

MACHINE LEARNING BASED PREDICTIONS OF SALES LEADS: PROOF OF QUALITY FROM POLISH BUSINESS-TO-BUSINESS COMPANY

Tomasz Woźniakowski  <https://orcid.org/0000-0002-0779-4769>

Institute of Economics and Finance
Warsaw University of Life Sciences – SGGW
e-mail: Tomasz_Wozniakowski@sggw.edu.pl

Abstract: Most sales managers struggle with achieving high lead conversion, key to lowering marketing costs and improving sales efficiency. Existing research emphasizes costly large-scale methods, often inaccessible to SMEs. Meanwhile, IT SMEs in B2B face numerous low-value leads without predictive support. This study proves that AI (AutoML on Google Cloud) can cost-effectively predict sales opportunities. Using 1000 historical leads, it demonstrates accurate predictions, offering SMEs a practical tool and paving the way for further research.

Keywords: sales leads, sales leads conversion, sales lead value, artificial intelligence, automated machine learning

JEL Classification: A11; B16; B21; C45; D22; M31

INTRODUCTION

The digital transformation in business is driven, among other factors, by a huge increase in the amount of data and computing power [Małkowska et al. 2021]. The problem of sales procedures optimization becomes very intense, competitive and complex process. Sales win-propensity prediction is fundamental to effective sales management [Yan et al. 2015].

Large actors often use very sophisticated and expensive methods including big data analysis often utilizing artificial intelligence. [Yan et al. 2015] This seems to be the reason why most of researchers are focused on large scale operations through the use of large datasets [Plawgo et al. 2021], therefore describe time and

<https://doi.org/10.22630/MIBE.2025.26.2.7>

cost consuming methods of sales optimization which results in high entry threshold, very often unreachable for small and medium enterprises (SMEs).

The SME sector plays the leading role in labor creation, sales, and production of value-added goods and services in most countries [Ključnikov et al. 2022]. SMEs significantly contribute to the growth GDP dynamics. Small and medium-sized enterprises create a more intense competitive environment which translates into prices and quality of products and services [Oleksiuk et al. 2022]. Moreover, the high-tech sector was affected by globalization at a much faster pace than other sectors [Soniewicki et al. 2022]. Therefore, an author will argue that SMEs must reduce this technology gap [Hyder et al. 2022].

The article emphasizes the real-life experiment of the Business To Business (B2B) market model in which small enterprises often compete with large ones, which appears to be very strong in IT sector [Mang'unyi et al. 2019]. The SMEs can offer a very narrow but at the same time very deep range of competencies, effectively gaining serious positions and even an advantage in selected niches. Therefore they have to provide very intense marketing activities which results in high numbers of sales leads (SL), in most cases without any kind of sophisticated value prediction support. To maintain a competitive advantage on the market, companies must efficiently adapt their operations [Nogalski et al. 2020].

The aim of the paper is to prove: there is a cost-effective method of using Artificial Intelligence (Machine Learning) for successful predictions of sales in SME operating in B2B model.

The study is focused on complex data taken from Polish business-to-business company in IT sector. Although the standard European Union classification divides SMEs into three categories regarding their size (micro, small and medium-sized enterprises) [Civelek et al. 2021], the researcher classified the company as small since it has less around 30 employees.

Researched sample spans last 5 years of its more than 30 years existence on the market. Data analysis are empowered by couple real-time experiments conducted in last 6 months of year 2022 when results of sales predictions done by models, AI and more traditional approach have been compared.

The paper examines the complexity of the sales activities: 1) financial value of sales leads, 2) success probabilities defined by status, 3) intensity of contacts, 4) geographical factors, 5) interpersonal factors. The research is finally focused on data from 1000 historical sales leads (SL). The data has been classified by proposed AI model. The automated machine learning (AutoML) binary classification model on Google Cloud Platform has been created for experiment, trained by prepared data and used for sales lead future value prediction and confronted to historical outcomes of conversion rates and values.

The rest of the paper is structured as follows: Literature review will outline this research's theoretical background. Research methodology section will develop the research hypotheses, the data, used models, and the paper's research methods. The results from the analyses are illustrated and explained in another section. Then

the paper will discuss this study's results and arguments for them, including some existing and possible implementations. Finally, the conclusion part of the research paper will summarize main points.

LITERATURE REVIEW

Several queries in scientific resources have been done in order to define current knowledge in areas similar to the topic of the paper. They addressed subjects of:

- machine learning in predictive sales pipeline or sales leads analytics,
- machine learning models in sales predictions,
- business analytics for sales pipeline management in the software industry,
- sales forecasting machine-learning techniques,
- B2B sales predictive modeling: machine-learning approach.

Contrary to the author's expectations, few work has been done or released for involving predictive modeling in sales pipeline analytics, presenting similar focus on sales predictions, sales leads in B2B model. However, some research articles have been found, covering similar areas to the paper and share similar attitude to major problem.

The paper by Yan et al. [2015] presents profile-specific, two-dimensional Hawkes processes model, developed to capture the influence from seller's activities on their sales leads to the win outcome, coupled with lead's personalized profiles.

Another work by Bohanec et al. [2017] presents state-of-the-art black-box prediction model which is a novel use of this methodology inside an intelligent system in a real-world case of business-to-business (B2B) sales forecasting.

In general, many of literature found in topic shares the same attitude to an author of this paper:

1. focus on sales leads as a process reflecting whole history and outcome of sales process [Kaplan et al. 1996],
2. focus on B2B model as most suitable for SLs management [Zalocco et al. 2009],
3. use of Machine Learning as most appropriate method for sales forecasting [Huo 2021]
4. use of experiment in real-world case [Rohaani et al. 2022].

However, the most important differences are:

1. in every case, presented solution is large scale solution for big enterprises,
2. none of researchers analyses financial value of SL as a function of time and probability to support the AI model predictions,
3. they are mostly focused on binary result of the whole process (successful / unsuccessful lead).

It needs to be underlined that the first from list above is the most important obstacle to consider proposed by mentioned authors methods in a scale of SME. Yan

et al. even state that their model is deployed and in continual use to a large, global, B2B multinational technology enterprise listed at “Fortune 500”. It took three years and work of eight researchers to establish, train and deploy model according to their findings. The cost of such endeavor is almost equal to the whole year income of the SME company that has been described in given paper. It is obviously impossible to make use of such solutions by SMEs.

Also the key components of given subject has been studied in literature separately in order to observe current findings and understanding of terms: sales leads, sales pipeline, pipeline funnel, artificial intelligence models, automated machine learning and others.

Sales lead can be described as an entity (business or a person) that is not yet a client, however, there are strong signs that they can be. What’s more, sales leads can also be in a form of data, which allows to identify those potential future consumers.

The ideas of a "consumer funnel" or "sales funnel" were created by Peterson [1959], which was first visual image and description of mentioned phenomenon. In order to achieve a result (sales), it is necessary to get results at each of the stages. Peterson [1959] identified four phases of the sale:

- The initial phase is to attract attention;
- The phase of adaptation of the (offer to the needs of the client) – to arouse interest;
- Illustration /explanation phase – to form a desire;
- The selection phase is to achieve action [Stoop 2009].

SL ‘travels’ in time changing its parameters (mostly probability of success as it goes through different stages towards deal) and the average conversion rate of SLs to sales deals is around 10% in the B2B sector [Eitle et al. 2019].

In recent years the topic of Artificial Intelligence (AI) has been widely spoken about, across all different fields. Machine Learning works as a connecting bridge between AI and Data science allowing to automatically draw conclusions from the raw data provided and test hypothesis.

According to D. Guiterrez, we can distinguish two forms of machine learning used in data science: supervised and unsupervised learning [Gutierrez 2015]. Automated machine learning, also called automated ML or Auto ML, is the process of automating time-consuming and repetitive tasks of developing machine learning models. Preprocessing, training and evaluation is an experimental and iterative process that requires several trials before satisfactory results are achieved. Since these tasks tend to be repetitive, AutoML can help automate these steps. In addition to automation, optimization methods are used in the learning and evaluation process to search for and select algorithms and hyperparameters [Shen et al. 2021].

Automated machine learning is the process of automating tasks related to applying machine learning to real-world problems. AutoML potentially covers every stage, from a raw dataset to building a ready-to-deploy machine learning model. It is the next step in the development of artificial intelligence. The work uses the

AutoML service on Google Cloud Platform. AutoML is using Combined Algorithm Selection and Hyperparameter Optimization. According to Feurer et al. the goal of CASH is to find a common algorithm and hyperparameter settings that minimize the training dataset given a set of algorithms and hyperparameters of those algorithms [Feurer et al. 2015]. The authors define AutoML as the problem of automatically (without human input) producing test set predictions for a new dataset within a fixed computational budget. Therefore, the arguments for AutoML use in case of this paper are:

- minimization of the training data set,
- developer must set limits on the resources used in the AutoML optimization process. This budget usually consists of one or a combination of CPU/GPU usage, uptime, and memory usage,
- model access via simple REST-API and JSON.

AutoML limitations:

- AutoML algorithms rely on data to be clean and relevant. Data cleansing and feature engineering are not yet supported by any AutoML approach,
- developing a well-functioning solution can take quite a long time and in extreme cases it may not be possible.

In given example the AutoML runs on Google Cloud Platform, which is a suite of cloud computing services that runs on the same infrastructure that Google uses internally in its end-user products, such as Google Search, Gmail, Google Drive and YouTube.

RESEARCH METHODOLOGY

The company, chosen as a real-life research object has been established in 1989 in Poland. It is an IT company that produces and implements its own software in area of finance and budgeting for other companies and organizations. It has been chosen by certain criteria:

- Operates in B2B model,
- Belongs to SME (small enterprises less than 50 employees, net revenues less than or equal to EUR 10 million),
- Has minimum thousand historical sales leads including full record of its variability and final outcome.

The material selected from the company is historical database of 1496 records containing information on initial state of SL and its result.

Table 1. Example of a record including SL data

Voivodship	Income Planned	Date Planned	Pipeline Funnel	Company Employee	Success
Kujawsko-Pomorskie	12500	2	P4	JK	false

Source: own work based on company's MySQL database with sales leads (SL) data

The data will be used to train model of automated machine learning using binary classification which predicts a binary outcome. In simplest words it is going to predict if new SL is most likely a success or not. After several surveys done among sales persons, several features has been selected. Features are how the model identifies patterns to make predictions. In general they need to be relevant to the problem. Another goal was to minimize its number to reduce the cost.

As a result a dataset has been constructed. It contains transaction details shown in Table1. In general, the more training examples you have, the better your outcome is. The amount of records also scales with the complexity of the problem. Usually for classification problems we need to have 50 x the number of features, but since Google Cloud requires minimum amount of 1000 records, an author decided to use this amount. It could leave the room for more features, but it has to be stressed that the general idea is to predict outcome at initial stage of SL where amount of information on features is limited. It has been decided that finally a record would contain:

- Voivodship - geographical region of a client,
- Income Planned – planned financial flow at the end of process,
- Date Planned - denoted in months prediction of process duration,
- Pipeline Funnel - success probability connected to stage at which the SL is (from P5 - lowest to 0 - certain)
- Company Employee – initials of sales person in charge of SL process,
- Product Name – product acronym from available class,
- How Many Products – how many products certain client has implemented in the past.

THE HYPOTHESES:

1. There is an affordable method of using automated machine learning (binary classification) for sales predictions in small company within the budget less than 1% of its income yearly and with accuracy more than 80%.
2. There is an algorithm which utilizes predictions from above mentioned AutoML binary classification model to calculate the predicted income based on current sales leads in different stages of its life-time. The prediction accuracy would also be more than 80%.

RESEARCH METHODS:

Ad 1. The data have been implemented to the model. The parameters of binary classifications have been set and model was tested. In evaluation phase model metrics have been reviewed with special focus on confidence score and features relevance. Then the model was deployed and made available for use. For 6 months each new sales lead has been tested and received prediction result from the model. It was decided to be invisible for sales persons in order not to interfere with normal procedures. After completion of its life-cycle the prediction result has been confronted with real-life result to measure accuracy. The computation and resources cost has been automatically calculated by Google Cloud Platform and compared to company income in given period.

Ad 2. The method is to find a formula that may calculate predicted income of the company based on the pile of SLs in different stages. Usually such methods are based on every SL probability and predicted value and deliver very poor accuracy. The conceptual work around algorithms and composite measures have been done in order to find improved formula which would return value closest to real.

RESULTS

The data prepared for the model has one thousand records and has a structure shown in the Table 2. It is being exported to GCP in form of CSV file. After importing it had to be verified if the columns contained proper data format numerical or categorical and determine which field is meant to be target output – in our case “success”. Another check has to be made on each feature if the distinct values number is correct, ex. Pipeline Funnel is a probability measure from P5 to 0 then distinct value supposed to be 6. The parameters of binary classifications have been set and model was tested.

Table 2. Example of a record including SL data

AUC PR	AUC ROC	Accuracy	Log Loss	Score Threshold
0.97	0.99	89.35%	0.016	0.5

Source: own work based on company’s MySQL database with sales leads (SL) data

Table 3. Accuracy of implemented binary classification. Model metrics are generated based on the less common being the positive class

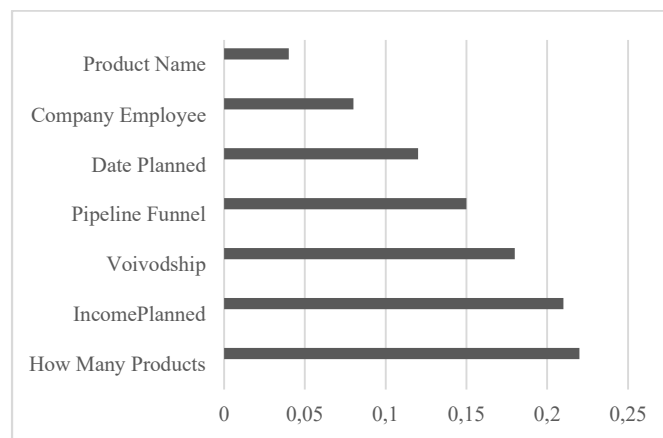
Voivodship	Income Planned	Date Planned	Pipeline Funnel	Company Employee	Product Name	How Many Products	Success
Kujawsko-Pomorskie	12500	2	P4	JK	PZP	2	false

Source: own work based on results shown in Google Cloud Platform model:

aut_ml_example_20190420105640

The model accuracy appeared to be reasonably high: 89,35% and feature importance distribution shows that there is no ‘noise’ in data as shown at the Figure 1. However small importance of product category was a surprise. After deploying the model to GCP user has to feed it with new data for prediction. It has been done by JSON format. Therefore it has been easily connected with existing system for registering SL data status by REST-API web service by http request type; GET. The data of each new sales lead has been stored to database.

Figure 1. Feature importance distribution of implemented binary classification model



Source: own work based on results shown in Google Cloud Platform model:
aut_ml_example_20190420105640

For 6 months each new sales lead has been tested and received prediction result based on described, automated machine learning binary classification model.

Table 4. Accuracy of predictions in given timeframe

	SL created	SL closed	False prediction	True prediction
July 2022	12	7	1	6
August 2022	26	18	2	16
September 2002	34	45	0	45
October 2022	91	57	11	46
November 2022	45	82	17	65
December 2022	22	21	0	21
	Totals:	230	31	199
	Accuracy:	86.52%		

Source: own work based on results taken from Google Cloud Platform model:
aut_ml_example_20190420105640 and stored in company's MySQL database

It was decided to be invisible and inaccessible for sales persons in order not to interfere with normal procedures and achieve highest purity of comparison. After

completion of its life-cycle the prediction result has been confronted with real-life result to measure accuracy as shown in Table 4.

The computation and resources cost has been automatically calculated by Google Cloud Platform and compared to company income in given period.

The average cost of quarterly use of the system is around 480 PLN which meets the criterion of total cost less then 1% of an income, which (together with accuracy) proves the hypothesis number one.

However such prediction even though it is very accurate brings only limited values to the company (selects most promising leads to focus efforts). Very important measure from stakeholder perspective is accurate sales value prediction in order to plan the budget and targets. The method is to find a formula that may calculate predicted income of the company based on the pile of SLs data in different stages. Among the most commonly used features of SLs we find (some of them have not been used in binary classification AutoML model):

- expected sales value V_{ps}
- real sales value V_{rs}
- expected completion date T_{ps}
- real completion date T_{rs}
- customer details
- product details
- salesperson details
- input probability of execution P_k
- output probability of execution P_0

Among the less analyzed features, there are some characteristics of the SL life cycle expressed in probability variability. You can also specify SL monetary value parameters, e.g. expected margin. In our case, for simplicity, we assume a relatively constant average margin. Thus, the amount of the revenue is of stakeholders interest.

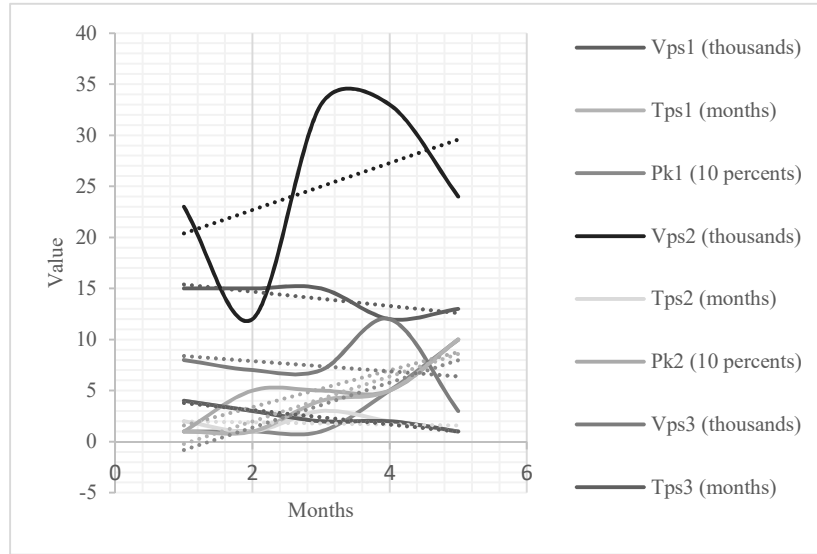
Table 5. Feature values of three chosen SLs over 5 months period

	month 1	month 2	month 3	month 4	month 5
V_{ps1} (thousands)	15	15	15	12	13
T_{ps1} (months)	4	3	2	2	1
P_{k1} (10 percents)	1	1	1	5	10
V_{ps2} (thousands)	23	12	33	33	24
T_{ps2} (months)	2	1	3	2	1
P_{k2} (10 percents)	1	5	5	5	10
V_{ps3} (thousands)	8	7	7	12	3
T_{ps3} (months)	4	3	2	2	1
P_{k3} (10 percents)	1	1	4	5	10

Source: own work based on results taken from company's MySQL database

There are usually around 6 levels of probability. In absolute terms, it is usually assumed that 5 to 10 percent of the SL turns into real sales. Further considerations are aimed at finding relationships between data that affect the predicted value of the sales funnel in the most effective way.

Figure 2. Feature values of three chosen SLs over 5 months period



Source: own work based on results taken from company's MySQL database

The gathered data (Table 5 and Figure 2) shows the great variability and diversity of SL features behavior over time. We clearly see that finding a proper formula for value prediction is extremely complicated without AI models. Traditional methods of calculating SLs value are based on probability and predicted value exclusively and treat probability as a weight for expected value. Sum of all elements gives predicted total value:

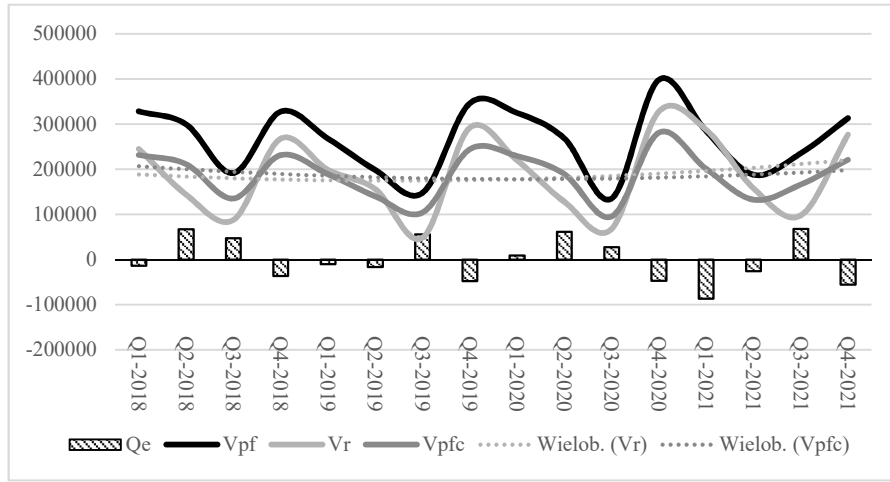
$$V_{pf} = \sum_{i=0}^n (V_{psi} \cdot P_{ki}) \quad (1)$$

Given formula has been tested on historical data from year 2018 to 2021. The assumption is to calculate the expected value of sales at specific points in time, which are the beginnings of the quarters, and the forecast is to apply to all sales for the coming quarter. Therefore, the formula (1) includes only those leads that at given moment have a parameter of expected execution time less than three months. The comparison to each quarter real sales value resulted in calculation on average error of prediction. As shown on Figure 3 the forecasts were quite inaccurate and mostly too optimistic. It seems to be general rule since as it was stated in literature review, usually 90% of SL end up with no sales which results in strong positive feedback on predictions from such a large base of SLs. Therefore we add coefficient which is calculated from average error (σ_M) in percents, updated at the end of each quarter :

$$(T_{ps} \leq 3) \rightarrow V_{pfc} = \sum_{i=0}^n (V_{psi} \cdot P_{ki}) \cdot (1 - \sigma_M) \quad (2)$$

As a result we see that overall trend of predictions and real values of sales in quarters are very similar, however in most periods the predictions have significant errors (Q_e) to real sales value (V_r) and overall accuracy is much less then requested 80%.

Figure 3. Comparison of predictions of sales against real value and ultimate errors in quarters



Source: own work based on results taken from company's MySQL database

Thus an idea appeared to implement the predictions results of AutoML binary classification model into the formula. It was modified in such a way that SLs that have a value of success $P_{aml} = 1$, so they were marked as successful by AutoML binary classification model where excluded from previous formula (2) and calculated differently, using the value of probability P_{true} , calculated as a percentage of successfully predicted sales leads by AutoML (updated after each quarter).

$$(T_{ps} \leq 3) \cap (P_{aml} \neq 1) \rightarrow V_{pfc1} = \sum_{i=0}^n (V_{psi} \cdot P_{ki}) \cdot (1 - \sigma_M) \quad (3)$$

$$(T_{ps} \leq 3) \cap (P_{aml} = 1) \rightarrow V_{pfc2} = \sum_{i=0}^n (V_{psi} \cdot P_{true}) \quad (4)$$

$$V_{pfr} = V_{pfc1} + V_{pfc2} \quad (5)$$

The value of V_{pfr} has been tested on quarters when the AutoML binary classification model has been implemented, so in last two quarters of year 2022. The results are shown in Table 6 and prove hypothesis number two with the reservations described in the discussion.

Table 6. Comparison of predictions of sales against real value with errors and accuracy in last two quarters of year 2022

	Q3-2022	Q4-2022
Vr	127567	321876
Vpfr	147235	289765
Qe	-19668	32111
Accuracy	84.58%	90.02%
Accuracy for both quarters:	87.30%	

Source: own work based on results taken from Google Cloud Platform model: aut_ml_example_20190420105640 and stored in company's MySQL database

DISCUSSION

The major two questions raised by stakeholders of the company were: how to identify sales leads (SLs) that promise better outcomes and how to calculate accurate income predictions at least one quarter ahead. Both hypotheses in the paper were formulated in a similar way and the research achieved its intended goals. However, several points need further clarification before an assessment of the overall outcome can be made.

During the observation of historical data and interviews with sales staff, dependencies were observed indicating that customers from certain regions (e.g., Wielkopolskie and Dolnośląskie voivodeships) had a higher likelihood of conversion. This may be connected to market saturation and the regional diffusion of product information. Some of these factors were not directly included in the predictive formula but were implicitly captured by the binary classification model.

It is worth noting that model validation procedures could be extended in future studies. Although the model achieved a high accuracy score (89.35%), further validation using a clear training/test set split and cross-validation would enhance the robustness of the results. Additionally, possible overfitting should be examined, as the dataset - while sufficient for an initial proof of concept - was relatively small and based on data from one company only.

Some limitations also relate to the specific nature of the case study. The results are drawn from one small IT company operating in Poland, which limits the generalizability of findings to other industries or regions. Furthermore, seasonal effects were not explicitly modeled, and their potential influence on quarterly sales volume remains to be verified statistically. In future, incorporating seasonality and additional classifiers (e.g., quarter identifiers) could improve predictive accuracy and model interpretability.

Overall, despite these limitations, the research provides strong evidence that cost-effective, automated machine learning can substantially enhance decision-making for SMEs in the B2B sector.

CONCLUSIONS

The research confirmed that an affordable and accurate application of automated machine learning (AutoML) can be used to predict sales lead outcomes and estimate company income within the SME sector. The study proved that a binary classification model trained on a dataset of 1000 records achieved high predictive accuracy at a cost below 1% of the company's annual income.

The process of model validation included training on historical data and evaluation based on accuracy metrics. Future extensions should apply explicit training/test data splits, use k-fold cross-validation, and monitor for possible overfitting. Such steps will strengthen the reliability and generalizability of the proposed method.

The paper also introduced an income prediction formula that integrates probabilistic weights with AI-based classification results. The combination of traditional probabilistic modeling and machine learning outcomes proved to improve forecasting accuracy, confirming the second hypothesis.

Nevertheless, the current solution is subject to several limitations, such as the narrow scope of one company, limited dataset size, and potential lack of generalizability to other business environments. Further research should include a wider range of companies, industries, and datasets to build a more universal framework.

In conclusion, the findings indicate that AutoML-based predictive modeling can serve as a practical, low-cost decision-support tool for SMEs. With proper validation and scaling, it holds promise for broader applications in sales management and beyond.

REFERENCES

- Bohanec M., Kljajić-Borštnar M., Robnik-Šikonja M. (2017) Explaining Machine Learning Models in Sales Predictions. *Expert Systems with Applications*, 71(1), 416-428. <https://doi.org/10.1016/j.eswa.2016.11.010>
- Civelek M., Ključnikov A., Fialova V., Folvarčná A., Stoch M. (2021) How Innovativeness of Family-Owned SMEs Differ Depending on Their Characteristics?. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 16(2), 413-428. <https://doi.org/10.24136/eq.2021.015>
- Eitle V., Buxmann P. (2019) Business Analytics for Sales Pipeline Management in the Software Industry: A Machine Learning Perspective. Retrieved from <http://hdl.handle.net/10125/59541> (19.01.2023)
- Feurer M., Klein A., Egensperger K., Springenberg J., Blum M., Hutter F. (2015) Efficient and Robust Automated Machine Learning. Retrieved from <https://papers.nips.cc/paper/2015/file/11d0e6287202fcd83f79975ec59a3a6-Paper.pdf> (19.01.2023)
- Gutierrez D. D. (2015) Machine Learning and Data Science: An Introduction to Statistical Learning Methods with R. Retrieved from <https://github.com/linux08/machine-learning-books/blob/master/Machine%20Learning%20and%20Data%20Science%20>

- %20An%20Introduction%20to%20Statistical%20Learning%20Methods%20with%20R%202015.pdf (19.01.2023)
- Huo Z. (2021) Sales Prediction Based on Machine Learning. 2021 2nd International Conference on E-Commerce and Internet Technology (ECIT), 410-415. <https://doi.org/10.1109/ECIT52743.2021.00093>
- Hyder A. S., Sundström A., Chowdhury E. H. (2022) Knowledge of Network-Based Market Orientation for the Internationalization of Disruptive Innovation in SMEs. *Central European Management Journal*, 30(3), 36-60. <https://doi.org/10.7206/cemj.2658-0845.81>
- Kaplan R. S., Norton D. P. (1996) Linking the Balanced Scorecard to Strategy. *California Management Review*, 39(1), 53-79. <https://doi.org/10.2307/41165876>
- Ključnikov A., Civelek M., Klimeš C., Farana R. (2022) Export Risk Perceptions of SMEs in Selected Visegrad Countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 17(1), 173-190. <https://doi.org/10.24136/eq.2022.007>
- Małkowska, A., Urbaniec, M., & Kosała, M. (2021) The Impact of Digital Transformation on European Countries: Insights from a Comparative Analysis. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 16(2), 325-355. <https://doi.org/10.24136/eq.2021.012>
- Mang'unyi E., Govender K. (2019) Exploring the Development Strategies of Rural Small, Medium and Micro Enterprises. *Central European Management Journal*, 27(4), 95-123. <https://doi.org/10.7206/cemj.2658-0845.11>
- Nogalski B., Niewiadomski P., Szpitter A. (2020) Agility Versus Flexibility? The Perception of Business Model Maturity in Agricultural Machinery Sector Manufacturing Companies. *Central European Management Journal*, 28(3), 57-97. <https://doi.org/10.7206/cemj.2658-0845.27>
- Oleksiuk A., Pleśniak A. (2022) Environment Characteristics and Internationalization of SMEs: Insights from a Polish and Finnish Sample. *Central European Management Journal*, 30(3), 175-194. <https://doi.org/10.7206/cemj.2658-0845.86>
- Peterson A. F. (1959) *Pharmaceutical Selling, "Detailing," and Sales Training*. New York, Heathcote-Woodbridge.
- Plawgo B., Ertman A. (2021) Competency Needs of Industry 4.0 Companies. *Central European Management Journal*, 29(4), 172-195. <https://doi.org/10.7206/cemj.2658-0845.64>
- Rohaani D., Topan E., Groothuis-Oudshoorn C. G. (2022) Using Supervised Machine Learning for B2B Sales Forecasting: A Case Study of Spare Parts Sales Forecasting at an After-Sales Service Provider. *Expert Systems with Applications*, 188, <https://doi.org/10.1016/j.eswa.2021.115925>
- Shen Z., Zhang Y., Lu J., Xu J., Xiao G. (2020) A Novel Time Series Forecasting Model with Deep Learning. *Neurocomputing*, 396, 302-313. <https://doi.org/10.1016/j.neucom.2018.12.084>
- Sleep S., Dixon A. L., DeCarlo T., Lam S. K. (2020) The Business-to-Business Inside Sales Force: Roles, Configurations and Research Agenda. *European Journal of Marketing*, 54(5), 1025-1060. <https://doi.org/10.1108/EJM-06-2018-0416>
- Stoop J. (2009) Developing a Reference Model for KPI and Dashboard Reporting in Sales & Marketing. Retrieved from <http://essay.utwente.nl/60699/> (13.01.2023)

- Soniewicki M. (2022) Is the Importance of Market Orientation Growing? A Study of High-Tech Manufacturing Companies. *Central European Management Journal*, 30(4), 150-174. <https://doi.org/10.7206/cemj.2658-0845.93>
- Zallocco R., Bolman-Pullins E., Mallin M. L. (2009) A Re-Examination of B2B Sales Performance. *Journal of Business & Industrial Marketing*, 24(8), 598-610. <https://doi.org/10.1108/08858620910999466>
- Yan J., Zhang C., Zha H., Gong M., Sun C., Huang J., Chu S., Yang X. (2015) On Machine Learning towards Predictive Sales Pipeline Analytics. *Proceedings of the AAAI Conference on Artificial Intelligence*, 29(1), 1945-1951. <https://doi.org/10.1609/aaai.v29i1.9455>