

Multidisciplinary
SCIENTIFIC JOURNAL
OF MARITIME RESEARCH



University of Rijeka
FACULTY OF MARITIME STUDIES

Multidisciplinarni
znanstveni časopis
POMORSTVO

<https://doi.org/10.31217/p.35.1.11>

Predictive analytics as a tool of controlling in decision making process in the marina industry

Uwe Lebefromm

University of Rijeka, Faculty of Economics and Business, Ivana Filipovića 4, 51000 Rijeka, Croatia, e-mail: uwe.lebefromm@sap.com

ABSTRACT

This paper is dealing with predictive modeling based on predictive analytics using computer application system and the usage of the prediction results for decision-making processes. Usually the prediction is based on the experience of decision makers, but the aim of this study is to explain and proof higher predictive efficiency when using predictive analytics based on machine learning as well as more accurate future-oriented business decisions. The marina industry in Croatia is used for this research because of its complexity and necessity to predict future events that influence company success with reliable accuracy. The information for decision-making were obtained from the customer database recorded manually over the past 30 years and according to data from December 2020. The optimized prediction by the vector machine and statistical theory based on the Bayes theorem is used to support more accurate prediction. The quantitative research was carried out using the SAP Predictive Analytics (SAP PA) computer application. The results of prediction models are a perfect basis for making future-oriented strategic and tactical decisions. This research proves that, with knowledge obtained from the results of prediction models it is possible to improve the identification of the target group among applicants and customers that contribute to company success. The research provides a theoretical and an empirical contribution in the usage of predictive analytics in the marina industry in Croatia.

ARTICLE INFO

Review article
Received 12 April 2021
Accepted 13 May 2021

Key words:
Business decision
Predictive analytics
Decision making
Controlling
Marina industry

1 Introduction

The development of information technology leads to the extensive digitalization of data and the use of digital data technologies in administration, management and controlling of the enterprises all over the world. This does not apply to routine work, but also tactical and strategic planning and the resulting management decisions supported by controlling. The tasks of controlling have changed significantly from the historical collection and evaluation of data to a decision support system. The possibility to use big data, the Internet of things, machine learning and other potential tools of the data science gives controllers a role of predictors. [37], [54, 27-48], [19]. However, this also requires controllers to ensure the consistency and comparability of diverse data and analysis models within the company [48, 8-19], [49, 160], [50]. Artificial intelligence systems and experts who supply these systems

with decision-relevant data, parameterize them correctly based on their own experience. The interpretation of predicted results by specialists using their knowledge leads to the development of effective and efficient tactical and strategic plans and decision proposals [15, 59], [22, 94-95]. Modern controlling must therefore break new ground. Experience and expertise must be supplemented by the usage of expert systems [23, 362-366], [44, 43]. This study follows this path. Pattern recognition in the data and the related development of forecast models are the basis for a system of quick and sustainable decisions.

The figure above gives an overview of the progress of reporting in controlling. Traditionally, controlling management received standard reports. This flood of numbers in the reports was outlined using graphics to visualize the core of the statements. With further development of the reporting system the possibility was created to select ad

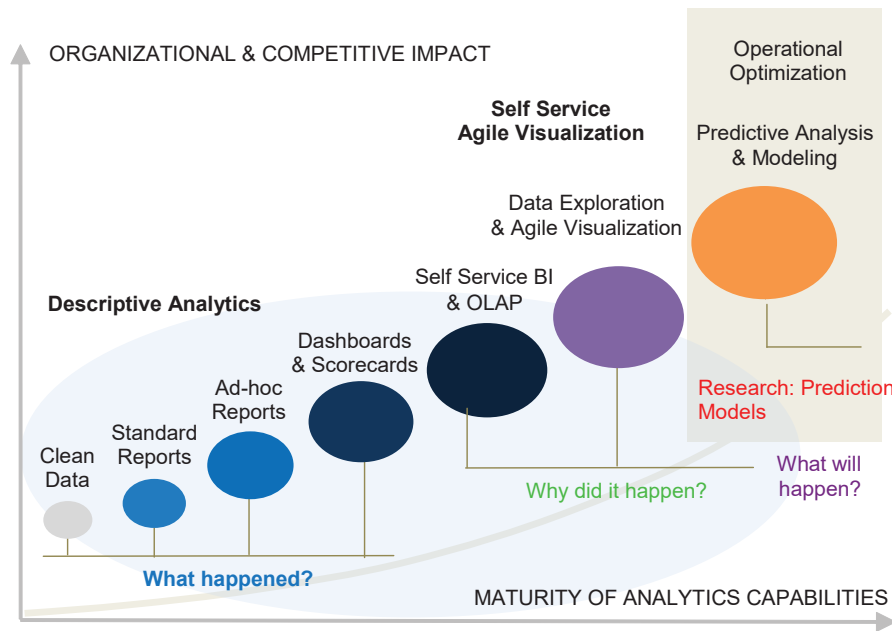


Figure 1 The progress of analytics

Source: Bakhshaliyeva [1], 2017, pp. 64. Modified by the author.

hoc figures and calculate key figures. Additional information systems have been created for better visualization and faster recording of key figures relevant to decision-making. This area also includes the Balanced Scorecard developed by Norton and Kaplan [29]. With the development of the systems based on a business warehouse system, a data pool was available in which data can be extracted from different systems. This enables interdisciplinary reporting. The entire know-how of a company was incorporated in a single reporting system, called Business Intelligence (BI). Given that, only the characteristics and key figures are extracted from the source systems, the name Online Analytical Processing (OLAP) was defined. The usability of the reporting systems was developed to such an extent that experts were no longer necessary for reporting. Management could serve itself. User-friendly data exploration and agile visualization with dynamic dashboards support the fast and up-to-date evaluation of business developments [25, 83]. However, in all cases, it is a question of reporting on historical data. A look into the future is essential if a company wants to remain competitive. This brings up the question of why predictive analytics is better. The availability of big data makes it possible to discover relationships between the data and be prove them [31], [46], [47, 863]. Another aspect is the availability of data from social media. With these data-as-a-service offers, companies can make their marketing campaigns more intelligent. Interrelationships between the company and market processes can be made visible and useful for business decisions. In this way, suitable target groups for the company can be better identified and marketing budgets can be used in a target group-oriented manner.

A company that does not use the new technology of data science and the associated possibilities for making future-oriented decisions will be overtaken by competition. This research has proven that predictive analytics does significantly improve the level of information of controlling and management [7, 35-37], [19], [20], [40, 63]. This lead to a changed decision-making behavior [2, 136], [6, 432-459], [9, 53-148], [13, 43-47], [43]. It would also has an impact to the processes in controlling and controlling decisions [18, 21], [36], [42], [51], [52], [27, 57], [42], [50], [38], [39].

The subject of this research is the marina industry in Croatia. Marina industry is a dynamic and complex industry, a very important competitive product of Croatia which tend to be sustainable [21], [26]. Considering the length of the coast nautical tourism is very competitive branch in Croatia with high turnover in the past few years [21], [26, 59-66], [28, 175-186], [33], [41], [45]. However, it is necessary to improve management of marinas in the direction of more affordable pricing of berths, raising the level of service in the marinas and making other improvements necessary for a better positioning in the Mediterranean. Several marina companies are competing. Therefore, offers must be made available to the interested parties in a short time. Long-term decision-making processes about the allocation of berths lead to applicants turning to competitors. The decision as to which prospective customer should be offered a berth depends on how reliable the prospective customer is and what agreement duration and agreement volume can be expected. Two tasks had to be mastered to achieve this goal. The first task was to understand the data available. With an understanding of the

data, the parameters for generating prediction models in the computer application could be set in such a way that the explanatory variables with the highest significance could be found. The second task was to create the prediction models and evaluate the results. The evaluation of the prediction models referred to performance indicators measuring the prediction confidence and prediction power of a model. These indicators are explained in the following chapter.

2 Theoretical background

The explanations of the theoretical background firstly refer to the statistical foundations of machine learning, developed as a statistical learning theory by Vladimir N. Vapnik. In the second step, the statistical learning theory was implemented in a programmed algorithm, the machine learning of the support vector machine. The third aspect is the transfer of the statistical learning theory to the machine learning algorithm which is the support vector machine. Machine learning is therefore used for pattern recognition in big data. Pattern recognition is an important theoretical subject for recognizing hidden information in the data. It combines operational research and machine learning. Operational research simulates and optimizes business processes using algorithms and quantitative methods. The machine learning uses the so-called learning algorithms to generalize recognized patterns in the data. The generalization leads to the fact that the recognized patterns can also be transferred to future data. Therefore, reliable statements can be made about future events when future data occurs [1, 22]. The relationships between the observed characteristics contained in the data can be very complex. This high complexity reduces the informative value of identified relationships since the associated boundary conditions are too special. Complexity reduction required in the statistical theory was developed by Vladimir M. Vapnik through his statistical learning theory. The basic idea is to replace the true risk in the forecasting through the empirical risk. If reality is described by a model that is too complex, the model cannot be applied to new future data because the model does not generalize enough. Lack of transferability of complex models was solved by Vladimir N. Vapnik, who found a way to reduce non-robustness and non-reliability of prediction models with the invention of the statistical learning theory [55], [56], [57, 207-249]. It leads to a compromise of data quality and robustness. The next step was the implementation of statistical learning theory in programmable learning algorithms. The algorithm used in this research in the computer application is the support vector machine. In the simplest of terms, pattern recognition is the separation of marked patterns in two classes. The support vector machine (SVM) is an algorithm with linear programming doing just that [10, 273-297], [11, 2-10], [17], [24], [30, 323], [32, 22]. Using the classification for pattern recognition, a set of Q classes that will consist of several classes q_i will

be used. The classification approximates the assignment of input variables x_i to discrete output variables y_j , which defines the class q_i . The variable x is then assigned to the class with the highest probability. The assignment of characteristics to a class is done by a classifier, which is an algorithm in the form of a mathematical-statistical function. The results of the prediction models are used to investigate decision-making behavior. Following the Bayes' theorem, a decision based on previous knowledge (a priori) has lower correctness than a decision which is based on additional decision-relevant information and therefore defined as a posteriori decision [3, 370-418], [4], [5, 80-83], [9, 53-148], [12, 260-274], [16, 376], [51], [52, 117-122], [53].

Subject of this research is the marina industry in Croatia. The scenario concerns management decisions and controllers support to these decisions regarding the conclusion of berth agreements with applicants. The objective of this research is the creation of predictive models about the behavior of applicants and customers of the marina industry in Croatia, providing a future-oriented data basis. The expected customer behavior relates to the conclusion of long-term berth rental agreements, and the parties credit rating. A distinction is made between corporate customers and private customers. These research objectives are also the target variables of the generated prediction models.

The database refers to data which include customer data of a marina company in Croatia. These data basically represent the explanatory variables for a prediction model, such as the age and citizenship of the applicants but also the length and age of their vessels. When the model is generated, the program runs the learning algorithm which is the Support Vector Machine. When the machine learning process was initiated, the computer application methods like clustering or classification were set. Characteristic values that are in a similar correlation to the target variable are grouped into clusters or classes. Classification is what is known as supervised learning since the classes are predefined. On the other hand, in the case of clustering, on the other hand, the very algorithm forms the clusters. This is called unsupervised learning. If the target variable is a continuous variable (e.g. sales volume), the computer application uses regression. Classification is used, if the target variable is a binary variable (e.g. extension of the contract YES or NO).

The research aims to calculate the probability that, under certain conditions, applicants and later customers of the marina company will extend their rental agreement for a berth. In generating prediction models, agreement extension is the central target variable of the model. The explanatory variables result from customer data, which were kindly made available by the marina company in Croatia. The explanatory variables result from the customer data, which was kindly made available by a marina company in Croatia. For the prediction of agreement extension, the main explanatory variables turned out to the

vessel length, vessel age, vessel type and certain marinas preferred by the customers.

The applied quantitative research was performed using the computer application called SAP Predictive Analytics (SAP PA). The support vector machine algorithm was programmed by the company KXEN into the KXEN algorithm. Using SAP PA, the prediction models were created with setting the parameters in such a way that best quality prediction results could be generated. The parameterization consisted of setting the target variable, excluding irrelevant variables that could contaminate the result. This was done through the successive evaluation of the key performance indicators KI and KR, which calculate the quality of the prediction model. The accuracy of the model indicates the influence of explanatory variables on the target variable. The performance of the predictive models can be measured by two key performance indicators [6], [7], [8].

1. Predictive power (KI) or predictive capability, which indicates the accuracy of the model. KI has a range of 0.0 to almost 1.0, usually 0.998. The higher the value of KI, the more significant the explanatory variables are. However, not all available explanatory variables of the database should be used. Only the essential variables that contribute to the target variable should be considered. This are the variables with the highest correlation to the target variable.
2. Prediction confidence (KR) or reliability of the prognosis characterizes the statistical robustness of the model. KR has a range of 0.0 to 1.0. Only a value above 0.95 shows that it is a robust model [8, 124]. A model is robust, when the same quality of prediction can be expected using new data with an unknown result [1, 31].

The abbreviations KI and KR come from the American company KXEN INC., which was founded in 1998. Therefore,

KI stands for “KXEN Information Contribution”. KR stands for “KXEN Robustness”. The results of the forecasts are visualized by graphics to the model.

The estimation curve graph results from the model calculation. The random curve graph is the result of randomly observed values. This means that, forecast-wise, it is coincidental whether an observed value was found or not. The probability of a “hit” is therefore 50% for the random curve. The definition of a forecast model should increase the probability of finding a “hit”, if possible, by more than 95%. The robustness of the forecast model is the applicability to statistical data, the result of which is unknown, which occurs through the KR calculation.

The performance indicators KI and KR are explained using the numerical example in the figure above. A forecasting model contains the parameters with which the target customers, i.e. those who are interested in the company’s products, should be found. The model parameters include personal characteristics of the person, for example, age, education, occupation, marital status, etc. The forecast model is calculating the influence of the parameters on the interest in their own products. Once the key parameters have been identified, product marketing can be targeted. In the following example, the total population is 18 million people. The target group in the population is about 6 million people, which is 33%. If the selection of members of the population to find the target group was random, only 50% of the target group would be found. If the prediction results were used to select members of the population, the hit rate would be significantly higher. If the perfect prediction model were to be generated, the whole target group of 33% of the total population would be found with the selection of 33% of the total population. From a business perspective, target customers

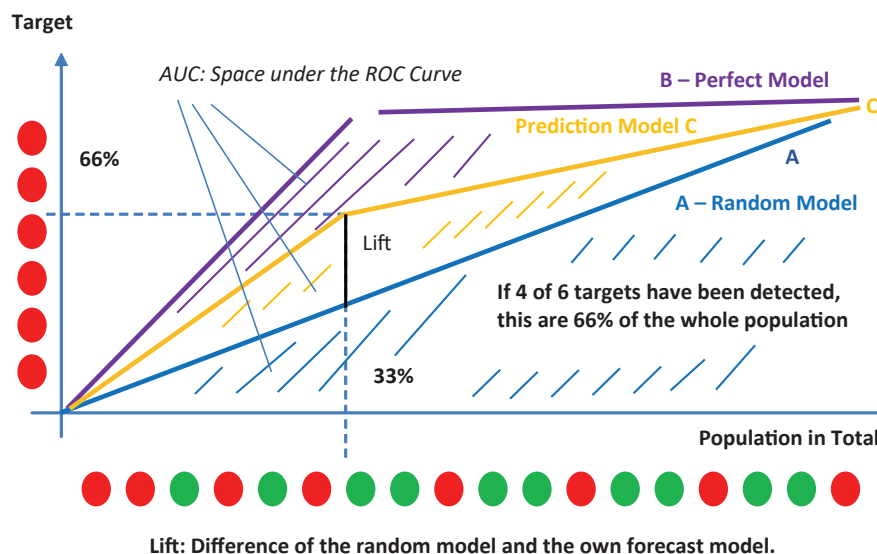


Figure 2 Prediction Model and Perfect Model

can be found faster, and marketing costs can be reduced. Revenues can be increased by using forecast models. KI is calculated as follows:

$$\text{Predictive Power: KI} = C/(A + B + C)$$

The areas $A + B + C$ shown in the previous figure represent the population. The ratio of the area of the forecasting model C and the perfect model B to the population $A + B + C$ is the measure of the predictive accuracy of the model and thus the result of model training (estimation). The ratio of the area of the forecasting model C to the population $A + B + C$ is a measure for the verification of the forecasting accuracy of the model (validation). As part of the validation process, it is intended to examine, regarding to the known statistical results, whether the predictions calculated by the model on the known statistical results can be confirmed. The robustness of the forecast model, is the applicability to statistical data whose result is not known, takes place via the calculation of KR. KR results as:

$$\text{Prediction Confidence (KR)} \approx 1 - B/(A + B + C)$$

The areas are calculated mathematically using the determined integral. The computer application uses the so-called rectangle methods as an approximation. For this purpose, the area of the function is divided into large number of rectangles whose areas are summed up. It results in a sum of rectangles, which are above the function and a sum of rectangles that lie below the function. The mean value is formed from both values. The curve therefore is named ROC – curve (Receiver Operating Characteristics) [34].

3 Methods

To be able to calculate a reliable forecast about the sales progress, it is of a decisive advantage for the competitive Croatian marina industry should be able to make reliable predictions about future customer behavior. The results of this research show that reliable statements about customer behavior can be made by using predictive analytics. The results show that from the large number of customer characteristics, those having a significant influence on the customers decisions. It is important to filter out less important features, because no decision-making constellations would be possible with an unmanageable number of influencing factors.

Partitioning method of clustering. The partitioning method of clustering divides a quantity of data into k clusters, whereby an observed characteristic belongs to exactly one cluster. The features observed are points in an n -dimensional Euclidean vector space. A cluster is represented by a centroid. The goal of the partitioning is to divide the data set into k partitions in such a way that the sum of the squared distances from the cluster centroids μ_i is minimal. Mathematically, this corresponds to the optimization of the function [14, 51]:

$$SSE = \sum_{x \in C} \text{dist}(x, \mu_i)^2$$

SSE – Sum of Squared Distance to Euclid Center.

C_i – cluster i .

x – observed variable value of the explanatory variables.

μ_i – cluster centroid

$\text{dist}(x, \mu_i)^2$ – Squared distance of observed variable x_j to the cluster centroid μ_i

The smaller the sum of the distances between the observed values and the centroid of a cluster, the more compact the cluster is. The variables used in this research are: x_1 = Vessel Type; x_2 = Vessel Age; x_3 = Vessel Length; x_4 = Marina. No Filter. With the use of the learning machine SAP Predictive Analytics, the contract extension was set as the target variable.

Method of Machine Learning. The process in generating the models in this research is divided into a learning phase (ESTIMATION) and a validation phase (VALIDATION). The system divided the data set randomly into two subsets:

- Estimation subset. In the first step, the variables which has not been excluded manually are tested about its significance to the target variable. If the significance is zero or very low, the algorithm will exclude the variables with the lowest significance as explanatory variables. A parameter can be used to set the degree of correlation at which an explanatory variable should be excluded. In the next iteration, a new model will be created using the remaining explanatory variables.
- Validation subset. In the next step, the created models are evaluated using the evaluation data set. The best model with the highest values for predictive power KI and prediction confidence KR will be selected for the third step.

The cutting strategy could be set manually. Charbert et al. [8, 163] propose to use the default cutting strategy. Generally, 75% of data records are assigned to the estimation data subset and 25% to the validation data subset. At model with minimum complexity is created first. The errors related to the prediction of the target variables with the actual values of the target variables known in the training data-set are calculated. After the first generated model, a second model is generated by including additional variables, which has the complexity of $h + 1$. The second model is validated. If the error rate decreases, a third model is generated with the complexity $h + 2$. As soon as the error rate increases with another generated model, the previous model is determined by the algorithm as the best predictive model. In the third step, the performance indicators are calculated for the best model using the tested data set [30], [31], [32].

Gaining knowledge from predictive models. In times of a global pandemic, which has been spreading worldwide since 2020, it has become more difficult to ac-

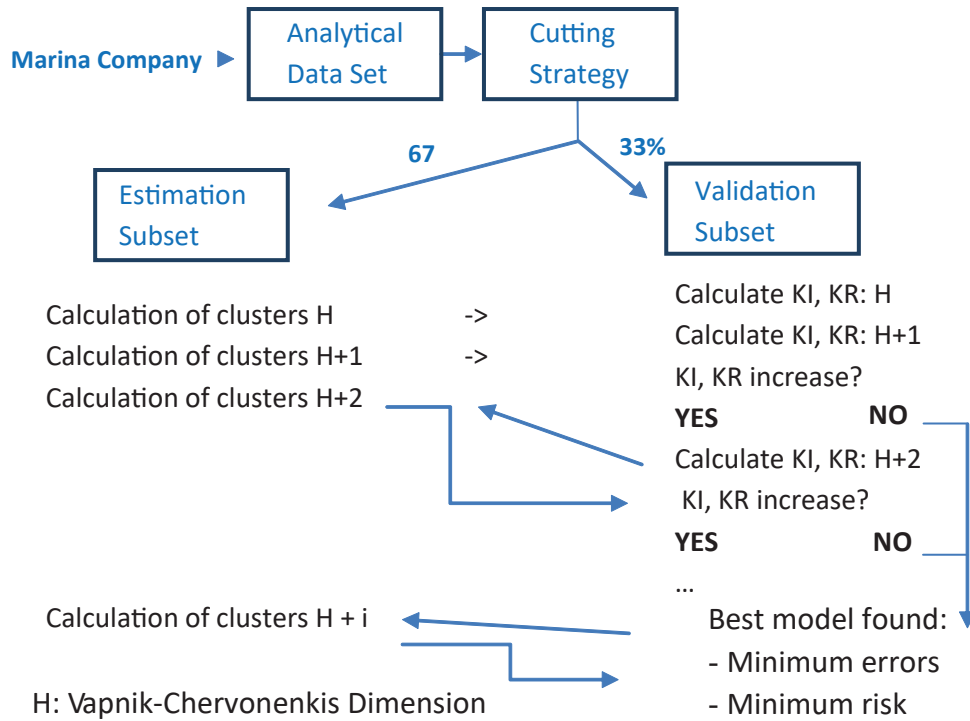


Figure 3 Machine Learning Process

Source: Author

quire new customers. In this context, and primarily with the goal to maintain long-term contractual relationships with customers, it is interesting to analyze the characteristics of private applicants and customers, who tend to extend existing agreements. With knowledge of the customer characteristics, marketing can be designed for this specific target group. The management gets a better basic for the calculation of the expected sales volume in the future if they know which of the existing agreements will be extended [35]. If, in the current difficult times of a pandemic, the focus is on extending the existing agreements, it is important to identify the customers who are most likely to extend their agreements. It is hypothesized that applicants and customers have similar characteristics with which the target group – customers with the intention of long-term contracts – can be identified. The investigation is focused on the explanatory variables for vessels. This occurs because the system has put customer groups PRIVATE and FIRM together in a single group and therefore does not have any influence on the target variable from this feature.

4 Research results

The significance of explanatory variables for the target variable “agreement extension”. The sailing boat vessel type is predominant in terms of long-term customer relationships. This could be explained by the fact, that not every marina is suitable for sailing boats. If customers did

find a marina with good wind conditions, they are more likely to hold onto the berth. The catamaran and motorboat vessel types have a negative impact on the extension of the existing agreements. Catamarans play a subordinate role regarding the proportion of vessel types. Motor yachts are less dependent on wind conditions than sailing boats. Regarding the size of the vessels, the result is the same as in previous investigations. Medium-sized vessels in the range of 9 to 15 meters are clear favorites when it comes to extending existing agreements. Boats of this length represent 77% of the total. This could also be an input for berth dimensioning. Owners of sailing boats, who prefer marinas suitable for sailing boats, could show that they are interested in free berths with a compatible berth length. Another aspect is the inclination towards charter boats. A study of the average length of chartered boats would be helpful here. In addition to owners of medium-length boats, owners of middle-aged boats are also the strongest group of clients who tend to extend existing agreements. Owners of newer boats are obviously more inclined to explore new areas and call at other ports accordingly.

It is interesting to examine the ports with an advantage in agreement extension. The preferred ports for agreement extensions are, for example, marinas with good wind conditions. It can be concluded that good wind conditions are of great importance for sailboats.

Contribution of the Variables. A clear correlation with sailing boats, suitable marinas and the average length and average age of the boats.

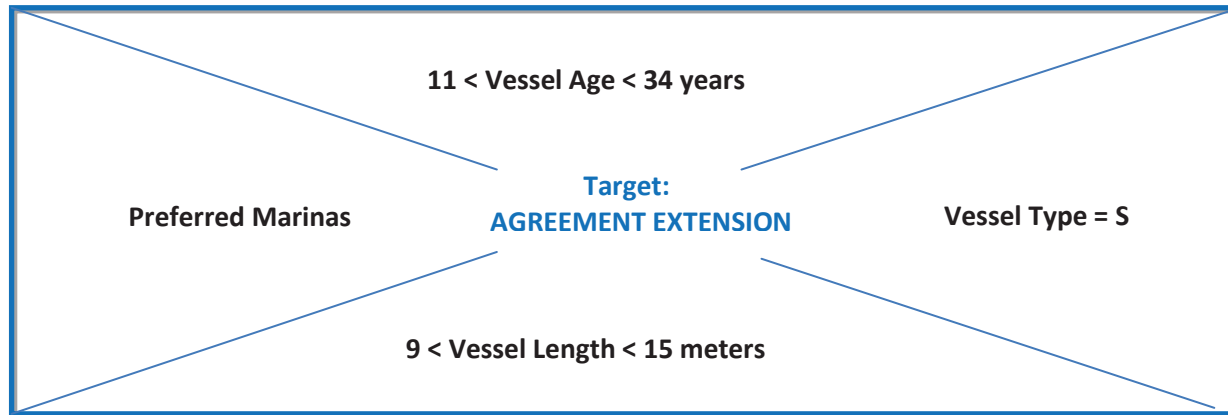


Figure 4 Explanatory variables with positive influence on target AGREEMENT EXTENSION

Source: Author

Explanatory variables: VESSEL AGE – VESSEL TYPE – VESSEL LENGTH – MARINA

Target variable: AGREEMENT EXTENSION. A marketing campaign aiming to extend the agreements should be addressed to clients who are using sailing boats with a length between 9 and 15 meters long and between 11 and 34 years old. The marinas which are preferred should be advertised. Note: The names of the marinas are not given exactly here to keep confidence. However, the properties shown should not discourage constant progress monitoring and result updating, while making available new data set and correcting them if necessary. Decisions and tendencies on the allocation of berths as well as the concept of marketing campaigns can be made according to the target group. Finally, controlling can provide a better basis for company planning of the expected sales volume.

5 Discussion

Predictive analytics is the innovative tool that helps controlling take future-oriented decisions. Reliable knowledge of customers decisions regarding the conclusion of agreements and agreement extensions provide a better basis for calculating future sales. Secured sales forecast and sales planning, in turn, provide a very good basis for calculating the development of financial opportunities and the associated investment opportunities. A company that can calculate its business development more precisely using innovative methods has a clear competitive advantage.

Outlook: The following projects could further develop these research results:

- With the knowledge of the prediction model, a guideline-based questionnaire could provide the a-posteriori better decisions with the knowledge of the prediction model [3], [4], [5], [9], [28].
- Continued analysis of further decision-making processes, for example in the planning of marketing campaigns, the search for suitable locations for new marinas or the question of a combination of individual

products such as a charter boat, berth, skipper and maintenance service for the ships.

- Surveying other marina companies, national and international. Such a survey coverage could further increase the applicability of stochastic decision models.
- Use of modern forecast information systems, for example SAP ANALYTICS CLOUD.

However, it should be emphasized that the concept of a decision logic in the form of stochastic decision models enables the models to be programmed in an information application system and thus used as a decision support system for controlling and management. This would make a further contribution to digitalization in controlling.

Funding: The research presented in the manuscript did not receive any external funding.

Acknowledgments: A big thank you to Prof. Dr. Neda Vitezić for the management of the organization to obtain the data necessary for the research. A big thank to the controllers of the marina company for the cooperation to make the data for prediction available.

References

- [1] Bakhshaliyeva, N., Chen, J. L., Dommer, U., Samlenski, E., Schmedt, H., Schulze, N., Wilczek, R.: SAP Predictive Analytics – Vorausschauende Analyse mit SAP, Rheinwerk, Bonn. 2017. pp. 22.
- [2] Bamberg, G., Coenenberg, A., Krapp, M.: Betriebswirtschaftliche Entscheidungslehre. 15th edition, 2012. München. pp. 136.
- [3] Bayes, Th., 1763: Bayes, Th.: *An Essay towards solving a Problem in the Doctrine of Chances*. In: *Philosophical Transactions*. Band 53, 1763, pp. 370–418. An essay solving a problem in the doctrine of chances. By the late rev. Mr. Bayes F.R.S. Communicated by Price, R. in a letter to John Canton A.M.F.R.S.
- [4] Bayes, Th., 2008: Versuch zur Lösung eines Problems der Wahrscheinlichkeitsrechnung. Hrsg.: H. Timerding. Leipzig. 1908.

- [5] Beck, H.: Behavioral Economics. Eine Einführung. Springer-Gabler, Wiesbaden. 2014. pp. 80–83.
- [6] Busemeyer, J. R. & Townsend, J. T.: Decision field theory: A dynamic cognitive approach to decision-making in an uncertain environment. *Psychological Review*, 100, pp. 432–459. 1993.
- [7] Butsmann, J., Crumbach, M., Franke, J., Köhler, B., Morgenthaler, J.: SAP S/4HANA Embedded Analytics – Architektur, Funktionen, Anwendungen. Rheinwerk, Bonn. 2019. pp. 35–37.
- [8] Charbert, A.; Forster, A.; Tessier, L.; Vezzosi, P.: SAP Predictive Analysis – The Comprehensive Guide. 2017. Boston, USA. pp. 163.
- [9] **Clinical Laboratory Management Review: Bayesian Analysis: An objective scientific approach to better decisions.** 1999. pp. 53–148. Published in Europe PMC, online science platform.
- [10] Cortes, C., Vapnik, V.: **Support Vector Networks.** Kluwer Academic Publishers. Boston (USA). In: *Machine Learning*, 20, 1995. pp. 273–297.
- [11] Dianzi, Keji, Daxue, Xuebao: **An Overview of Theory and Algorithm of Support Vector Machines.** In: *Journal of the University of Electronic Science and Technology of China*. Vol. 40/1, pp. 2–10. Year 2011.
- [12] Diederich, A.: **Dynamic Stochastic Models for Decision Making under Time Constraints.** In: *Journal of Mathematical Psychology* 41, pp. 260–274, 1997
- [13] Dörsam, P.: *Grundlagen der Entscheidungstheorie.* 6. Edition. Heidenau. 2013. pp. 43–47.
- [14] Ester, M., Sander, J.: *Knowledge Discovery in Databases. Techniken und Anwendungen.* Springer. Berlin. 2013. pp. 34, 111, 160.
- [15] Feindt, M. und Kerzel, U.: *Prognosen bewerten,* Springer Gabler, 2015, pp. 59.
- [16] Fenton, Norman; Neil, Martin: **Risk Assessment and Decision Analysis with Bayesian Networks.** School of Electronic Engineering and Computer Science. Queen Mary University of London. 2nd Edition, 2019. pp. 376.
- [17] Fischer, J.: **Support Vector Machines (SVM).** Seminar „Statistische Lerntheorie und ihre Anwendungen“. Universität Ulm. June 12th 2007.
- [18] Gänßlein, S.: **Die Digitalisierung ist eine Chance für den Controlling.** In: Horvath, P.: *Controlling, Zeitschrift für erfolgsorientierte Unternehmenssteuerung, Sonderausgabe* 2017. pp. 21.
- [19] Gleich, R.; Tschandl, M. (Hrsg.): **Digitalisierung und Controlling. Technologien, Instrumente, Praxisbeispiele.** Haufe. Freiburg, München, Stuttgart. 2018.
- [20] Gleich, R., Grönke, K., Kirschmann, M., Leyk, J.: **Controlling und Big Data. Anforderungen, Auswirkungen, Lösungen.** Haufe. Freiburg, München. 2014.
- [21] Gračan, D.; Gregorić, M.; Martinić, T.: **NAUTICAL TOURISM IN CROATIA: CURRENT SITUATION AND OUTLOOK; Conference Paper** 2016. pp. 1. In Research Gate, <https://www.researchgate.net/publication/325181470>.
- [22] Hölscher, Bert F.: **Digitales Dilemma – Unternehmen im Spannungsfeld zwischen Effizienz und Innovation.** Arkadia, Hamburg. 2017. pp. 94–95.
- [23] Horváth, P., Gleich, R., Seiter M.: *Controlling.* 2015. 13. Auflage. pp. 362–366.
- [24] Hurwitz, J., Kirsch, D.: **Machine Learning for Dummies. Understanding Machine learning fundamentals. Make sense of machine learning algorithms. Build your data science team.** Jon Wiley & Sons Corp. Hoboken (NE), 2018.
- [25] Hyndman Rob J., Athanasopoulos, G.: **Forecasting – Principles and Practice.** A comprehensive introduction to the latest Forecasting methods using R. Learn to improve your forecast accuracy using dozens of real data Examples. Publisher OTEX. Melbourne 2014. pp. 83.
- [26] Ivanić, K., Hadžić Perić, A., Mohorović, D., **Nautical Tourism: Generator of Croatian Economy Development.** *Scientific Journal of Maritime Research* 32, pp. 59–66, 2018.
- [27] Jeschke, B. G.: **Entscheidungsorientiertes Management. Einführung in eine konzeptionell fundierte pragmatische Entscheidungsfindung.** Berlin, Boston. 2017. pp. 57.
- [28] Jugović, A., Kovačić, M., Hadžić, A.: **Sustainable development model for nautical tourism ports,** *Tourism and Hospitality Management*, Vol. 17, No. 2, pp. 175–186, 2011.
- [29] Kaplan, R. S., Norton, D. P.: **Balanced Scorecard, Strategien erfolgreich umsetzen.** Schaeffer, Poeschel. 2018. Stuttgart.
- [30] Kelleher, J. D., Mac Namee, B., D’Arcy, A.: **Machine Learning for Predictive Data Analytics, Algorithms, Worked Examples and Case Studies.** Cambridge, London. 2015. pp. 323.
- [31] Kießwetter, M., Vahlkamp, D.: **Data Mining in SAP Netweaver BI.** Gallileo Press, Bonn, 2007.
- [32] Kovalczyk, A.: **Support Vector Machines Succinctly.** Publisher: Syncfusion. Morrisville (USA), 2017. pp. 22.
- [33] Kovačić, M.: **Criteria for selecting a location for a port of nautical tourism.** Research Gate. 2009.
- [34] Krzanowski, W. J.: Hand, D. J.: **ROC Curves for Continuous Data.** In: *Monographs on Statistics and Applied Probability* 111. CRC Press Taylor & Francis Group. London, New York. 2009.
- [35] Kucukvar, M. Noori, M., Egilmez, G., Tatri, O.: **Stochastic decision modeling for sustainable pavement designs.** Springer, Berlin, Heidelberg. 2014.
- [36] Kuik, Sw. S., Kaihara, T., Fuji, N.: **Stochastic Decision Model of the Remanufactured Product with Warranty.** *Proceedings of the International Multi-Conference of Engineers and Computer Scientists.* Vol II, IMECS, March 18–20, Hong Kong. 2015.
- [37] Langmann, Chr.: **Digitalisierung im Controlling.** Springer Verlag, Heidelberg. 2019.
- [38] Laux, H., Gellenkirsch, R.M., Schlenk-Mathes, H.: **Bildung eines Wahrscheinlichkeitsurteils und Bewertung von Informationen.** 9. Auflage. Springer Verlag, Heidelberg. 2014.
- [39] Laux, H., Gillenkirch, R. M., Schenk-Mathes, H.Y.: **Entscheidungstheorie.** 10th Edition. Frankfurt am Main. 2018.
- [40] Lawrence, K. D., Klimberg, R. K.: **Advances. Business and Management Forecasting.** Vol. 12. Emerald Publishing. United Kingdom, North America, Japan, India, Malaysia, China. 2018. pp. 63.
- [41] Luković, T.: **Nautical Tourism.** Dubronik, 2013.
- [42] Möller, K.: **Controlling-Prozessmodell 2.0. Leitfaden für die Beschreibung und Gestaltung von Controlling-Prozessen.** 2. Auflage. Haufe-Verlag. Freiburg – München – Stuttgart. 2017.
- [43] Morris, D.: **BAYES THEOREM. A visual introduction for beginners.** Blue Windmill Media. Printed by Amazon, Leipzig. 2017.

- [44] Nasca, D., Munck, J.-Chr., Gleich, R.: Controlling-Hauptprozesse: Einfluss der digitalen Transformation. Investigation April – June 2018. In: Gleich, R., et al., 2018. Digitalisierung und Controlling. pp. 73.
- [45] Peović, K.: Die politische und wirtschaftliche Situation Kroatiens – die Peripherie Europas heute. Blog. Published in: “transform europe“. Vienna. 2018.
- [46] Runkler, A.: Data Mining – Modelle und Algorithmen intelligenter Datenanalyse. 2nd Edition. Springer-Vieweg, Wiesbaden. 2015.
- [47] Russell, St., Norvig, P.: Künstliche Intelligenz. Ein moderner Ansatz. 3. Aufl. Pearson DE, München. 2012. pp. 863.
- [48] Schäffer, U.: **Levers of Organizational Resilience. In: Controlling & Management Review. Zeitschrift für Controlling & Management.** 64, Jahrgang. Ausgabe 6-7. 2020. pp. 8–19.
- [49] Schäffer, U., Weber, J.: **Digitalization will radically change controlling.** In Schäffer, U: Behavioral Controlling- Anniversary Volume in Honor of Jürgen Weber, Springer. 2019. pp. 160.
- [50] Steinke, K.-H.; Schmidt, W.: Auf dem Weg zum Controlling 4.0 – Leitfaden für Controlling im Wandel. Haufe-Lexware, Freiburg. 2017.
- [51] Stigler, St. M.: **Who discovered Bayes Theorem. The American Statistician.** 1983. Vol. 37. No. 4.
- [52] Stigler, St. M.: Richard Price, the first Bayesian. *Statistical Science*, 2018. Vol. 33. No. 1. pp. 117–122. *Institute of Mathematical Statistics*.
- [53] Subbalakshmi, G., Ramesh, K., Chinna Rao M.: **Decision Support in Heart Disease Prediction System using Naive Bayes.** *Indian Journal of Computer Science and Engineering (IJCSE).* Vol. 2. No. 2. 2011.
- [54] Tschandl, M., Kogleck, R.: **Controller als Innovatoren: Von der Digitalisierungs-Roadmap zum neuen Geschäftsmodell.** In: Gleich, R., et al.: *Digitalisierung und Controlling – Technologien – Instrumente – Praxisbeispiele.* Haufe. 2018. pp. 27–48.
- [55] Vapnik, V. N.: **Inductive principals of statistics and learning theory.** In: *Yearbook of the Academy of Science of the USSR on Recognition, Classification and Forecasting (in Russian).* Vol. 1, Nauka, Moscow. (English translation: V.N. Vapnik, *Inductive principles of statistics and learning theory.* In: *Mathematical Perspectives on Neural Networks*, Smolensky, Mozer and Rumelhart, eds., Lawrence, Erlbaum Associates. Springer New York. 1996.
- [56] Vapnik, V. N.: **Statistical Learning Theory.** A Volume in the *Wiley Series on Adaptive and Learning Systems for Signal Processing, Communications, and Control.* Simon Haykin, Series Editor: Wiley. 1998. Reprint 2018. London. New Delhi.
- [57] Vapnik, V. N., Chevronekis, A. Ya.: **The necessary and sufficient conditions for consistency of the method of empirical risk minimization.** *Yearbook of the Academy of Science of the USSR on Recognition, Classification and Forecasting.* Vol. 2, Nauka, Moscow, pp. 207–249 (in Russian). English translation: V. N. Vapnik and A. Ya. Chevronekis (1991). **The necessary and sufficient conditions for consistency of the method of empirical risk minimization.** *Pattern Recognition. Image Anal.*1(3), pp. 284–305.