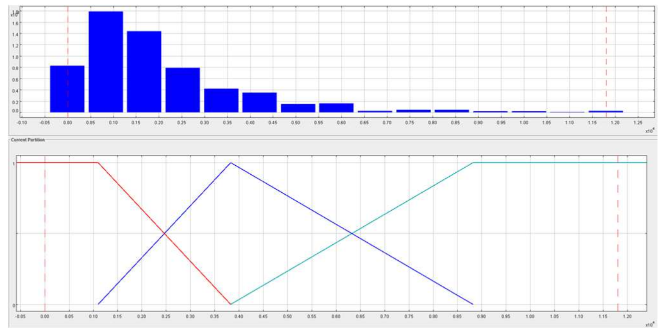

Fig. 3. Histogram and fuzzy subsets for input variable mouse_distance.

D. Fuzzy rules generation for the Mamdani model

In order to generate fuzzy rules, we used the fuzzy decision trees (FDT) algorithm. FDT sorts inputs according to their importance with the goal of minimizing entropy, the same as in C4.5 algorithm proposed by Quinlan [38]. In the next step, FDT translates a tree into a quite general incomplete rules set, due to considering only a subset of input variables.

For purchase intent rule generation, we have set the following parameters of the FDT algorithm: Maximum tree depth: 7; Minimum significant level which is the minimum matching degree for an item to be considered as belonging to the node: 0.2; Leaf minimum cardinality: 10. Tolerance threshold which represents the tolerance on the matching degree to the node majority class: 0.1; Minimum entropy/deviance gain: 0.001; Coverage threshold: 0.9; Prune: enabled; Split whole subtree/leaf: enabled.

As the intermediate result, the FDT algorithm produced 371 rules. After pruning, a fuzzy model with 120 rules was generated. To reduce the complexity of the resulting fuzzy model without decreasing its accuracy, we applied the simplification process to the fuzzy rules. Optimization was performed as a cyclical process containing two steps: rule base simplification and data base reduction. During the optimization, unused labels were removed and adjacent labels were used together. After the simplification process, we obtained a fuzzy model consisting of 42 rules with 100% coverage characterized by the following metrics: accuracy = 65.3%, error cases (EC) = 208, ambiguity cases (AC total) = 45, ambiguity cases error (AC error) = 13, mean square classification error = 0.084 and zero unclassified cases (UC).

Confusion matrix reveals that predicting low purchase intention has a higher precision but lower sensitivity than predicting high purchase intention (Table III). The F-measure shown is a harmonic mean of the precision and recall.

TABLE III. CONFUSION MATRIX WITH ACCURACY METRICS FOR THE MAMDANI MODEL.

Interest mark	TP	FP	TN	FN	Precision	Sensitivity	F-measure
low	111	30	290	183	0.787	0.378	0.51
high	290	183	111	30	0.613	0.906	0.731

Due to the limited range of the training data available, and a high number of resulting rules, some of them are not very promising, which shows that extra features, in particular other than coming from human-website interaction, could be useful for enhanced fuzzy modeling. Ten rules with the highest weights are presented in Table IV. Thanks to the granular approach, rules are easy explainable and understandable by humans. For example, rule 3 shows that even high review_length with low rel_user_activ_time and low tab_activ_time will result in a low level of purchase intent. If prod_review_time is high and rel_user_activ time is average and page_height is average, then we may assume a high level of purchase intent (rule 5).

TABLE IV. TEN HEAVIEST FUZZY RULES IN THE MAMDANI MODEL.

No.	Rule	Weight
1	IF mouse_clicks is very high AND rel_user_activ_time is average AND page_height is low THEN mark is low	0.35556
2	IF review_length is low AND rel_user_activ_time is low AND tab_activ_time is low THEN mark is low	0.29630
3	IF review_length is high AND rel_user_activ_time is low AND tab_activ_time is low THEN mark is low	0.29630
4	IF prod_review_time is high AND rel_user_activ_time is average AND page_height is high THEN mark is low	0.29630
5	IF prod_review_time is high AND rel_user_activ_time is average AND page_height is average THEN mark is high	0.29630
6	IF mouse_clicks is high AND mouse_distance is high AND rel_user_activ_time is low AND tab_activ_time is average THEN mark is low	0.23704
7	IF mouse_clicks is very low AND rel_user_activ_time is high AND tab_activ_time is high AND prod_recommend_time is low THEN mark is low	0.23704
8	IF mouse_clicks is high AND mouse_distance is low AND rel_user_activ_time is low AND tab_activ_time is average THEN mark is high	0.23704
9	IF review_length is average AND mouse_clicks is low OR average AND rel_user_activ_time is low AND tab_activ_time is low THEN mark is low	0.17778
10	IF mouse_clicks is average OR high AND rel_user_activ_time is high AND tab_activ_time is high AND prod_recommend_time is low THEN mark is low	0.17778

There are 42 rules and 206 premises in total, which accounts to a very high rule base structural dimension score equal one. All of the rules have three or more out of total number of ten inputs, which results in a very high rule base structural complexity measure equal one. There are ten inputs and 38 labels used in the rule base, which accounts to a very high rule base conceptual dimension score equal one. The percentage of elementary labels used in the rule base is 71%, which is rather a good measure. The percentage of OR composite labels used in the rule base is only 7.9%, while NOT operator is used only in 21% of rules. Rule base conceptual complexity index is only 0.179. Overall rule based structural and conceptual assessment scores are very high as they both equal one. The total number of labels defined in the knowledge base is only 32, which together with partition interpretability result in very high data base interpretability score equal one [46,47]. As a result of strict criteria, overall rules interpretability index is very low equal zero. But from the human perspective, the results of the most important rules are quite easily interpretable. Five most important rules consist only of three input variables from total ten inputs, then three rules consist of four inputs and only two rules consist of five input variables. The number of inputs in ten rules with the highest weight is still rather low and results in good comprehension.

The number of rule pairs which can be fired at the same time is equal to 580. The number of rule trios in the same situation is equal to 3945, whereas the number of rule sets (4 or more rules) which can be fired at the same time is equal to 722056, which overall constitutes a rather low local explanation index equal 0.25.

Theoretically, the interpretability index is very low, but compared to the complexity of rules and models generated with other machine learning algorithms (e.g. decision trees, neural networks), 42 rules with 38 labels constitute an acceptable base.

It needs to be emphasized that the generated fuzzy rule-based system can serve only as an evaluation of user behavior and purchase intent phenomenon. Due to the limited amount of data which well simulated the real behavior of customers, no data split for training and testing was performed. Therefore, the system cannot be used as a prediction tool extrapolated onto an unknown data realm. Therefore, we are aware that this is only a preliminary verification of the fuzzy approach to model purchase intent based on user tracking for e-commerce recommenders, in particular whether this approach may provide actionable, comprehensible insight.

E. *Results for the Sugeno model*

We have also generated the Sugeno model with two rules based on the same data. The average accuracy on the test data set based on the 10-fold stratified cross-validation is 0.58, while Kappa statistics is 0.15, which means that the obtained model is better than a random one. The average F1 score for interest mark low is 0.52, while for interest mark high is 0.62.

Below we present the best model (for one partition), with accuracy of 0.72 and Kappa of 0.44. The rules of the model are shown below, and we show also one membership value for variable mouse_distance (Figure 4). This membership is a result of fitting a Gaussian membership function to the cluster membership values, obtained by k-means clustering.

If review_length is $cl_{1,1}$ and mouse_clicks is $cl_{1,2}$ and prod_review_time is $cl_{1,3}$ and mouse_distance is $cl_{1,4}$ and rel_user_activ_time is $cl_{1,5}$ and rel_prod_image_time is $cl_{1,6}$ and rel_vertical_scroll is $cl_{1,7}$ and page_height is $cl_{1,8}$ and tab_activ_time is $cl_{1,9}$ and prod_recommend_time is $cl_{1,10}$ then interest mark is 0.08 x review_length + 0.18 x mouse_distance - 0.29 x rel_prod_image_time + 0.55 x prod_recommend_ time.

If review_length is $cl_{2,1}$ and mouse_clicks is $cl_{2,2}$ and prod_review_time is $cl_{2,3}$ and mouse_distance is $cl_{2,4}$ and rel_user_activ_time is $cl_{2,5}$ and rel_prod_image_time is $cl_{2,6}$ and rel_vertical_scroll is $cl_{2,7}$ and page_height is $cl_{2,8}$ and tab_activ_time is $cl_{2,9}$ and prod_recommend_time is $cl_{2,10}$ then interest mark is 0.03 x review_length + 0.11 x mouse_distance + 0.09 x rel_prod_image_time + 0.55 x prod_recommend_time.

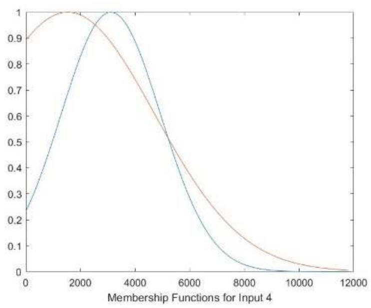

Fig. 4. Membership function for mouse_distance.

Naturally, those rules are more difficult to interpret than the rules of the Mamdani model, yet the predictive quality of this model is better. Figure 5 shows the surface visualization for the Sugeno model, where interesting saddle type behavior can be observed.

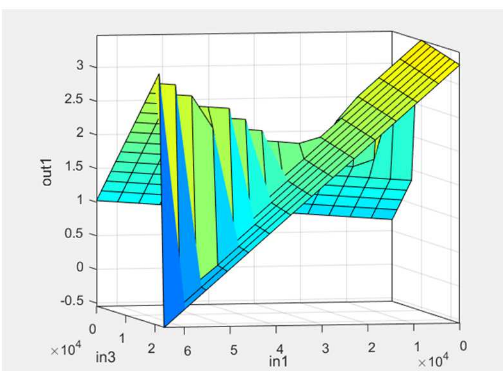

Fig. 5. Surface view for the Sugeno model. Review_length (in1) vs prod_review_time (in 3) .

IV. Conclusions

Recommender systems are commonly used to attract attention of customers towards products being in line with their preferences, and are supposed to result in higher conversion rates, purchase volumes and client retention. In this paper, a fuzzy rule-based system has been developed as an attempt for modeling purchase intent useful in developing personalized recommender systems in e-commerce stores.

The decision making of shopping online has a lot of subjectivity and requires a competent model to implement it. The application of computational intelligence is a good idea to develop an interpretable model. In this paper, a fuzzy model was proposed to implement purchase intent assessment of major e-commerce shops in Poland. Ten important parameters were used in this decision making.

We have created two models, the Mamdani model, with the purpose to explain the user behavior, and the Sugeno model to test the prediction capabilities. For the Mamdani model, most of the parameters' values translated into three fuzzy sets (low, average, high), which allows for easy understandability. As a final result, 42 rules were generated with 100% coverage and accuracy on a training test of 65.3%. The interpretability index is very low, yet compared to the complexity of rules and models generated with other machine learning algorithms (e.g. decision trees, neural networks) the result of 42 rules with 38 fuzzy labels is acceptable. It is worth noting that the most important rules use low number of input variables ranging from three for rules with the highest weights to five for two rules with the lowest weight. The low interpretability may result from too low scale of the study. Repeating the study at a larger scale for a longer period would give as better understanding of factors indicating purchase intent in e-commerce stores. However, due to the high number of explanatory variables involved in our work, it may be inevitable that many rules will still be fired simultaneously. For the Sugeno model, 10-fold stratified cross-validation accuracy is 0.58, which is satisfactory, considering the fact that it was based only on behavioral data gathered with ECPM. The average F1 score for interest mark high was better than for interest mark low (0.62 vs. 0.52).

In future research, we are planning to perform a larger-scale study of factors indicating product interest and purchase intent in cooperation with e-commerce stores. We are going to extend the number of factors which could explain purchase intentions in a wider way. We plan to include features reflecting not only behavioral data from mouse or eye tracking, but also sources which brought a customer to the shop (e.g. search engine advertisement, price comparer) and information describing the products themselves such as price, price comparison to other products, discount percentage/promotion, novelty (number of days in the offer or on the market), availability level etc. We plan to observe additional transaction-related actions such as adding product to a wish list. Those features should let us improve the understanding of key elements indicating intent toward product purchase and possibly obtain better modeling results.

References

[1] R. Burke, "Hybrid Recommender Systems: Survey and Experiments". User Modeling and User-Adapted Interaction 12 (4), 2002, pp. 331-370.

[2] P. Sulikowski, T. Zdziebko, D. Turzyński, E. Kańtoch, "Human-website interaction monitoring in recommender systems". Procedia Computer Science 2018, Vol. 126, pp. 1587-1596.

[3] P. Sulikowski, T. Zdziebko, D. Turzyński, "Modeling online user product interest for recommender systems and ergonomics studies". Concurrency and Computation Practice and Experience 31, Issue 22, 2019, e4301, pp. 1-9.

[4] P. Sulikowski, T. Zdziebko, "Horizontal vs. Vertical Recommendation Zones Evaluation Using Behavior Tracking". Applied Sciences 2021, Vol. 11(1), 56.

[5] J. Jankowski, P. Ziemba, J. Wątróbski, P. Kazienko, "Towards the Tradeoff Between Online Marketing Resources Exploitation and the User Experience with the Use of Eye Tracking". In: Nguyen N.T., 7. awiński B., Fujita H., Hong T.P.: Intelligent Information and Database Systems. 8th Asian Conference, ACIIDS 2016, Da Nang, Vietnam, March 14-16, 2016, Proceedings, Part I. Lecture Notes in Artificial Intelligence, vol. 9621. Berlin: Springer, 2016, pp. 330-343.

[6] A. Shahriar, M. Moon, H. Mahmud, K. Hasan, "Online Product Recommendation System by Using Eye Gaze Data". Proceedings of the International Conference on Computing Advancements, 2020.

[7] P. Sulikowski: “Evaluation of Varying Visual Intensity and Position of a Recommendation in a Recommending Interface Towards Reducing Habituation and Improving Sales”. In: Chao KM., Jiang L., Hussain O., Ma SP., Fei X. (eds.): Advances in E-Business Engineering for Ubiquitous Computing, Proceedings of the 16th International Conference on E-Business Engineering, ICEBE 2019, Shanghai, China, 11–13 October 2019, Lecture Notes on Data Engineering and Communications Technologies, 2020, Vol. 41, pp. 208–218.

[8] P. Sulikowski; T. Zdziebko; K. Coussement; K. Dyczkowski; K. Kluza; K. Sachpazidu-Wójcicka, „Gaze and Event Tracking for Evaluation of Recommendation-Driven Purchase”. Sensors 2021, 21, 1381.

[9] K. Bortko, P. Bartków, J. Jankowski, D. Kuras, P. Sulikowski, “Multi-criteria Evaluation of Recommending Interfaces towards Habituation Reduction and Limited Negative Impact on User Experience”. Procedia Computer Science 2019, Vol. 159, pp. 2240-2248.

[10] Li Zhang, Tao Qin, PiQiang Teng, “An Improved Collaborative Filtering Algorithm Based on User Interest”, Journal of Software, Vol. 9, No. 4, Apr. 2014.

[11] P. Sulikowski, T. Zdziebko, “Deep Learning-Enhanced Framework for Performance Evaluation of a Recommending Interface with Varied Recommendation Position and Intensity Based on Eye-Tracking Equipment Data Processing”. Electronics 2020, Vol. 9(2), 266.

[12] B.Ahirwadkar, S.N. Deshmukh, “Deep Neural Networks for Recommender Systems”. International Journal of Innovative Technology and Exploring Engineering, 8(12), 2019, pp. 4838–4842.

[13] S. Zhang, L. Yao, A. Sun and Y. Tay, “Deep Learning Based Recommender System”, ACM Computing Surveys, vol. 52, no. 1, 2019, pp. 1-38.

[14] D. Jannach, L. Lerche, M. Zanker, „Recommending based on implicit feedback. In Social Information Access”. In Brusilovsky P., He D. (eds), Social Information Access. Lecture Notes in Computer Science, vol 10100. Springer: Cham, Switzerland, 2018, pp. 510–569.

[15] L. Lerche, Using Implicit Feedback for Recommender Systems: Characteristics, Applications, and Challenges. Ph.D. Thesis, Technische Universität Dortmund, Germany, December 2016.

[16] Q. Zhao, F.M. Harper, G. Adomavicius, J.A. Konstan, “Explicit or implicit feedback? engagement or satisfaction?: A field experiment on machine-learning-based recommender systems”. In Proceedings of the 33rd Annual ACM Symposium on Applied Computing, Pau, France, 9–13 April 2018; ACM: New York, NY, USA, 2018, pp. 1331–1340.

[17] M. Zhou, Z. Ding, J. Tang, D. Yin., “Micro behaviors: A new perspective in e-commerce recommender systems”. In Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, Los Angeles, CA, USA, 5–9 February 2018, pp. 727–735.

[18] G. Tian, J. Wang, K. He, C. Sun, Y. Tian, “Integrating implicit feedbacks for time-aware web service recommendations”. Inf. Syst. G. Front. 2017, 19, 75–89.

[19] G. Velayathan, S. Yamada, “Behavior Based Web Page Evaluation”. Journal of Web Engineering, Vol. 1, No.1, 2005.

[20] G. Adomavicius, T. Alexander: “Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions”. IEEE Transactions on Knowledge and Data Engineering, Vol. 17, No. 6: 734–749, 2005.

[21] Y. Ku and Y. Tai, “What Happens When Recommendation System Meets Reputation System? The Impact of Recommendation Information on Purchase Intention”. 46th Hawaii International Conference on System Sciences, Wailea, Maui, HI, 2013, pp. 1376-1383.

[22] L.A. Zadeh, “Fuzzy sets”, Inform. and Control 1965, 8, 338-353.

[23] A. Piegat. Fuzzy Modeling and Control. Physica–Verlag: Heidelberg, Germany, 2001.

[24] S.Y. Kottala, “An empirical and fuzzy logic approach to product quality and purchase intention of customers in two wheelers”. Pacific Science Review B: Humanities and Social Sciences, Vol.1(1), 2015, pp. 57-69.

[25] M. Nilashi., O.B. Ibrahim, “A Model for Detecting Customer Level Intentions to Purchase in B2C Websites Using TOPSIS and Fuzzy Logic Rule-Based System”. Arab J Sci Eng 2014, 39, 1907–1922.

[26] L.C. Schaupp, F. Belanger, “A conjoint analysis of online consumer satisfaction”. J. Electron. Commer. Res. 2005, 6(2), 95–111.

[27] R. Oktavina, A.N. Nasution, “Consumer Involvement Analysis Using Fuzzy Mathematical Approach (Case Study: West Java, Indonesia)”. Proceedings of the 8th World Congress of the Academy for Global Business Advancement (AGBA) 2011. Advances in Global Business Research, 2011, Vol. 8, No. 1, pp. 145-157.

[28] S. Akuma, R. Iqbal, C. Jayne, F. Doctor, “Comparative analysis of relevance feedback methods based on two user studies”. Computers in Human Behavior, 60, 2016, pp.138-146.

[29] P. Boi, G. Fenu, L. Spano, V. Vargiu, „Reconstructing User’s Attention on the Web through Mouse Movements and Perception-Based Content Identification”. ACM Trans Appl Percept, 13(3), 2016, pp. 1-21.

[30] C. Mencar, M. Lucarelli, C. Castiello, and A.M. Fanelli, “Design of strong fuzzy partitions from cuts”. In Proceedings of the 8th Conference of the European Society for Fuzzy Logic and Technology, Advances in Intelligent Systems Research, 2013, pp. 424–431.

[31] S. Guillaume, B. Charnomordic, “Fuzzy inference systems: An integrated modeling environment for collaboration between expert knowledge and data using FisPro”. Expert Systems with Applications, 39(10):8744-8755, 2012.

[32] R. Alcala, J. Alcala-Fdez, J. Casillas, O. Cordon, F. Herrera, “Hybrid learning models to get the interpretability-accuracy trade-off in fuzzy modelling”. Soft Comput., 10 (2006), pp. 717-734.

[33] D.P. Pancho, J.M. Alonso, L. Magdalena, “Quest for interpretability-accuracy trade-off supported by Fingrams into the fuzzy modeling tool GUAJE”, International Journal of Computational Intelligence Systems, 6(sup1):46-60, 2013.

[34] J.M. Alonso, L. Magdalena, “Generating understandable and accurate fuzzy rule-based systems in a Java environment”, Lecture Notes in Artificial Intelligence, 9th International Workshop on Fuzzy Logic and Applications, LNAI 6857, 212-219, 2011.

[35] J.M. Alonso and L. Magdalena, “HILK++: an interpretability-guided fuzzy modeling methodology for learning readable and comprehensible fuzzy rule-based classifiers”, Soft Computing, 15(10):1959-1980, 2011.

[36] J.-S. R. Jang, “ANFIS: Adaptive-Network-based Fuzzy Inference Systems”, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 23, No. 3, May 1993, pp. 665-685.

[37] F.T. Liu, K.M. Ting, Z.-H. Zhou. “:Isolation-based anomaly detection”. ACM Transactions on Knowledge Discovery from Data, 6.1 (2012): 3.

[38] J.R. Quinlan, C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, 1993.

[39] W. Wang, Y. Zhang, “On fuzzy cluster validity indices”. Fuzzy Sets and Systems, vol.158, pp. 2095–2117, 2007.

[40] M. Halkidi, Y.Batistakis, M.Vazirgiannis, Clustering validity checking methods: part II, ACM Sigmod Record , vol.31(3), pp. 19-27, 2002.

[41] H. He, Y. Tan, K. Fujimoto, “Estimation of optimal cluster number for fuzzy clustering with combined fuzzy entropy index”, 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Vancouver, BC, 2016, pp. 697-703.

[42] l.-J. Li, Y.-L Liang, „A Hierarchical Fuzzy Clustering Algorithm”, 2010 International Conference on Computer Application and System Modeling (ICCASM 2010), Taiyuan, 2010, pp. V12-248-V12-25.

[43] J.C. Dunn, “Well-Separated Clusters and Optimal Fuzzy Partitions”. Journal of Cybernetics 1974, 4 (1): 95–104.

[44] J.C. Bezdek, “Cluster validity with fuzzy sets”, Journal of Cybernetics, vol.3, pp.58– 78, 1974.

[45] M.Y. Chen, “Establishing interpretable fuzzy models from numeric data”, In Proceedings of the 4th IEEE World Congress on Intelligent Control and Automation, 1857–1861, 2002.

[46] C. Mencar, C. Castiello, R. Cannone, A.M. Fanelli, “Interpretability assessment of fuzzy knowledge bases: a cointension based approach”, International Journal of Approximate Reasoning, 52(4):501-518, 2011.

[47] M. J. Gacto, R. Alcala, F. Herrera, “Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures”, Information Sciences, 181(20):4340-4360, 2011