DOI: 10.15849/IJASCA.240730.02

Int. J. Advance Soft Compu. Appl, Vol. 16, No. 2, July 2024
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Toward a Responsive System for Precision Marketing Based on RFM Model, Deep learning and Features Importance Ranking: A Case Study of Morocco

Nouhaila El Koufi, Abdessamad Belangour

Laboratory of Information Technology and Modeling LTIM, Hassan II University, Casablanca 28830, Morocco
Corresponding author: elkoufi.nouhaila1@gmail.com

Abstract

The omnichannel business model has undergone a revolutionary transformation, propelled by technological advancements. This evolution has effectively blurred the boundaries between physical and digital commerce, enabling seamless integration across various platforms and touchpoints. Utilizing advanced technologies like AI, big data analytics, and Internet of Things (IoT) devices, businesses now possess the capability to tailor customer experiences, streamline supply chain operations, and optimize overall efficiency to unprecedented levels. This transformative shift has empowered companies to effectively respond to the dynamic needs of contemporary consumers, who demand a cohesive, convenient, and personalized journey across all channels, whether online, in-store, or via mobile devices. Harnessing the wealth of data generated from these channels offers invaluable insights for strategic decision-making and further innovation. In this paper, we introduce a response system designed to utilize advanced feature engineering techniques, including the RFM model. Our aim is to provide valuable managerial insights. The dataset used in this study was gathered from a Moroccan omnichannel business based in Rabat, employing sophisticated data collection and labeling techniques. Subsequently, the data underwent preprocessing, cleaning, and balancing, and significant features were identified through advanced methods. To assess the effectiveness of our approach, we conducted experiments comparing the performance of our proposed deep learning model with decision trees and Naive Bayes, both with and without genetic algorithms. The results of these experiments consistently demonstrated that our proposed model outperformed the other methods.

Keywords: *Deep learning, Data mining, Features Engineering, Multiclass, Precision Marketing, RFM, Omni-channel Business.*

1 Introduction

Omnichannel strategies have become pivotal in driving economic growth for countries worldwide. By seamlessly integrating various channels such as online, mobile, and physical stores, businesses can enhance customer experiences, improve operational efficiency, and ultimately bolster their bottom line. The interconnected nature of

Received 10 February 2024; Accepted 31 May 2024

omnichannel approaches not only facilitates smoother transactions but also enables businesses to gather rich data insights on consumer behavior. This wealth of data, when effectively leveraged, equips enterprises and policymakers with the necessary tools to make informed decisions, refine marketing strategies, and drive overall economic productivity [1]. Countries that prioritize and invest in omnichannel infrastructures are better positioned to adapt to evolving consumer preferences, foster innovation, and sustain long-term economic competitiveness in an increasingly digital landscape.

Moreover, the utilization of advanced analytics techniques such as deep learning and the RFM (Recency, Frequency, Monetary) model plays a pivotal role in delving deeper into customer behaviors. Deep learning algorithms, fueled by vast amounts of data, have the capability to uncover intricate patterns and correlations within customer datasets that traditional methods might overlook [2]. By harnessing the power of deep learning, businesses can gain invaluable insights into customer preferences, purchasing habits, and sentiment analysis, enabling them to tailor products and services more effectively to meet individual needs. Concurrently, the RFM model offers a structured framework for segmenting customers based on their recency, frequency, and monetary value of transactions. This segmentation enables businesses to prioritize high-value customers, personalize marketing strategies, and optimize resource allocation, thereby maximizing profitability and enhancing customer satisfaction. Through the synergistic integration of omnichannel strategies with advanced analytics techniques, countries can foster a thriving economic ecosystem driven by data-driven decision-making and customer-centric innovation [3].

The objective of this study is to introduce a response system for the implementation of precision marketing in the business sector. This framework covers a range of activities, such as customer classification, optimization of resource allocation, and management of inventory. To illustrate the practical use of this framework, we carried out a real-world case study with data from a Moroccan omni-channel retail business. The main goal of this study is to help retail managers identify unique traits of high-value customers and offer personalized precision marketing strategies. This method is expected to notably reduce advertising costs and enhance overall marketing effectiveness.

The rest of the document is organized as follows: Section 2 reviews related literature. Section 3 offers background information relevant to the foundation of this paper. Section 4 introduces, details, and examines the proposed system. Section 5 covers the results and experimental phase, along with a discussion on the implementation outcomes of the proposed system. Finally, Section 6 presents the conclusion and future directions for this study.

2 Literature Review

This literature review delves into a plethora of research endeavors undertaken by diverse scholars, focusing on the utilization of data mining and deep learning techniques for customer classification. In this section, we offer a succinct overview of select recent studies that have contributed to this field, shedding light on the innovative approaches and methodologies employed to address the classification challenges.

Sanjaya et al. [4] employed the decision tree method and customer data analysis to devise a precision marketing model for SME e-commerce. They demonstrated the efficacy of their solution using SME e-commerce data from orebae.com, revealing that the model effectively tailors marketing strategies to customer preferences, resulting in increased sales and reduced marketing costs. The findings underscore the positive outcomes achieved by the proposed model. In [5], Jue Chen proposed a classification model for precision marketing based on an enhanced Apriori algorithm, utilizing historical behavior data from e-commerce websites. The study highlighted significant improvements in computational efficiency, bias reduction, and accuracy enhancement attributable to the enhanced Apriori algorithm. Yulan Zhen [6] employed the decision tree algorithm to accurately classify customers into two classes, showcasing the efficacy of the proposed solution through a case study involving an e-commerce company's customer data. In [7], Wang and colleagues proposed the M-GNA-XGBoost model to predict product sales in online shops, integrating LSTM, GAN, and XGBoost techniques. This method's effectiveness was confirmed through real-world data, with a case study centered on the Jingdong store. El Koufi et al. [8] presented the Potential Customer Prediction Algorithm (PCPA), a machine learning-based algorithm for predicting potential clients and delivering personalized recommendations, using real-world data from a Moroccan bank. In [9], Cheng developed an evaluation index system (EIS) for precision marketing in commerce, utilizing an improved attention-interest-desire-memory-action (ADIMA) model. Principal component analysis (PCA) was used to identify key evaluation indices, which were then assessed with an artificial neural network (ANN) and optimized through k-means clustering (KMC). Additionally, in [10], a modified K-means clustering algorithm, enhanced with the silhouette coefficient, was presented to boost clustering accuracy and overall K-means performance. This refined algorithm was applied to precision marketing for ETC credit cards to pinpoint key factors influencing customer adoption.

3 Background

In this section, we offer a brief summary of the primary methods used in this paper. We clearly describe the RFM model, data mining techniques, and the deep learning approach applied for customer classification.

3.1 RFM

The RFM model serves as a powerful instrument for distinguishing among users based on their purchasing patterns. By analyzing three critical behavioral metrics, Recency, Frequency, and Monetary Value of purchases, this method identifies high-performing clients with excellence across these dimensions. Specifically:

Recent users show a propensity to repeat their purchasing behavior, with the Recency (R) score increasing as their activity becomes more current.

Users who engage in frequent purchases often demonstrate a tendency for repeat buying, leading to a higher Frequency (F) score corresponding to their activity level.

Customers who spend more, denoted as Monetary Value, within a specified timeframe are more likely to make additional purchases. Consequently, the Monetary (M) score escalates alongside increased spending.

This comprehensive model offers a nuanced understanding of user behavior, empowering businesses to effectively target and engage their most valuable customers.

3.2 Deep learning method

A deep neural network operates as a robust solution for tackling intricate challenges, often without extensive prerequisites. This sophisticated architecture, characterized by multiple layers, demands meticulous attention to various training parameters. Elements like network size, learning rate, and initial weights play pivotal roles in fine-tuning the model's predictive capabilities. By strategically adjusting these parameters, the neural network can effectively navigate through complex data patterns to deliver accurate output predictions. Consider $X = (X_1, X_2, \ldots X_n)$ as the input vector and Y as the corresponding output vector. Through the mapping of inputs to outputs, a learning function Y can be defined as: $Y = f(X)$, where each X_i belongs to the set of real numbers and each Y_i is binary, taking values from $\{0, 1\}$. This function encapsulates the relationship between the input features and the target output, enabling the model to learn and make predictions based on the provided data. To ensure that the output units become linearly separable by transforming the input space X of each layer, including a bias input, an activation function is employed. This function serves the purpose of introducing non-linearity into the network, allowing it to capture complex patterns and relationships within the data. By applying the activation function to the transformed input space, the neural network can effectively learn and represent intricate decision boundaries, enhancing its ability to accurately classify or predict outcomes.

3.3 Data mining

Here, we provide a brief overview of commonly employed methods for classification of customer value.

Naive bayes [11]: The naive bayes model is a simplistic yet effective classifier that leverages Bayes' Theorem to make predictions based on probability calculations.

Decision tree [12]: A decision tree model adopts a hierarchical, tree-like structure to facilitate decision-making. It generates classification rules by recursively partitioning the input space into subsets based on feature attributes, ultimately leading to well-defined classification outcomes.

4 Proposed Method

In this section, we introduce our system, which utilizes data mining techniques, deep learning, and the RFM model for learning purposes. Fig. 1 depicts the architecture of our system used in this study.

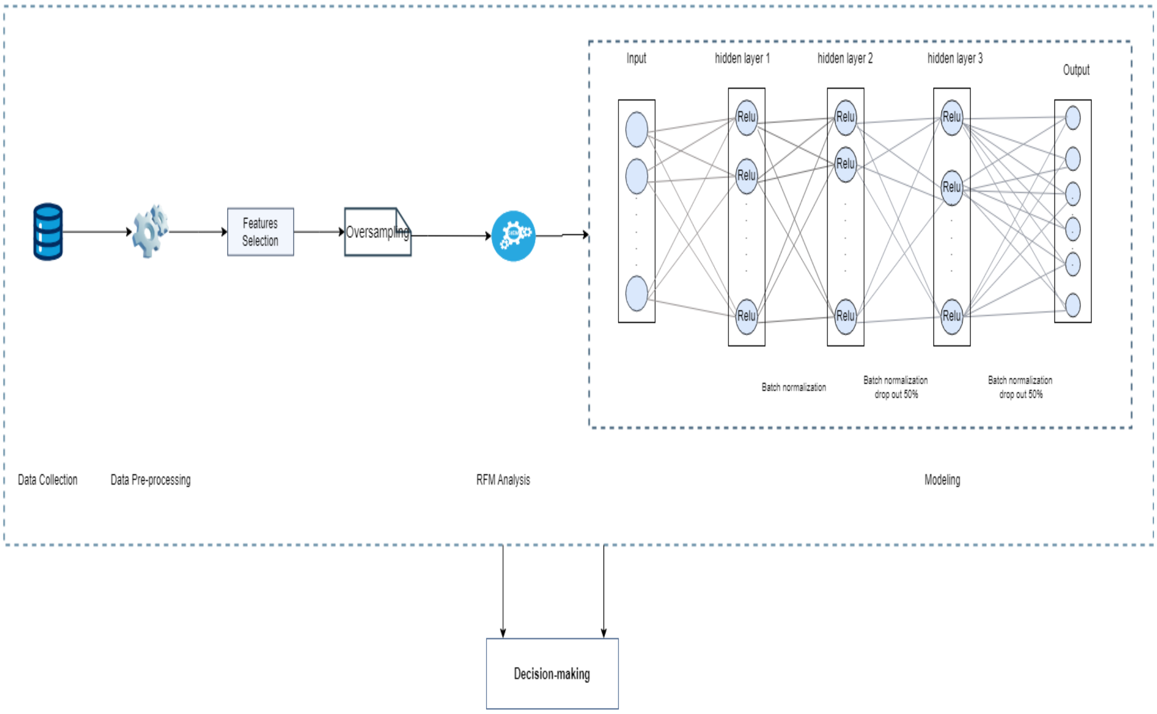

Fig. 1: Proposed system architecture

4.1 Data collection

Data collection primarily consists of gathering valuable information from the initial transactional data. The data collected in this study is real-world data from a Moroccan Omni-channel business. This business generally sells natural products, especially gluten-free products for people suffering from gluten sensitivity, and sugar-free products dedicated to people who suffer from diabetes. The dataset comprises 9841 records, each representing a customer transaction over a four months period. It contains valuable information about customer and user consumption habits, as well as some information about RFM value. The business operates both a local store in Rabat and an e-commerce website for product sales. Due to the original customization of their database system to meet specific requirements, we applied various techniques to adapt the data and extract features to align with our research objectives.

4.2 Data pre-processing

Data preprocessing stands as a vital step encompassing all necessary actions conducted before the modeling phase. Data preprocessing stage is the most challenging and time-consuming stage [13]. It involves several necessary steps, including data cleaning, transformation, and feature engineering, which are crucial for ensuring the data's quality and relevance to the modeling phase. Its purpose is to refine, reshape, and condition raw data to render it suitable for training a model. This process holds immense importance because the quality of the data used for training directly impacts the accuracy and performance of the model. The data preprocessing phase encompasses a range of techniques, including addressing missing values, eliminating outliers, scaling or normalizing data, and choosing or deriving features. These methods ensure that the data employed for training is both precise and complete, aligning seamlessly with the intended modeling task. Effective data preprocessing plays a pivotal role in striving for optimal results during the modeling phase. By dedicating time and effort to data preprocessing, it becomes possible to elevate the quality and dependability of the model, ultimately leading to good results and deeper insights. In this study the commonly employed approach in data cleansing is used which involves the systematic cleaning of variables. In this method, inaccurate or misspelled data points within each feature are removed from the dataset based on defined criteria. These criteria involve conditions like verifying that the minimum and maximum values are within acceptable ranges, maintaining variance and standard deviation below a set limit, and ensuring the absence of inaccurately entered values in the dataset. Additionally, we eliminated attributes that contained a various number of missing values, such as description (with 454 missing value). By removing these attributes, the percentage of missing values in the dataset was reduced. We also considered the data transformation which refers to the act of changing data from one format to another. Properly validating and organizing data can enhance its quality and safeguard applications against problems such as null values, duplicate records, inaccurate indexing, and format incompatibilities.

4.3 Oversampling

The imbalance of data can negatively affect the performance of the proposed model, causing it to favor the majority class and underperform on the minority class during training. To address this, we opted for a data-level approach to balance the class distribution. We will use the Synthetic Minority Over-sampling Technique (SMOTE), a data augmentation method for tabular data, which creates synthetic samples for the minority class by duplicating existing samples without adding new information.

4.4 Modeling

Our system takes X as input and normalizes it to a subset of features denoted as $X = \{g_1, g_2, \ldots, g_n\}$, where each $\boldsymbol{g_i}$ falls within the range 0 to 1. Subsequently, it classifies customers into six categories based on RFM. For classification, we use Decision Trees and Naive Bayes classifiers. To improve the results, we employ a genetic algorithm for feature selection. Additionally, we use a deep neural network for classification. This network comprises five layers: an input layer, three hidden layers with 512 neurons each and fully connected Rectified Linear Units (ReLU), and a softmax output layer with six classes. We utilize batch normalization to speed up pattern recognition and dropout as a regularization method, applying l2-norm constraints on weight vectors over 30 epochs. Back-propagation is used to maximize the likelihood during training with cross-entropy loss.

4.5 Evaluation metrics

When evaluating the efficiency of classification models, accuracy is a common metric, but it can be misleading with imbalanced data. In such cases, accuracy may overestimate performance by favoring the majority class. To address this issue, we use alternative metrics like the ROC AUC score and the F-measure. The ROC AUC provides a deeper insight into class discrimination, while the F-measure balances precision and recall, offering a more comprehensive evaluation of the model's effectiveness.

Table 1: Evaluation metrics

Metric	Formula
Precision	$Pr = \frac{TP}{TP + FP}$

F-score	$Fs = \frac{(2 * Pr * Rec)}{(Rec + Pr)}$
Recall	$Rec = \frac{TP}{TP + FN}$

4.6 ROC-AUC

The performance of the model on the positive and negative classes of the test set has been measured using the ROC-AUC curve. A higher score indicates better performance. The ROC-AUC metric is preferred over accuracy, precision, and recall as they are not robust to changes in class distribution. The ROC-AUC metric is a widely used ranking evaluation technique, also known as ROC or the global classifier performance metric, as it can measure different classification schemes to compare overall performance. The other metrics might not perform well if the test set changes its distribution of positive and negative instances. However, the ROC curve is insensitive to changes in the proportion of positive and negative instances and class distribution.

5 Results, Analysis and Discussions

For our case study, we utilized data from a Moroccan omni-channel business located in Rabat, which has been in operation since 2018. The dataset comprises 9841 records with 15 features, collected over a four-month period. This business employs two methods for product distribution: through an e-commerce website and from its local store. Our study primarily focused on the data sourced from their e-commerce website. Because the e-commerce database was initially designed to cater to the company's specific needs, we employed various techniques to adapt the data and extract features in line with our research objectives. The dataset collected includes information about customer habits, consumption patterns, and demands as well as some information about RFM value. The customers are classified into six categories. For the purpose of machine understanding, we represent these six classes using the values 0, 1, 3, 4, 5 and 6. After gathering the data, the subsequent stage involves data preparation. Initially, the data is scrutinized for any missing values, revealing 1236 missing values in the description of product column. Consequently, this column is eliminated from the dataset before advancing to additional evaluations. Additionally, the customer ID column is excluded as it does not hold any predictive significance for customer behavior. Also the attributes that unnecessary are deleted like customer name and customer phone number columns.

Firstly, the performance of two well-known classifiers, Naïve Bayes and Decision Tree, was compared in terms of precision, F1-score, and recall. Table 2 illustrates the comparison results between the two classifiers. The Decision Tree demonstrated better performance compared to Naïve Bayes.

Next, the performance of the same models was measured in another scenario by applying a feature selection method involving a heuristic search method for solving optimization problems, specifically the Genetic Algorithm, to assess any modifications in the results. Additionally, we compared these two models with the proposed deep learning model in terms of four evaluation metrics.

After analyzing the performance of the two models with and without feature selection, it became evident that their performance excelled when feature selection was applied. As illustrated in Table 3 and Fig. 2, our proposed model exhibited the best results. The precision values for all models ranged from 75.6% to 87.3%, with our proposed model achieving the highest precision score of 87.3%. It outperformed the other models, demonstrating a precision of 87.3%, a recall of 89.5%, and an F1-score of 90.1%. These results strongly indicate its superior performance.

To assess the models' ability to distinguish between positive and negative classes, we used the ROC-AUC metric, which is widely regarded for its thorough evaluation of model performance. A higher ROC-AUC score reflects better overall performance in differentiating between positive and negative classes, underscoring the model's effectiveness in discrimination. Fig. 2 illustrates the ROC-AUC curves for the three models examined in this study. The results of the comparison show that our proposed model attained the highest score of 93.1%, showcasing its superior performance. Following this was the Decision Tree with Genetic Algorithm, scoring 87.7%, and Naïve Bayes with Genetic Algorithm, scoring 80.1%.

Table 2: performance of models

Model	Precision (%)	F1-score (%)	Recall (%)
Naïve bayes	74.3	72.1	73.1
Decision tree	80.3	76.8	78.6

Table 3: performance of models

Model	Precision (%)	F1-score (%)	Recall (%)	ROC-AUC score (%)
Deep neural netwoek	87.3	89.5	90.1	93.1
Naïve bayes + genetic algorithm	75.6	83.2	78.9	80.1
Decision tree + genetic algorithm	81.2	86.4	92.8	87.7

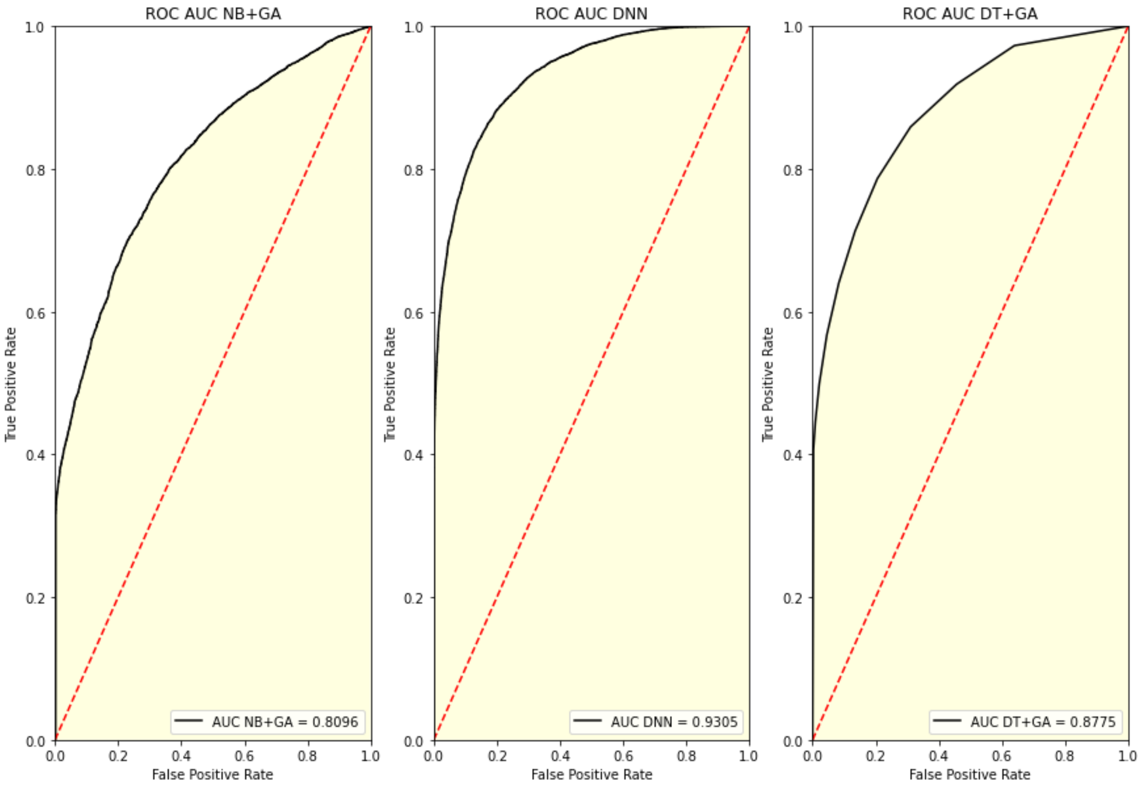

Fig. 2: ROC-AUC of models

6 Conclusion

Precision marketing equips marketers with insights into customer behaviors, enabling them to deliver the right message to the right person at the right time through the appropriate channel. Identifying the target customer is a crucial aspect of precision marketing. In this study, we proposed a response system that leverages an advanced feature engineering method, the RFM model, and a deep learning model, showing significant potential for enhancing managerial insights. Real data, collected and labeled using sophisticated techniques, were obtained from a Moroccan omni-channel business in Rabat. The data underwent preprocessing, cleaning, and balancing, with key features selected through advanced methods. To demonstrate the effectiveness of our approach, we compared the performance of our deep neural network model with decision trees and Naive Bayes, both with and without genetic algorithms. The experimental results indicated that the deep neural network model outperformed the others.

To further enhance this research, the application of big data analysis is suggested. Utilizing social network analysis to assess customer preferences in the context of retail services is recommended. Insights gained from this analysis can contribute to reducing customer churn and attracting potential clients. Additionally, incorporating diverse datasets could bolster the credibility of the results.

References

[1] El Koufi, N., & Belangour, A. (2024). Toward a decision-making system based on artificial intelligence for precision marketing: A case study of Morocco. Journal of Open Innovation: Technology, Market, and Complexity, 10(1), 1–10.

[2] Fan, L. (2023). Research on Precision Marketing Strategy of Commercial Consumer Products Based on Big Data Mining of Customer Consumption. Journal of the Institution of Engineers (India): Series C, 104(1), 163-168.

[3] El Koufi, N., & Belangour, A. (2024). Research Intelligent Precision Marketing Of Insurance Based On Explainable Machine Learning: A Case Study Of An Insurance Company. Journal of Theoretical and Applied Information Technology, 102(6), 1–10.

[4] Sanjaya, F. I., & Kalifia, A. D. (2023). Precision Marketing Model using Decision Tree on e-commerce Case Study Orebae.com. Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 7(5), 1033-1039.

[5] Chen, J. (2023, June). Construction of e-commerce platform precision marketing classification model based on big data algorithm. In Sixth International Conference on Intelligent Computing, Communication, and Devices (ICCD) Vol. 12703, pp. 565-571.

[6] Zheng, Y. (2020). Decision tree algorithm for precision marketing via network channel. Computer systems science and engineering, 35(4), 293-298.

[7] Wang, S., & Yang, Y. (2021). M-GAN-XGBOOST model for sales prediction and precision marketing strategy making of each product in online stores. Data Technologies and Applications, 55(5), 749-770.

[8] El Koufi, N., Belangour, A., & Sadiq, M. (2022). Research on precision marketing based on big data analysis and machine learning: case study of morocco. International Journal of Advanced Computer Science and Applications. 13(10), 58–63.

[9] Cheng, J. (2020). An Evaluation Strategy for Commercial Precision Marketing Based on Artificial Neural Network. Revue d'Intelligence Artificielle, 34(5).

[10] Tang, X., Cheng, C., & Xu, L. (2021). Research and Application of Precision Marketing Algorithms for ETC Credit Card Based on Telecom Big Data. The 7th International Conference on Signal and Information Processing, Networking and Computers (ICSINC), Springer, 1–6.

[11] El Koufi, N., & Belangour, A. (2023). Artificial intelligence techniques applied in precision marketing: a survey. In The International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), IEEE. 3(456), 1–5.

[12] El Koufi, N., & Belangour, A. (2023). Artificial Intelligence Techniques in Precision Marketing: A Multi-criteria Analysis and Comparative Study. In International Conference on Computing, Intelligence and Data Analytics? Springer. 1–7.

[13] El Koufi, N., Belangour, A., & Sadiq, M. (2023). Machine learning application in precision marketing: a systematic literature review and comparative study. In The International Conference on Artificial Intelligence and Smart Environment, Springer. 635(60), 601–607.

Notes on contributors

Nouhaila El Koufi is PhD student at the Hassan II University in Casablanca, Morocco. She is member of the laboratory: Information Technology and Modeling. Her main research interests include Artificial Intelligence, Big data analysis, Social network analysis, and Recommendation system.

Abdessamad Belangour is a full professor at the Department of Mathematics and computer science, Hassan II University in Casablanca, Morocco. His main teaching and research interests include Artificial Intelligence, model-driven engineering approaches and their applications on new emerging technologies such as big data, business intelligence, cloud computing, internet of things, real-time embedded systems, etc.