

IRSTI 28.23.01

S. Daulylov¹, A. Bogdanchikov²

^{1,2}Suleyman Demirel University, Kaskelen, Kazakhstan

INTRODUCE PREDICTIVE ANALYTICS USING THE NEXT BEST ACTION (NBA) MODELS INTO THE BANKING SYSTEM

Abstract. NBA - is an approach in which each client is initially considered purely individual. It has a close correlation with Predictive analysis. Predictive or prognostic analytics is a set of techniques and methods for analyzing data to build a forecast of future events. The banking system is currently using the method to obtain certain business results from its customers and has increased loyalty, increased income, found new growth points, etc. The classical model of marketing was rather different, it repelled from its existing product line and its parameters. But new models repels from customer's inclination to purchase a particular product. The aim of this project is to investigate the field of deposit accounts of the banking system by using NBA approach and to show the benefits and possible outcomes. This approach was tested on various aspects of the banking system and showed a number of solutions which can predict the probability of a customer to create a term deposit account.

Keywords: Next Best Action, Predictive analysis, implementation, Banking System, Term Deposit, Predictive analysis.

Аңдатпа. NBA - бұл әр клиент бастапқыда жеке тұлға болып саналатын тәсіл. Болжамдық талдаумен тығыз байланыс бар. Болжалды немесе болжамдық аналитика - бұл болашақ оқиғалардың болжамын құру үшін мәліметтерді талдау әдістері мен әдістерінің жиынтығы. Қазіргі уақытта банк жүйесі өз клиенттерінен белгілі бір бизнес нәтижелерін алу әдісін қолданады және адалдықты арттырды, кірісті ұлғайтты, жаңа өсу нүктелерін тапты және т.с. маркетингтің классикалық моделі мүлдем өзгеше болды, ол өзінің қолданыстағы өнімі мен параметрлерінен өзгеше болды. Бірақ жаңа модельдер тұтынушының белгілі бір өнімді сатып алуға деген бейімділігінен арылтады. Бұл жобаның мақсаты - NBA тәсілін қолдана отырып, банктік жүйенің депозиттік шоттары саласын зерттеу және оның артықшылықтары мен мүмкін нәтижелерін көрсету. Бұл тәсіл банктік жүйенің әртүрлі аспектілері бойынша сыналды және клиенттің мерзімді депозиттік шотты құру ықтималдығын болжайтын бірқатар шешімдерді көрсетті.

Түйін сөздер: Next Best Action, болжамды талдау, енгізу, банк жүйесі, мерзімді салым, болжамды талдау.

Аннотация. NBA - это подход, при котором каждый клиент изначально считается сугубо индивидуальным. Это имеет тесную связь с прогнозным анализом. Прогнозирующая или прогностическая аналитика - это набор методов и методов анализа данных для составления прогноза будущих событий. Банковская система в настоящее время использует метод для получения определенных бизнес-результатов от своих клиентов и имеет повышенную лояльность, увеличенный доход, найденные новые точки роста и т.д. Классическая модель маркетинга была несколько иной, она отталкивалась от существующей линейки продуктов и ее параметров. Но новые модели отталкиваются от склонности покупателя покупать конкретный товар. Целью данного проекта является исследование области депозитных счетов банковской системы с использованием подхода NBA и демонстрация преимуществ и возможных результатов. Этот подход был опробован в различных аспектах банковской системы и показал ряд решений, которые могут предсказать вероятность того, что клиент создаст срочный депозитный счет.

Ключевые слова: Next Best Action, прогнозный анализ, внедрение, банковская система, срочный депозит, прогнозные анализы.

Introduction

The theme of emotion and customer experience has been worrying bankers for quite some time. The tightening of regulations and requirements prompted banks to look for new sources of competitive advantage, in the US for example, 92% of millennials claim to have a lack of confidence in the traditional banking system and increasingly use new services. Customer experience is an indirect characteristic that translates into financial results of a bank through customer loyalty, a desire to continue to serve all of them at the same bank even in conditions of high consumer awareness and low switching costs to other financial companies. 78% of bank managers around the world surveyed by the IBM Institute for Business Value argue that customer engagement and understanding customer's needs are key conditions for the best customer experience and successful creation of new products and services.

Aim and objectives of the research

As a successful, the whole process of the research, find the dataset which will contain attributes of Bank System. Then preprocessing data, which means

replace empty fields or putting default value. A good result will be comparison several prediction methods. Also it will be nice if each step can contain a pretty graph. In the future go to practice with a big amount of data and using a stack of technologies from Big Data tools.

Background of literature review

At the beginning of the research, have been found a couple of interesting scientific papers which are a little be related to our topic, also some solutions that are ready for use. There are several finished offers, from free to paid.

One of the companies in New Zealand used a lot of data for improving credit scoring models. They construct a baseline model based solely on the existing scoring features obtained from the loan application form, and a second baseline model based solely on the new bank statement-derived features. A combined feature model is then created by augmenting the application form features with the new bank statement derived features [1].

Microsoft offers Veripark's [2] Next Best Action (N.B.A.) is a leading customer-centric technology that considers all the possible actions during a customer interaction and recommends the next best one thereby increasing the likelihood of positive response.

RapidMiner [3] is a software platform for analytics teams that unites data prep, machine learning, and predictive model deployment, also one of the solutions is Next Best Action. Their slogan is "One platform, does everything", they mean visual workflow, team collaboration, model management, deployment and work with technologies like Hadoop and Spark.

Cloud Sense [4] platform has a good interface, tutorials and before using it, you can see the demo version. For an organization that uses Salesforce, a most popular CRM system, CloudSense easily can be integrated.

Pega [5] they're a software solution called Pegasystems. Good company with a couple of success cases. In the platform, they have Customer Decision Hub, Intelligent Guided Selling, Contextual Next-Best-Action Marketing Intelligent Guidance for Customer Service.

NGData [6] offers the solution of Next Best Action as a whole marketing system. This decision is an excellent solution for the organization that works with clients and they need a CRM System.

Jacada [7] prefer a flexible system for working with their platform. Also, they have a nice intelligent chatbot. Their Agent Desktop is prime real-estate in the bid to win and keep customers.

Also, early in this year, there was a good article [8], about how Sberbank works with BigData and what kind of tools they are using. Their tools and methods can be as good for a start.

Methods and materials

The main tool to support this project was the Jupyter notebook, which consists of various libraries for applying of the machine learning approach. Data was collected from the open-sourced repository “UCI: Machine Learning Repository” [9]. In this database were used 45211 instances and 21 attributes.

- I. Attribute Information - to some of them was presented detailed graphs for better representation (Figure 1).

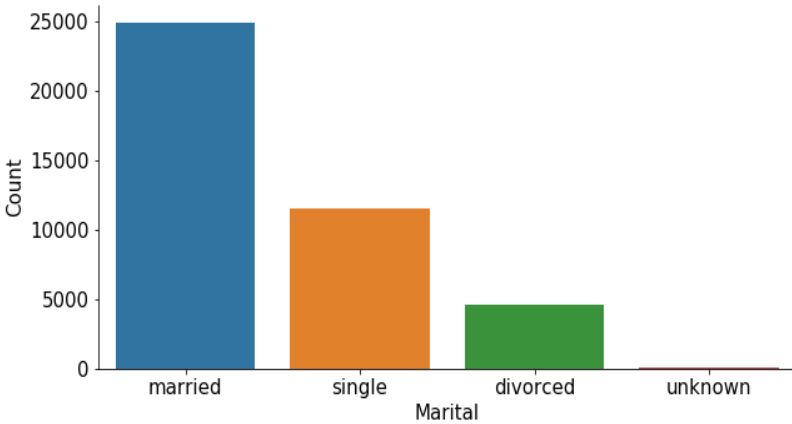
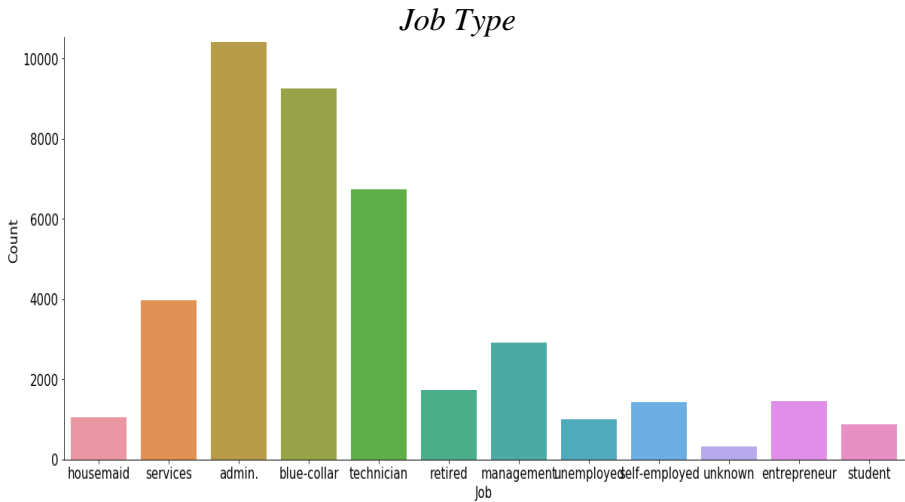


Figure 1. Attributes Information Marital Status



II. Implementation

Stage 1: Import of Library and Data

Stage 2: Preprocessing of data

Stage 3: Development of Predictive Model

After preprocessing data development of the model can finally start. Data was divided into standard proportions: train - 0.8, test - 0.2. Here will be comparing several types of machine learning classifiers:

1. K Nearest Neighbors Classifier [10]

An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If $k = 1$, then the object is simply assigned to the class of that single nearest neighbor. $K =$ will be choosing just brute-forcing neighbors from 1 to 25, and getting categories where the accuracy of prediction was tested. Thereby was selected an optimal number of neighbors.

2. Random Forest Classifier [11]

Random Forest Classifier is an ensemble algorithm. Ensembled algorithms are those which combine more than one algorithms of the same or different kind for classifying objects. Random forest classifier creates a set of decision trees from a randomly selected subset of the training set. It then aggregates the votes from different decision trees to decide the final class of the test object. Basic parameters to Random Forest Classifier can be a total number of trees to be generated and decision tree related parameters like minimum split criteria etc.

3. Naive Bayes Classification [12]

Naive Bayes classifier calculates the probabilities for every factor (here in case of email example would be Alice and Bob for given input feature). Then it selects the outcome with the highest probability.

Data and results

In this study were tested three types of classifiers. Random Forest classifier showed 90.0% of accuracy, NaiveBayes has shown 85.0% and highest among them is K Neighbors Classifier with remarkable 90.4% of accuracy. During the study 22 neighbors were selected due to their high accuracy. In comparison to more traditional methods of product-based marketing, this model allows finding potential customers with high accuracy due to a predictive selection of them (Figure 2).

The optimal number of neighbors is 22 with 90.4%

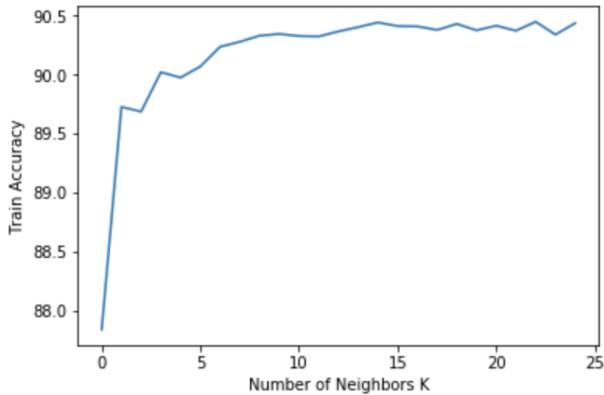


Figure 2. Selection of best numbers of neighbors

Discussion

In this paper, the NBA approach was presented and implemented into the banking system cases. The proposed approach correlates between Marketing and Data Science, to develop new models of working with clients. The performance of the NBA method in comparison to more traditional methods has shown remarkably well. In the future this might be used as the main model of working with clients. This is significantly increasing the popularization of Data Science among young developers and creates better ground for further investigation.

Conclusion

In this paper, the NBA approach was presented and implemented into the banking system cases. The proposed approach correlates between Marketing and Data Science, to develop new models of working with clients. The performance of the NBA method in comparison to more traditional methods has shown remarkably well. In the future this might be used as the main model of working with clients. This is significantly increasing the popularization of Data Science among young developers and creates better ground for further investigation.

References

- 1 Rory, P.B., Zhang, W., Asif Naeem, M. Improving a Credit Scoring Model by Incorporating Bank Statement Derived Features. Oct 30, 2016. Arxiv.org. Last modified on March 1, 2017. URL: <https://arxiv.org/abs/1611.00252v2>
- 2 Veripark. Veripark.com. Last accessed on Feb 27, 2020. URL: <https://www.veripark.com/products/veriloan>.

- 3 RapidMiner. Rapidminer.com. Last accessed on Feb 27, 2020. URL: <https://rapidminer.com/solutions/next-best-action/>.
- 4 CloudSense. Cloudsense.com. Last accessed on Feb 27, 2020. URL: <https://www.cloudsense.com/platform/genius-next-best-action/>.
- 5 Pega. Pega.com. Last accessed on Feb 27, 2020. URL: <https://www.pegacom.com/technology/next-best-action>.
- 6 NGData. Ngdata.com. Last accessed on Feb 27, 2020. URL: <https://www.ngdata.com/solutions>.
- 7 Jacada. Jacada.com. Last accessed Feb 27, 2020. - URL: <https://www.jacada.com/solutions/agent-next-best-action>.
- 8 Чирухин О. Спецпроекты в Сбербанк-Технологиях. Habr.com. Jan 30, 2018. Last accessed on Feb 27, 2020. URL: <https://habr.com/company/jugru/blog/347854/>
- 9 Moro, S., Cortez, P., Rita, P. A Data-Driven Approach to Predict the Success of Bank Telemarketing. Decision Support Systems, Elsevier. Archive.ics.uci.edu. June 2014. Last accessed on Feb 27, 2020. URL: <https://archive.ics.uci.edu/ml/datasets/Bank+Marketing>.
- 10 Harrison, O. Machine Learning Basics with the K-Nearest Neighbors Algorithm. Towardsdatascience.com. Sep 11, 2018. Last accessed on Feb 27, 2020. URL: <https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761>.
- 11 Yiu, T. Understanding Random Forest. Towardsdatascience.com. June 12, 2019. Last accessed on Feb 27, 2020. URL: <https://towardsdatascience.com/understanding-random-forest-58381e0602d2>.
- 12 Gandhi, R. Naive Bayes Classifier. Towardsdatascience.com. May 5, 2018. Last accessed on Feb 27, 2020. URL: <https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c>.