Gender Disparities in Customer Churn Rates: A Rough Neuro-Fuzzy Classifier-based Analysis

Magdalena M. Scherer
Częstochowa University of Technology
Faculty of Management
Czestochowa, Poland *magdalena.scherer@pcz.pl*

Robert K. Nowicki
Czestochowa University of Technology
Faculty of Computer Science and Artificial Intelligence
Czestochowa, Poland
AGH University of Krakow
Faculty of Computer Science
Krakow, Poland *robert.nowicki@pcz.pl*

Abstract

This study addresses customer churn prediction using a rough neuro-fuzzy classifier with CA defuzzification, focusing on gender differences. Using rough set theory, it handles missing data without imputation through lower and upper approximations. Feature importance is assessed both directly and via the classifier to identify gender-specific churn patterns, introducing the concept of conditional significance. The results show notable gender-based differences, with female customers more sensitive to feature availability. The approach, validated by ten-fold cross-validation, offers insights for gender-tailored churn reduction strategies.

Keywords: rough sets, rough neuro-fuzzy classifier, customer churn prediction, gender.

1. Introduction

In today's competitive market, customer retention is a key challenge, as churn leads to financial and reputational losses [6]. Although gender-based differences in consumer behavior are well documented, their impact on churn remains underexplored. This study examines gender disparities in churn using the Bank Customer Churn dataset. We employ a rough neuro-fuzzy classifier [9], [10]with LEM-2 rule induction [4], enhanced by fuzzification and ten-fold cross-validation. A novel conditional significance coefficient is introduced to assess the gender-specific impact of attribute availability. Our aim is to improve churn prediction and support gender-aware retention strategies. Prior work highlights gender's role in segmentation and decision-making [1], [2], [7], while hybrid models combining preprocessing and classification have shown improved predictive performance [8]. The paper is organized as follows. Section 2 describes the dataset and the preprocessing approach based on rough set theory. Section 3 presents the classifier design, the rule generation method, and the experimental results. Section 4 concludes the study and discusses key gender-related findings.

2. Customer Churn Dataset Processing

We use the bank customer data set [5] with ten thousand cases, as it contains a mix of categorical and numerical features. In [11] the significance of attributes was analyzed without regard to the gender value. The dataset is treated as a decision table, according to Pawlak's rough set theory [10]. Thanks to this, we can calculate the significance coefficients of particular attributes or their groups [3]. The significance coefficient obtains values $[0; 1]$ and the value 0 occurs when the

Table 1. Normalized significance coefficient computed directly from the data for bank churn across ten attributes for two genders and overall for two genders. In the experiments, Geography is treated as one-hot encoded three variables but these coefficients stay the same after encoding.

Feature	Significance			Feature	Significance		
	female	male	any		female	male	any
CreditScore	0.0068	0.0066	0.0134	IsActiveMember	0.0018	0.0004	0.0022
EstimatedSalary	0.0042	0.0063	0.0105	NumOfProducts	0.0012	0.0008	0.0020
Tenure	0.0051	0.0044	0.0095	HasCrCard	0.0004	0.0004	0.0008
Age	0.0036	0.0042	0.0078	Geography	0.0000	0.0000	0.0000
Balance	0.0018	0.0036	0.0054	Gender	0.0000	0.0000	0.0000

lack of attribute values does not affect the classification. The value 1 means that without values of these attributes the classification will be impossible. In this work, we extend this concept to the conditional significance coefficient. Thus, in the gender context, we derive the significance separately under one of two conditions: gender is male or female. The conditional significance values with respect to genders for our dataset are given in Table 1. In Section 3 we compute significance coefficients from ten created classifiers by removing all possible features apart one each time. Table 1 shows that attributes such as Geography and Gender are insignificant. This means that partitioning the data into equivalence classes while omitting either of these two attributes (but not both simultaneously) will not cause two identical samples to be classified into different categories. Consequently, either attribute can be removed without affecting the classification quality. We can reasonably expect that constructing a decision system without either the Geography or Gender attribute would yield the same performance as using all 10 conditional attributes. The Geography attribute becomes redundant when all other attributes are available, and the same applies to the Gender attribute. The data presented in Table 1 clearly indicates the presence of redundancy in the dataset.

3. Results

We used a classifier based on the architecture from [11]. The rule base was generated using the LEM-2 algorithm [4] in two phases: data quantization and rule induction. Quantization transformed numerical attributes into intervals while preserving the determinism of the decision table, ensuring perfect classification of the training samples which is an advantage over gradient and evolutionary methods. Approximately 4,000 interval-based rules were created. Gaussian fuzzy sets were applied with widths set to 60% of the original interval size. To reduce noise, 90% of the most informative samples were used. We computed average normalized feature significance coefficients (ANIC) for various feature availability scenarios. Conditional versions of these coefficients were also introduced and are illustrated in Fig. 2. While Tabl. 1 shows significance values computed directly from the data, Fig. 2 presents importance within the operational decision system. The rough set-based classifier [9] produces one of three outputs: 'yes', 'no', or 'I don't know', depending on available features and gender. Average prediction outcomes (correct (OK), unknown (NO), and incorrect (ERR)) are shown in Fig. 2 for all combinations of missing features.

4. Conclusions

We used a Rough Neuro-Fuzzy Classifier [9] with the LEM-2 algorithm [4] to predict bank customer churn, focusing on gender differences. We introduced a conditional attribute significance coefficient to quantify how omitting attributes affects classification. The impact varies by gender: removing IsActiveMember reduces accuracy 4.5 times more for females, while Balance affects males twice as much. When three or more features are missing, the quality of the classi-

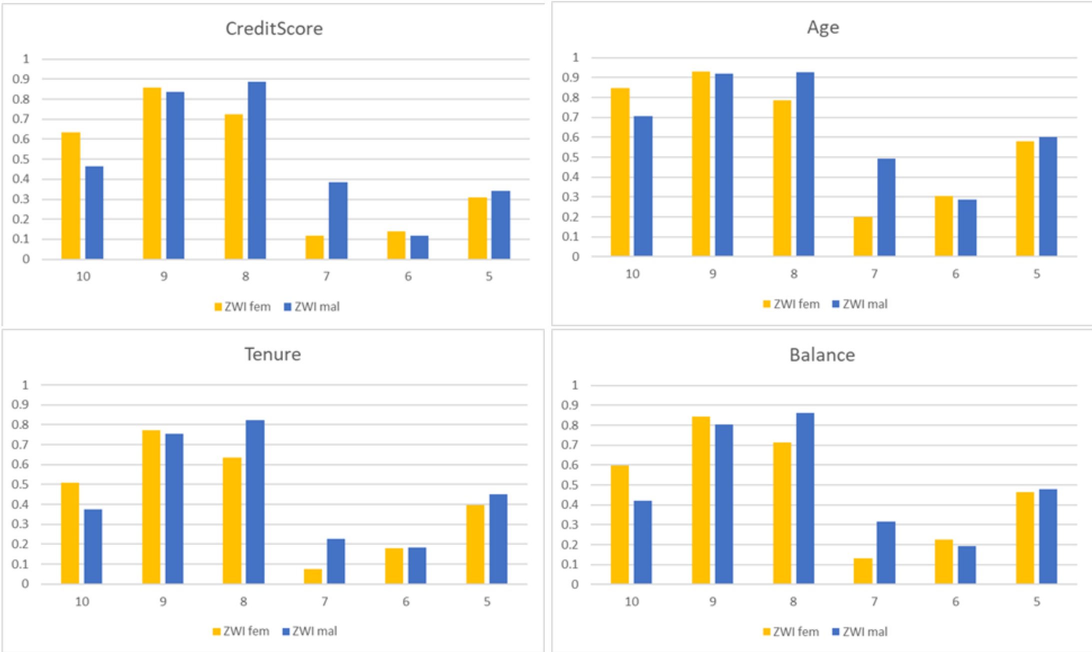

Fig. 1. Selected average normalized feature importance coefficient (ANIC) for the testing dataset. The first bar in each plot (yellow) represents the coefficients for females and the second one (blue) for males. Each plot shows average features importance when certain number of features is missing, apart from removing the feature denoted on a plot.

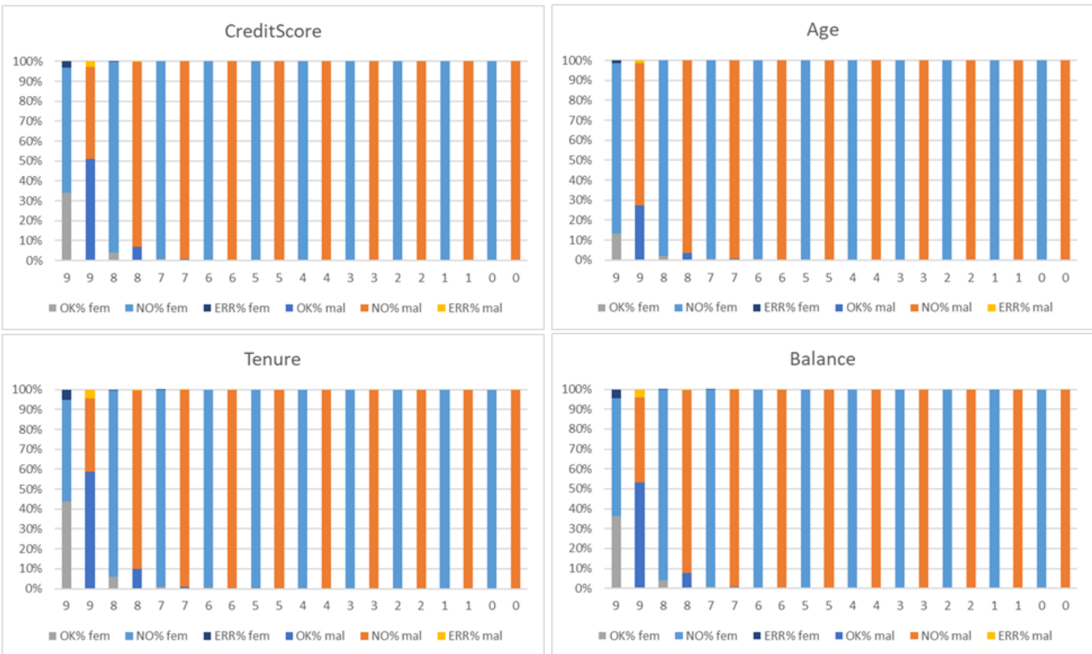

Fig. 2. Selected average accuracy in % (OK), do not know decision (NO), and wrong answers (ERR) for the testing dataset when the attribute value is unavailable. Each plot shows average predicted decisions for all combinations of removed features when certain number of features is missing, apart from removing the feature denoted on a plot.

fication decreases dramatically. Despite its low base importance, NumOfProducts becomes critical when other features are missing. Gender-based analysis reveals that missing CreditScore, Tenure, or IsActiveMember harms predictions for women more, while EstimatedSalary, Age, and Balance matter more for men. In general, missing data degrade the classification more severely for female customers.

Acknowledgment

R.N. gratefully acknowledges the funding support provided by the program "Excellence initiative—research university" for the AGH University in Krakow, as well as the ARTIQ project (UMO-2021/01/2/ST6/00004 and ARTIQ/0004/2021), and by the funds of the Polish Ministry of Science and Higher Education assigned to the AGH University of Krakow.

References

[1] Chen, Y., Yan, X., Fan, W., Gordon, M.: The joint moderating role of trust propensity and gender on consumers' online shopping behavior. Computers in Human Behavior 43, pp. 272–283 (2015)

[2] Chiu, Y.B., Lin, C.P., Tang, L.L.: Gender differs: assessing a model of online purchase intentions in e-tail service. International journal of service industry management 16(5), pp. 416–435 (2005)

[3] Cios, K.J., Pedrycz, W., Swiniarski, R.W.: Data mining methods for knowledge discovery, vol. 458. Springer Science & Business Media (2012)

[4] Grzymala-Busse, J.W.: Lers-a system for learning from examples based on rough sets. Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory pp. 3–18 (1992)

[5] Kaggle: Bank customer dataset. https://www.kaggle.com/datasets/adammaus/predicting-churn-for-bank-customers (2024), accessed April 2, 2024

[6] Khare, P., Arora, S.: Predicting customer churn in saas products using machine learning. International Research Journal of Engineering and Technology (IRJET) 11(5), pp. 754–765 (2024)

[7] Ladhari, R., Leclerc, A.: Building loyalty with online financial services customers: Is there a gender difference? Journal of Retailing and Consumer Services 20(6), pp. 560–569 (2013)

[8] Lenard, M.J., Madey, G.R., Alam, P.: The design and validation of a hybrid information system for the auditor's going concern decision. Journal of Management Information Systems 14(4), pp. 219–237 (1998)

[9] Nowicki, R.K.: Rough Set–Based Classification Systems. Springer International Publishing, Cham (2019)

[10] Pawlak, Z.: Rough sets. International Journal of Computer and Information Sciences 11(5), pp. 341–356 (1982)

[11] Scherer, M., Nowicki, R.: Customer churn prediction by rough neuro-fuzzy classifier with ca defuzzification. In: Marcinkowski, B., Przybylek, A., Jarzębowicz, A., Iivari, N., Insfran, E., Lang, M., Linger, H., Schneider, C. (eds.) Harnessing Opportunities: Reshaping ISD in the post-COVID-19 and Generative AI Era (ISD2024 Proceedings). University of Gdańsk, Gdańsk, Poland (2024)