NN-based predictive analytics for prepaid mobile services in telco sector

Predictive analytics; Telco sector; Prepaid mobile services; Customer behavior; Top-up propensity; Account balance; Neural Network; Machine Learning; Random Forest; Gradient Boosting Trees

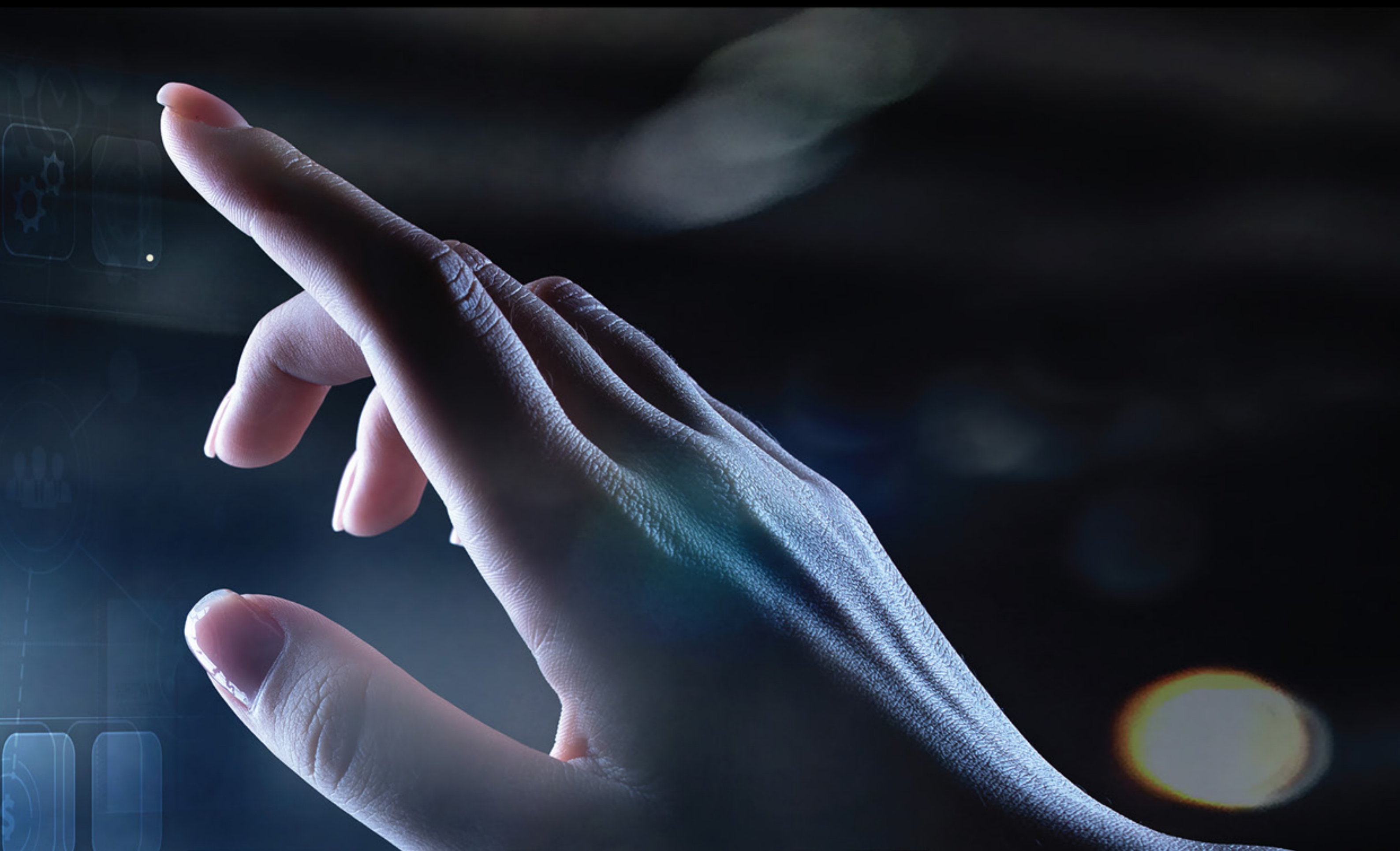
White paper

Version 1.0, April 2024

Abstract

This study delves into the application of predictive analytics within the telco sector, with a particular focus on prepaid mobile services. It investigates the prediction of customer behaviors, such as the top-up propensity within 2 to 4 days and account balance before top-ups. The goal is to empower telco operators with data-driven insights to tailor their marketing strategies more precisely.

This research evaluates the effectiveness of Neural Network (NN) models, underscored by rigorous hyperparameter tuning and cross-validation processes, against the traditional Machine Learning (ML) models, namely Random Forest (RF) and Gradient Boosting Trees (GBT), which are currently in production at Altice Labs. Innovatively, it incorporates pre-processing and feature selection techniques not previously used in traditional ML model development. The results demonstrate a significant performance leap of NN models over existing ML counterparts in accurately predicting customer actions. By providing telco operators with a more nuanced understanding of customer behavior patterns, this study offers insights into enhancing predictive models in the telco sector.



Introduction

In the ever-evolving telco sector, understanding and predicting customer behavior, especially in the realm of prepaid mobile services, is crucial for tailoring effective marketing strategies. Prepaid services, distinguished by customers paying in advance for telco services such as calls, SMS, MMS, and internet access, offer users control over their spending and avoid the contractual obligations associated with postpaid plans [1], [2]. However, the unpredictability of top-up events poses challenges for telco operators in their efforts to optimize marketing campaigns that meet individual customer needs [3], [4].

This article presents a comprehensive study focusing on leveraging Deep Learning (DL), more specifically Neural Networks (NNs) with more than two hidden layers, to enhance predictive models within the prepaid mobile services domain. Specifically, it investigates the propensity of customers to top-up within specific time frames (two, three, or four days) and estimates account balances before top-up events. By accurately predicting these aspects, telco operators can more effectively time their marketing communications, thereby increasing the relevance and efficiency of promotional efforts [6].

Our primary objective and challenge are to assess the viability of NNs in these specific use cases. For this purpose, we evaluate the performance of NN models, refined through extensive hyperparameter tuning and cross-validation, against established Machine Learning (ML) models, namely, Random Forest (RF) and Gradient Boosting Trees (GBT) [7], previously deployed by Altice Labs' FOCUS team. These traditional models have demonstrated their effectiveness and maturity by being fully integrated and deployed. Nevertheless, our study aims to explore the potential of NNs to surpass these models by incorporating innovative pre-processing and feature selection techniques not utilized in their development.

Motivated by recent shifts in the usage of prepaid mobile services in Portugal, where a notable decline was observed before a temporary uptick in 2021 due to eased pandemic-related travel restrictions [8], [9], and [10], this study addresses the urgent need to reverse the declining trend and encourage the adoption of prepaid services. Accurate predictions can significantly enhance customer experiences by allowing for the delivery of personalized offers at the most opportune moments, fostering customer loyalty and satisfaction [6]. Inadequate predictions could lead to missed marketing opportunities or, conversely, excessive outreach that may frustrate customers. The introduction of NNs into this context aims to provide a level of accuracy in predicting customer behavior that traditional ML models may not achieve, thus supporting telco operators in crafting more effective and personalized marketing campaigns.

This article outlines the process of designing data pipelines for the intake and transformation of relevant data, including customer top-up and balance history, which feeds into the predictive models. It also discusses the integration of performance indicators for model oversight and the employment of model validation techniques to mitigate risks associated with prediction errors. Through this study, we not only demonstrate the superior predictive capability of NN models but also offer insights into their potential to improve predictive analytics in the telco sector, thereby enabling operators to enhance customer retention and engagement through more personalized and attentive services.

State of the art

In the world of predictive analytics for the telco sector, particularly regarding prepaid mobile services, the challenge of accurately predicting customer top-up behavior is characterized by the limited number of publicly available studies on the topic. This gap primarily results from the proprietary nature of industry research, where findings are closely kept to maintain competitive advantage. However, by examining available predictive model research, valuable insights can be gathered to inform our investigation.

P. M. Alves et al.'s work [11] underscores the importance of detailed customer profiling in the telco industry to enhance service quality and retain customers, emphasizing predictive analytics' role in understanding and anticipating customer actions. This aligns with our study's focus on predicting top-up propensity and account balances before top-up. While Alves et al. applied sliding window Multiple Linear Regression for monthly top-up predictions, our research explores NN models for their potential to offer more accurate predictions in specified timeframes.

The application of Recency, Frequency, and Monetary value (RFM) analysis in previous studies for customer data analysis, particularly to identify top-up patterns [11], mirrors our approach to feature engineering. This methodological parallel extends to our consideration of seasonal trends in top-up behavior, informed by existing literature [11], to refine our predictive modeling.

Churn prediction studies, while distinct, provide a foundation from which to draw parallels and adapt methodologies for top-up behavior prediction. Bharadwaj et al.'s exploration of Logistic Regression and Multilayer Perceptron (MLP) models in churn prediction [12], achieving high accuracy, illustrates the potential of advanced predictive models to navigate the complexities of telco datasets. Similarly, the success of Deep NNs in outperforming traditional models in churn prediction studies hints at their applicability in predicting top-up behaviors [13].

Feature selection emerges as a critical step in refining predictive models, with methodologies aimed at identifying the most impactful variables for customer behavior prediction. The use of Mutual Information (MI) for feature selection, as done by Idris and Khan [14], aligns with our approach, as does the use of SHAP [15] as a feature selection technique by Dumitrache et al. [16], emphasizing the elimination of redundant features to enhance model performance.

Aryan Raj and Vetrithangam D. [17] bring a new perspective by using resampling techniques such as SMOTE [18] to address class imbalance issues. This aspect is particularly relevant to our problem, considering the similar challenges faced in dealing with an imbalanced dataset.

The reviewed literature, despite not directly addressing top-up behavior prediction, offers essential methodologies and insights that inform our approach. The relevance of churn prediction studies to our research lies in the shared goal of understanding and anticipating customer behavior to improve retention and satisfaction. This study seeks to extend these methodologies to the unique challenges of predicting top-up behavior and balance before top-up in prepaid mobile services, aiming to provide telco operators with actionable insights to tailor their marketing strategies effectively.

Methodology

Top-up propensity prediction in 2, 3, and 4 days (binary classification)

Our analysis utilizes 60 days of historical data, focusing on top-up and balanced history to generate aggregated features for our model. These features, including averages, standard deviations, counts, and maximum and minimum values, are specifically tailored to reflect customer engagement within this timeframe. This 60-day period balances inclusivity and relevance: longer periods might capture less engaged customers, diluting predictive accuracy, while shorter spans might exclude too many customers due to a smaller dataset. Therefore, we exclude customers who haven't topped up in the 60 days leading to our reference day. We set a specific reference day, typically at the start of a month, to anchor our analysis consistently. The concept of a ***top-up cycle***, represented by the orange and red periods in **Figure 1**, is central to our feature construction, defined as the time from a customer's last top-up before the reference day to their first top-up after it. For feature construction, we consider the period up to the day before the reference day, as shown by the brackets in **Figure 1**, resulting in the concept of the ***last top-up cycle***. The period following the reference day, including the day itself, is used for label construction. We also define the ***last n days*** as the critical three-day interval leading up to the reference day, specifically for feature creation. This interval focuses on capturing the most immediate and relevant customer behavior before a top-up event.

Labels are assigned based on whether a customer tops up within a set interval (2, 3, or 4 days) from the reference day, creating binary classification models. Customers topping up within this period are labeled '1', and those who don't are labeled '0'. This process establishes a binary classification problem aimed at predicting the propensity of a customer topping up within the specified timeframes. In **Figure 1**, the propensity periods are marked by dotted lines: purple for two days, green for three, and yellow for four, starting from the reference day.

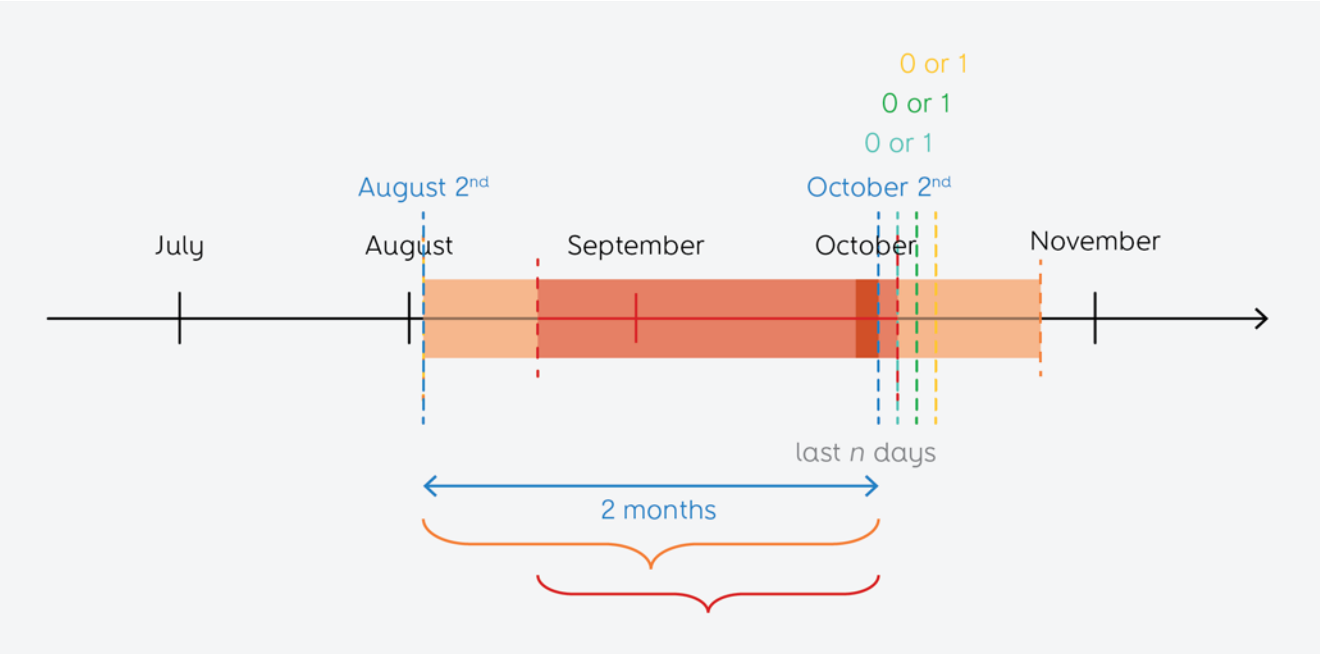


Figure 1 – Methodology example for top-up propensity within the next 2, 3, or 4 days

Evaluation metrics

Given the extremely unbalanced nature of our binary classification datasets, we must carefully select our evaluation metrics. Accuracy, for instance, is not suitable because it can misleadingly inflate performance by favoring the majority class, such as predicting all customers as *non-top-up*, which would yield high accuracy despite failing to identify any *top-up* cases. Therefore, we use the F1-Score as our metric. It balances precision and recall through their harmonic mean, providing a more accurate assessment for the minority class in our classification problem [19], [20].

Balance before top-up prediction (regression problem)

This approach for feature temporal constraints mirrors the top-up propensity prediction, focusing on the last two months leading up to a chosen reference day. The features considered here are identical to those in the top-up propensity prediction, emphasizing the relevance of both the *last top-up cycle* and the *last n days* leading up to the reference day.

For balance prediction, we label data based on the customer's account balance right before their first top-up following the reference day, which includes the day itself. Our model training involves analyzing the month after the reference day to identify each customer's first top-up and predict the balance. This approach results in fewer data samples for training compared to the top-up propensity predictions because it only considers customers who topped up in the two months before and the month following the reference day.

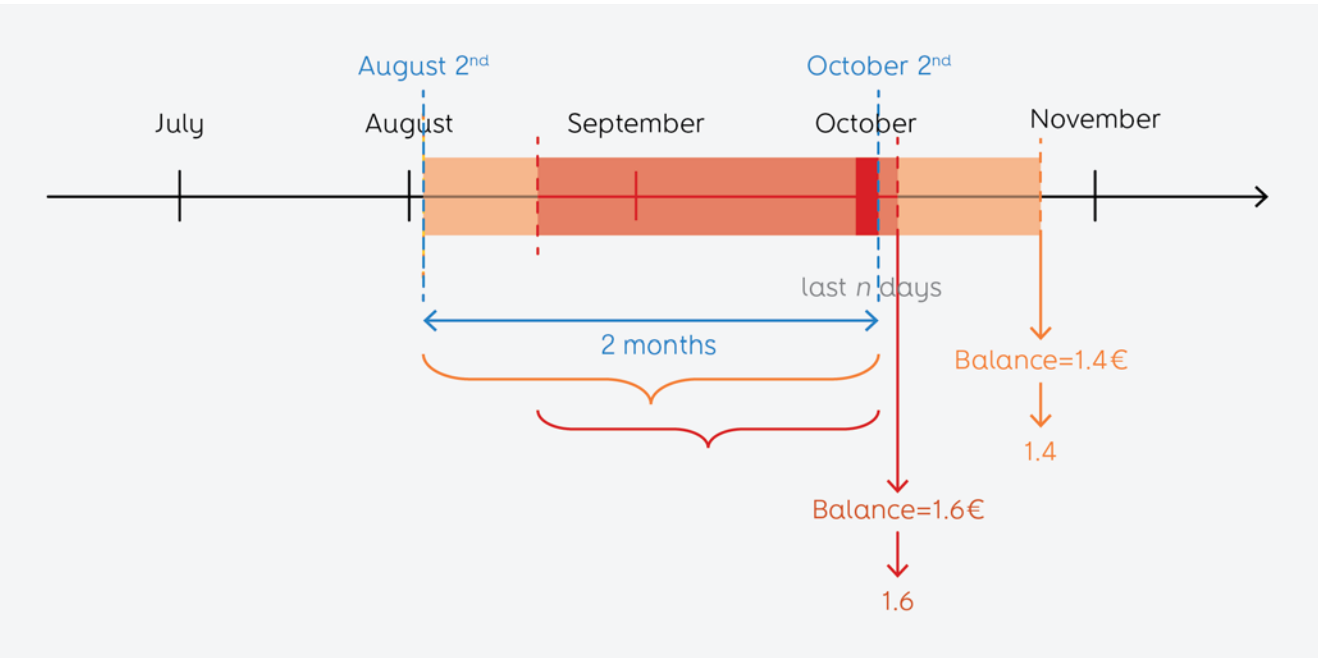


Figure 2 – Methodology for account balance before top-up

Evaluation metrics

In this regression problem, we employ two distinct evaluation metrics: the Mean Absolute Error (MAE) [21] and the Hit Ratio. MAE calculates the absolute difference between actual and predicted balances before top-up, directly measuring prediction accuracy. The Hit Ratio, on the other hand, assesses the proportion of instances where the predicted balance exceeds the actual pre-top-up balance, indicating our model's ability to anticipate top-ups based on balance evolution. However, relying solely on the Hit Ratio could be misleading. For instance, a model predicting a balance of 20 euros when the actual balance is 2 euros would still improve the Hit Ratio, despite the prediction being significantly off. Therefore, we aim to balance these metrics, ensuring our predictions not only anticipate actual top-ups but also accurately reflect the customers' real balances.

Daily performance testing

Top-up propensity

We evaluate the model's performance daily over the test month using a sliding window approach. Each day is a new reference point from which the model predicts customer top-up propensity within the next 2, 3, or 4 days. As the reference day shifts, the 60-day feature construction window adjusts accordingly, generating 31 distinct datasets for each day of the month. This method provides a detailed assessment of the model's consistency and accuracy across different times. We analyze the model's F1-Score, precision, and recall daily to identify performance trends or patterns needing attention.

Balance before top-up prediction

Like the top-up propensity model, we utilize a sliding window approach for feature construction, with a 60-day window that adjusts as the reference day progresses, creating 31 distinct datasets for each month's day. For balance prediction, we extend our analysis beyond the training set's constraints. For training, we considered only customers who topped up in the 60-day window before the reference day and the month after the reference day (including the reference day itself); for testing, we evaluated all customers who topped up at least once in the 60 days before the reference day. This adjustment allows us to observe the evolution of the predicted balance for each customer daily, offering a dynamic view of anticipated balances. Each day, we compare the predicted balance directly with the actual balance for customers who topped up on the reference day, aiming to measure our model's ability to anticipate imminent top-ups. We employ the Hit Ratio to assess our success in anticipating top-ups and the MAE to ensure the accuracy of our predicted balances.

Feature engineering

For our analysis, we strategically chose July 3rd as the training reference day to capture a high activity phase, ensuring our dataset, with features defined in the period from May 3rd to July 3rd, is enriched with diverse customer interactions. In contrast, our testing period spans from August 2nd to October 2nd, with October 2nd serving as the test reference date to examine performance across a different activity range.

Across both applications, our datasets include 58 aggregated features of customer top-up and balance behavior. Specifically, for top-up propensity, we generated three separate datasets tailored to 2, 3, and 4-day predictions, resulting in 58.472 entries for training and 60.428 for testing. Meanwhile, the balance prediction dataset is slightly smaller, with 28.559 entries for training and 25.242 for testing. This size discrepancy arises from our label generation approach for balance prediction, which includes only customers active in the subsequent month post-reference day, as detailed in Section 3. This criterion was applied consistently across training and testing phases to identify the most effective NN architecture. However, for daily model evaluations with the established architecture, the balance prediction use case expands to include a broader set of customers. Similar to the top-up propensity analysis, we include any customer who has made at least one top-up in the 60 days leading up to the reference day, thereby increasing the dataset's scope for this specific evaluation.

To optimize the performance of our NN architectures for each use case, we embarked on an extensive feature engineering and hyperparameter tuning process, including cross-validation. Our approach to feature selection was multifaceted: we evaluated the use of all 58 features, selected subsets based on Mutual Information (MI) scores [22], for higher relevance, and applied SHAP analysis [15], for further insight into feature importance. Specifically, if models performed better with all features, SHAP analysis helped pinpoint the most influential ones, guiding us to refine our feature subset based on SHAP values to potentially enhance model performance.

We also explored different normalization techniques, such as min-max normalization and standardization [23], to improve our models' input data handling. For the top-up propensity case, we experimented with the datasets' original class distribution (approximately 5% top-up customers), applying balancing strategies like SMOTE [18] and undersampling [24] to address the class imbalance. Moreover, we tested both standard binary cross-entropy [25] and weighted binary cross-entropy [26] loss functions for our classification models, aiming to improve class differentiation in this unbalanced scenario. This comprehensive testing strategy was designed to identify the most effective NN architecture and feature set for each specific use case.

Best performing Neural Networks

Top-up propensity use case

After extensive experimentation, we identified the most effective model architecture for predicting top-up propensity within 2 days. Applying the same architecture to the 3-day and 4-day predictions revealed that different sets of input features yielded optimal results for each timeframe. Min-max scaling and using the original class distribution without applying balancing techniques achieved the best performance. This effectiveness likely stems from similar feature distributions across classes, where balancing techniques failed to enhance class distinction, leading to overfitting on the training data and inadequate generalization to real-world testing scenarios.

For input features, utilizing all 58 features was most effective for the 2-day top-up propensity prediction. In contrast, for the 3-day and 4-day predictions, selecting a subset of features through SHAP analysis improved performance, resulting in the use of 25 features for 3-day and 29 features for 4-day predictions.

The optimal model architecture, consistent across the three prediction intervals, is detailed in **Figure 3**. This architecture was employed to train the model for 100 epochs at a batch size of 64, using the Adam optimizer with a learning rate of 0.001 and weighted binary cross-entropy as the loss function. To prevent overfitting, L1 regularization [27] with a learning rate of 0.001 was applied to the second hidden layer.

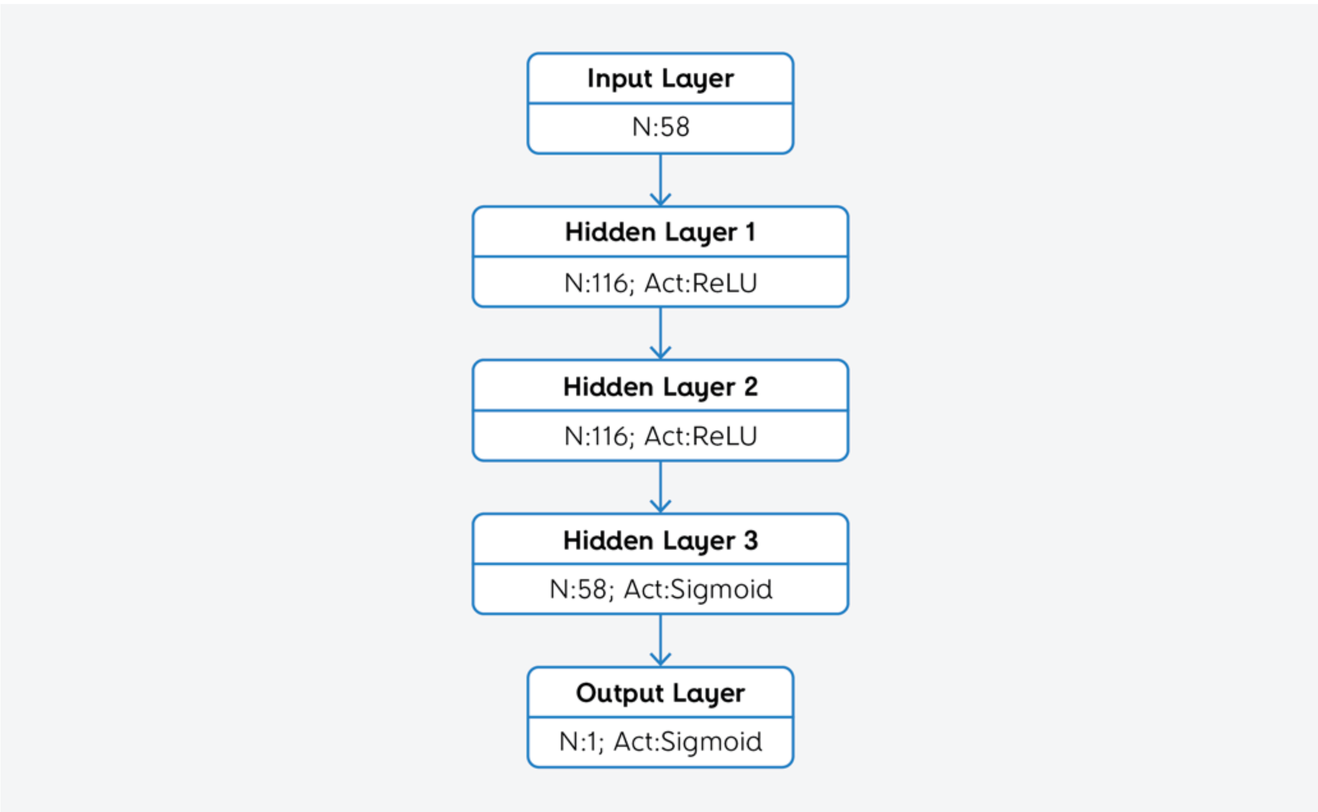


Figure 3 – Optimal architecture of the NN for the top-up propensity use case, showcasing layers, their number of Neurons (N) and Activation Functions (Act)

Balance before top-up use case

In the account balance prediction scenario, the model achieved superior performance using a subset of 11 features, selected based on Mutual Information (MI) scores rather than the full-feature set, and applied min-max scaling for data normalization. The architecture that proved most effective for this use case is depicted in **Figure 4**. The training involved 150 epochs with a batch size of 32, utilizing the Adam optimizer at a learning rate of 0.001. To combat overfitting, we applied L2 regularization [27] with a coefficient of 0.01 in both the first and second hidden layers. Additionally, to further ensure model generalization, early stopping [27] was implemented with a patience setting of 30 epochs, which effectively reduced the total number of training epochs required.

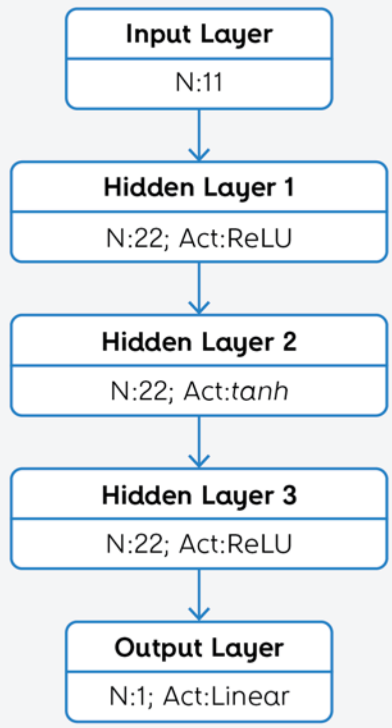


Figure 4 – Optimal architecture of the NN for the balance use case, showcasing layers, their number of Neurons (N), and Act. functions (Act)

Comparison between Neural Networks and traditional Machine Learning models daily performance

In this section, we evaluate the daily performance of our optimal NN models, as depicted in **Figures 3 and 4**, throughout the test month of October. We then compare these findings with the outcomes achieved by the FOCUS team using traditional ML models. This daily evaluation sheds light on the NNs' consistency and adaptability under various conditions, offering a grounded view of their real-world applicability.

The FOCUS team previously implemented a Random Forest (RF) classifier [28] for the top-up propensity prediction and a Gradient Boosting Trees (GBT) regressor [28] for the balance prediction, detailing their configurations in **Table 1**. They also selected July 3rd as the training reference day to determine the most effective models. However, our method diverges in feature selection; while the FOCUS team favored Information Value (IV) [29], we opted for Mutual Information (MI) and SHAP analysis. This strategic variance in feature selection has resulted in our NN models delivering more precise predictions.

Top-up propensity use case		Balance before top-up use case	
RF classifier		GBT regressor	
max_depth	11	learning_rate	0.1
criterion	entropy	criterion	friedman_mse
min_samples_split	2	min_samples_split	0.1
n_estimators	500	n_estimators	100
classification threshold	0.255	loss	absolute
		max_depth	7
		min_samples_leaf	0.05

Table 1 – Existing Machine Learning models for top-up propensity and balance prediction

Top-up propensity use case

In **Figure 5**, we observe the daily performance comparison between our NN (blue) from **Figure 3** and the previously developed RF classifier (red) over October. Our NN exhibits a performance pattern, performing better at the week's start and declining on weekends and on October 5th, a national holiday. This pattern, consistent across the 2, 3, and 4-day propensities, might stem from selecting July 3rd, a Monday, as our training day, potentially biasing the model towards early-week behaviors. Alternatively, it could reflect the naturally erratic weekend behavior.

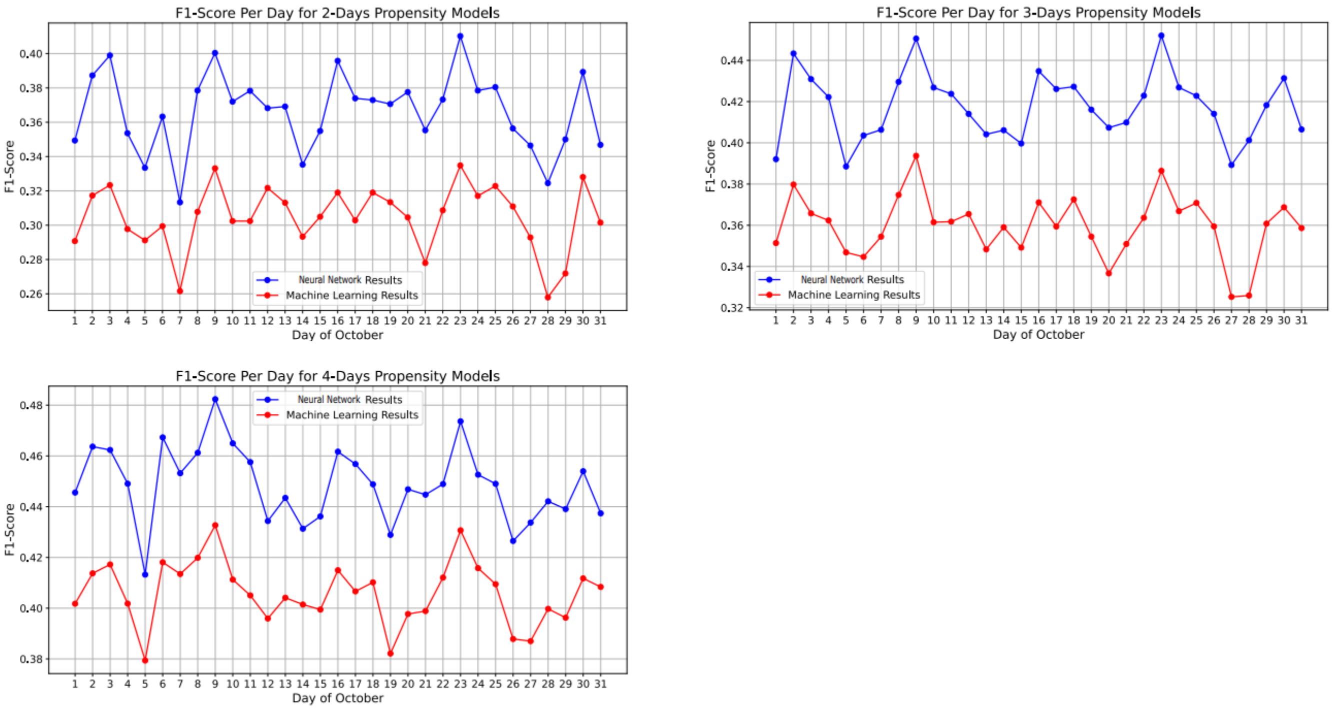


Figure 5 – F1-Score obtained with the NN classifier (blue line) and with the RF classifier (red line) for the propensity use case over the testing month of October

Our analysis reveals that the NN consistently outperformed the RF classifier across all three prediction time frames in October, indicating an overall performance improvement with the NN. This superiority is particularly notable in achieving higher precision without compromising recall, which is crucial for reducing false positives in top-up predictions. This success underscores the efficacy of our NN approach, benefiting from our unique feature selection and architectural design. Both models displayed similar temporal performance trends, suggesting a common sensitivity to time-related behavioral shifts.

Balance before top-up use case

In **Figure 6** (on the next page), we compare the Hit Ratio (left) and Mean Absolute Error (MAE) (right) between the GBT regressor (red) and our NN from **Figure 4** (blue) for the balance before the top-up use case, aiming for an optimal balance between these metrics.

Throughout October, our NN regressor consistently achieved a higher Hit Ratio and significantly lower MAE compared to the GBT regressor, indicating more accurate and reliable predictions. Notably, both the Hit Ratio and MAE were better at the month's start but tended to worsen as the month progressed. This pattern may stem from both models being trained on July 3rd, suggesting they may be better at capturing customer behavior typical of the month's beginning and less adept at generalizing across the entire month. Despite this, our NN regressor was able to maintain a high Hit Ratio and keep the MAE below 2 euros, underscoring its effectiveness in predicting the balance before top-up with remarkable accuracy.

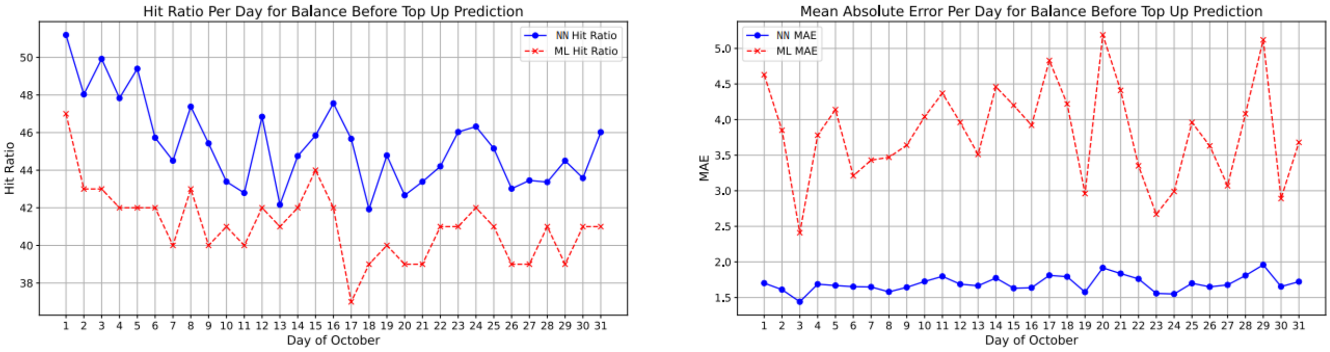


Figure 6 – Hit Ratio (left plot) and MAE (right plot) obtained with the NN regressor (blue line) and with the GBT regressor (red line) for the prediction of the account balance before top over the testing month of October

Conclusions and future work

Conclusions

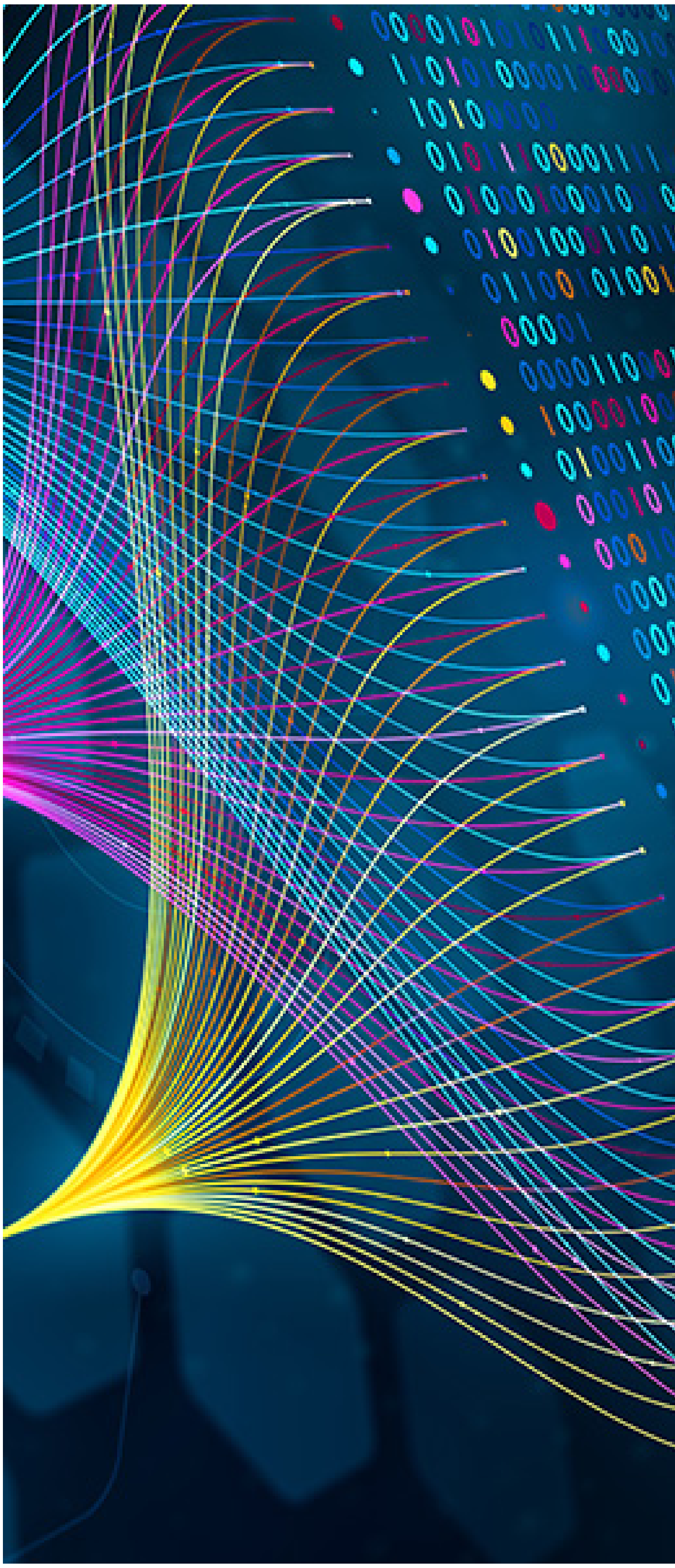
This article presented an in-depth exploration of predictive analytics within the telco sector, focusing on two primary objectives: predicting top-up propensity and predicting account balances before top-ups in prepaid mobile services. Our investigation involved a detailed comparison of NNs against traditional ML models previously utilized by the FOCUS team: an RF classifier for top-up propensity and a GBT regressor for balance prediction.

A significant portion of our research was dedicated to data analysis and feature selection, where we diverged from traditional IV methods in favor of MI scores and SHAP values. This approach allowed us to tailor features specifically to the needs of each predictive model, enhancing the NNs' ability to discern complex patterns within the data.

Through rigorous experimentation, including hyperparameter tuning and cross-validation, we identified optimal model architectures and preprocessing techniques. Our findings highlighted the superiority of NN models in both use cases, as evidenced by improved performance metrics over the ML models employed by the FOCUS team. Notably, our NN models demonstrated a remarkable ability to deliver precise predictions with lower error rates.

An interesting aspect of our research was the parallel daily performance trends observed between the ML and NN models, reflecting a sensitivity to temporal patterns in customer behavior. This underscored the importance of model training periods in predicting performance.

Reflecting on the broader implications of our work, this article illustrates the potential of NNs for enhancing predictive analytics within the telco sector. Our research underscores the value of adopting advanced analytical techniques to improve service offerings and customer understanding in an industry characterized by rapid technological advancements and evolving customer needs.



Future work

There are numerous avenues for future enhancements to our models. For the top-up propensity use case, we applied the same NN architecture across the three prediction intervals (2, 3, and 4 days). A potential improvement could involve fine-tuning the NN architecture separately for each prediction period, potentially enhancing performance by tailoring each model to its specific timeframe.

Additionally, adjusting the timing of training could address observed performance patterns. For instance, in the top-up propensity case, we noted superior performance at the beginning of the week compared to the end. This variation might be influenced by our choice of reference day for training. To explore this further, we could experiment with training models using reference days from both the start and the end of the week, aiming to capture a broader spectrum of customer behaviors throughout the week. Similarly, for the balance prediction use case, which showed improved performance at the month's start and declined towards the end, setting training reference days at the beginning, middle, and end of the month might enable our models to better capture the full range of monthly customer behaviors.

Integrating time series analysis into these use cases presents another promising direction. Transitioning from aggregated features to those that reflect the temporal evolution of customer behavior could allow the use of models like Long Short-Term Memory (LSTM) networks, potentially offering deeper insights into behavioral patterns.

These suggested areas of research build upon our current findings and are crucial for adapting to the dynamic nature of customer behavior in the telco sector. They promise to enhance the relevance and effectiveness of predictive models in this rapidly evolving industry.

References

[1] H. Bouwman, H. de Vos, and T. Haaker, ***Mobile Service Innovation and Business Models, Chapter 4***. Springer Berlin Heidelberg, 2008. doi: https://doi.org/10.1007/978-3-540-79238-3.

[2] Y. Lin, M. Chang, and H. C. Rao, "Mobile prepaid phone services," ***IEEE Personal Communications***, vol. 7, no. 3, pp. 6–14, 2000, doi: https://doi.org/10.1109/98.847918.

[3] A. Khalatyan, "Churn Management in Telecommunications: Challenging the innovative Capability of Data Mining Tools," ***iSCHANNEL***, 2010.

[4] Y. Zheng et al., "TelcoFlow: Visual exploration of collective behaviors based on telco data," in ***2016 IEEE International Conference on Big Data (Big Data)***, 2016, pp. 843–852. doi: https://doi.org/10.1109/BigData.2016.7840677.

[5] Altice Labs, "Altice Labs Homepage," Available at: http://www.alticelabs.com, Accessed: January 2024.

[6] S. Qiu, L. Wu, Y. Yang, and G. Zeng, "Offering the right incentive at the right time: Leveraging customer mental accounting to promote prepaid service," ***Ann Tour Res***, vol. 93, p. 103367, 2022, doi: https://doi.org/10.1016/j.annals.2022.103367.

[7] P. Singh, ***Learn PySpark: Build Python-based Machine Learning and Deep Learning Models***. Apress, 2019.

[8] ANACOM, "Mobile Internet traffic increased 39.5% in the 1st half of 2022," Available at: https://anacom.pt/render.jsp?contentId=1727788, Accessed: January 2024.

[9] ANACOM, "In 2022 the number of subscribers increased by 3.4%," Available at: https://www.anacom-consumidor.pt/-/em-2022-o-numerode-assinantes-aumentou-3-4-, Accessed: January 2024.

[10] ANACOM, "Mobile services," Available at: https://anacom.pt/render.jsp?contentId=1745046, Accessed: January 2024.

[11] P. M. Alves, R. Â. Filipe, and B. Malheiro, "Towards Top-Up Prediction on Telco Operators," in ***Progress in Artificial Intelligence***, G. Marreiros, F. S. Melo, N. Lau, H. Lopes Cardoso, and L. P. Reis, Eds., Springer International Publishing, 2021, pp. 573–583.

[12] S. Bharadwaj, B. S. Anil, A. Pahargarh, A. Pahargarh, P. S. Gowra, and S. Kumar, "Customer Churn Prediction in Mobile Networks using Logistic Regression and Multilayer Perceptron(MLP)," in ***2018 Second International Conference on Green Computing and Internet of Things (ICGCIoT)***, 2018, pp. 436–438. doi: 10.1109/ICGCIoT.2018.8752982.

[13] N. Putu Oka H and A. Setyo Arifin, "Telecommunication Service Subscriber Churn Likelihood Prediction Analysis Using Diverse Machine Learning Model," in ***2020 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT)***, 2020, pp. 24–29. doi: 10.1109/MECnIT48290.2020.9166584.

[14] A. Idris and A. Khan, “Customer churn prediction for telecommunication: Employing various various features selection techniques and tree based ensemble classifiers,” in ***2012 15th International Multitopic Conference (INMIC)***, 2012, pp. 23–27. doi: 10.1109/INMIC.2012.6511498.

[15] Y. Lai, W. Sun, J.-D. Schmöcker, K. Fukuda, and K. W. Axhausen, “Explaining a century of Swiss regional development by deep learning and SHAP values,” ***Environ Plan B Urban Anal City Sci***, vol. 50, no. 8, pp. 2238–2253, 2023, doi: 10.1177/23998083221116895.

[16] A. Dumitrache, A. A. M. Nastu, and S. Stancu, “Churn Prediction in Telecommunication Industry: Model Interpretability,” ***Journal of Eastern Europe Research in Business and Economics***, vol. 2020, p. Article ID 241442, 2020, doi: 10.5171/2020.241442.

[17] A. Raj and D. Vetrithangam, “Prediction of customer churn using resampling and ensemble classification algorithms on telecom dataset,” in ***2023 11th International Conference on Emerging Trends in Engineering & Technology - Signal and Information Processing (ICETET – SIP)***, 2023, pp. 1–7. doi: 10.1109/ICETET-SIP58143.2023.10151659.

[18] N. V Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic Minority Over-sampling Technique,” ***Journal of Artificial Intelligence Research***, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.

[19] Ž. . Vujovic, “Classification Model Evaluation Metrics,” ***International Journal of Advanced Computer Science and Applications***, vol. 12, no. 6, 2021, [Online]. Available: http://dx.doi.org/10.14569/IJACSA.2021.0120670.

[20] M. Grandini, E. Bagli, and G. Visani, “Metrics for Multi-Class Classification: an Overview.” 2020.

[21] V. Plevris, G. Solorzano, N. P. Bakas, and M. E. A. Ben Seghier, “Investigation of performance metrics in regression analysis and machine learning-based prediction models,” in ***8th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS Congress 2022)***, in ECCOMAS Congress;8th European Congress on Computational Methods in Applied Sciences and Engineering, vol. 8. European Community on Computational Methods in Applied Sciences, Mar. 2022, pp. 1–25. [Online]. Available: https://doi.org/10.23967/eccomas.2022.155.

[22] A. Kraskov, H. Stögbauer, and P. Grassberger, “Estimating mutual information,” ***Phys. Rev. E***, vol. 69, no. 6, p. 66138, 2004, doi: 10.1103/PhysRevE.69.066138.

[23] S. Vieira, W. H. L. Pinaya, and A. Mechelli, “Chapter 2 – Main concepts in machine learning,” in ***Machine Learning***, A. Mechelli and S. Vieira, Eds., Academic Press, 2020, pp. 21–44. doi: https://doi.org/10.1016/B978-0-12-815739-8.00002-X.

[24] R. Mohammed, J. Rawashdeh, and M. Abdullah, “Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results,” in ***2020 11th International Conference on Information and Communication Systems (ICICS)***, 2020, pp. 243–248. doi: 10.1109/ICICS49469.2020.239556.

[25] T. Chakraborty and U. Kumar, "Loss Function," in ***Encyclopedia of Mathematical Geosciences***, Springer International Publishing, 2023, pp. 774–779. [Online]. Available: https://link.springer.com/referencework/10.1007/978-3-030-85040-1?utm_medium=referral&utm_source=google_books&utm_campaign=3_pier05_buy_print&utm_content=en_08082017.

[26] S. Panchapagesan et al., "Multi–task learning and Weighted Cross–entropy for DNN–based Keyword Spotting," in ***Interspeech 2016***, 2016. [Online]. Available: https://www.amazon.science/publications/multi-task-learning-and-weighted-cross-entropy-for-dnn-based-keyword-spotting.

[27] P. J, ***TensorFlow Developer Certification Guide: Crack Google's official exam on getting skilled with managing production–grade ML models***. GitforGits, 2023.

[28] P. Singh, ***Learn PySpark: Build Python–based Machine Learning and Deep Learning Models***. Apress, 2019.

[29] D. Bhalla, "Weight of Evidence (WOE) and Information Value (IV) Explained." 2015.

Acronyms

DL	Deep Learning
GBT	Gradient Boosting Trees
IV	Information Value
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MI	Mutual Information
ML	Machine Learning
MLP	Multilayer Perceptron
MMS	Multimedia Messaging Service
NN	Neural Network
RF	Random Forest
RFM	Recency, Frequency, and Monetary value
SHAP	SHapley Additive exPlanations
SMOTE	Synthetic Minority Oversampling Technique
SMS	Short Message Service

Authors

Beatriz Mesquita

Professional Intern at FOCUS team / Junior Data Scientist

Altice Labs, Portugal

beatriz-s-goncalves@alticelabs.com

https://www.linkedin.com/in/beatriz-gonçalves-13195b210/

Bernardo Duarte

Senior Data Scientist

Altice Labs, Portugal

bernardo-x-duarte@alticelabs.com

https://www.linkedin.com/in/bernardoxnduarte/

Francisco Silva

Senior Data Scientist

Altice Labs, Portugal

francisco-c-silva@alticelabs.com

Petia Georgieva

Associate Professor with Aggregation of the Department of Electronics, Telecommunications, and Informatics at the University of Aveiro

DETI/IEETA, University of Aveiro / Instituto de Telecomunicações, Portugal

petia@ua.pt

https://www.linkedin.com/in/petia-georgieva-4b7992a3/

Contacts

Address

Rua Eng. José Ferreira Pinto Basto

3810 - 106 Aveiro (PORTUGAL)

Phone

+351 234 403 200

Media

contact@alticelabs.com

www.alticelabs.com