SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2026
ORGANIZATION AND MANAGEMENT SERIES NO. 246

THE APPLICATION OF THE CHAID ALGORITHM IN THE FIELD OF MANAGEMENT

Marlena SMUDA-KOCOŃ

University of Economics in Katowice; marlena.smuda-kocon@ue.katowice.pl, ORCID: 0000-0003-0468-5708

Purpose: The aim of this paper is to diagnose the main areas of usefulness of the Chi-squared Automatic Interaction Detector (CHAID) algorithm in explaining management problems and to assess its application possibilities in terms of dominant trends shaping functions of contemporary organisations.
Design/methodology/approach: The Systematic Literature Review (SLR) procedure was undertaken using the Web of Science. A set of publications was identified and selected, resulting in the development of a comprehensive database that served as a baseline for further conclusions.
Findings: The study identified key areas of application of the algorithm in management research, as well as the possibilities and limitations of its use in achieving measurable organisational effects. The results confirmed a steadily growing trend in the number of publications. However, predictive models using the algorithm are mainly used in segmentation research related to new customer acquisition, fraud detection and churn analysis. Despite the wide range of research using the CHAID algorithm, there are still areas where its untapped potential can be identified.
Practical implications: Machine learning algorithms (MLA) open up new opportunities for both management practice and theory. The application potential of these tools is particularly evident in relation to the analysis of large data sets and the personalisation of products and services. The detection of useful knowledge patterns using CHAID can contribute to increasing the effectiveness of operational and strategic activities in an organisation. Inspiring examples of classification using CHAID are provided by research in the financial, transport and service sectors.
Originality/value: The paper uncovers the possibilities of applying the algorithm in the field of management. These findings constitute a recommendation for decision-makers looking for tools to improve decision-making in specific areas of business operations. The results of a comprehensive review of the research prove that, thanks to its simplicity, the analysed algorithm can contribute to the discovery of non-obvious but useful patterns from raw data sets, while reducing costs and enabling a faster response to changing market conditions.
Keywords: algorithm CHAID, decision trees, digital transformation.
Category of the paper: literature review.

1. Introduction

Managerial decision-making requires searching for new ideas and tools that systematically improve this process and organise the complex reality of modern organisations. The need to structure business processes is widely emphasised in situations requiring rapid analysis of variable factors (Shamim, 2024), while maintaining objectivity and eliminating intuitive decision-making (Neiroukh et al., 2024). It is pointed out that companies using advanced data mining mechanisms, recognising patterns and useful regularities using machine learning tools, manage risk more effectively (Gubernat et al., 2025) and are able to instantly identify changing operating rules and customer needs and expectations. Furthermore, it is possible to discover new knowledge from available data, so data mining tools, which form the foundation of modern analytics, are booming (Wieczorkowski et al., 2021). Their diversity facilitates the researchers' work in building predictive models (cf. Łapczyński, 2010, p. 13) and enables organisations to model scenarios more effectively and increase the efficiency of management processes (Paramesha et al., 2024).

Scientific studies point to advanced data exploration methods, such as clustering, association rules, regression analysis, classification trees, random forest or neural networks (Wieczorkowski et al., 2021). The authors draw attention to their multidimensional application potential. Most partitioning algorithms are derived from three classic methods: CLS (Concept Learning Systems), AID (Automatic Interaction Detection) and CART (Classification and Regression Trees) (Breiman, 2017). Hou (2018) predicts the directions of development of the decision tree algorithm.

The Chi-squared Automatic Interaction Detector (CHAID) algorithm belongs to the largest family of data mining classification algorithms (Hou, 2018). It is widely emphasised that the method is distribution free (Hache et al., 2017) and fully non-parametric (Durica et al., 2019). In addition, the authors emphasize its simplicity and high level of interpretability, which creates opportunities for transposing research results into practical recommendations and setting further directions for scientific exploration. Interesting conclusions are provided by Smolarek's (2016) empirical research, according to which the most commonly used group of knowledge discovery algorithms are those that allow the construction of a classification model (such as CHAID), with as many as one third of the surveyed companies declaring that they use these classification methods (Smolarek, 2016, p. 100) in HR, among other areas.

Currently, however, it can be assumed that in the face of the increasing dynamics and complexity of phenomena shaping the management of modern organisations, the usefulness of most methods and tools (including the discussed CHAID) is being evaluated for suitability. The need to manage in conditions of information overload (Toffler, 1984) and the greater precision of analytical tools do not always, paradoxically, mean a more accurate diagnosis.

In this context, it is crucial to identify the areas of usefulness of particular methods and tools in explaining contemporary management problems.

Consequently, the aim of this article is to identify the current state of knowledge regarding the use of the selected CHAID algorithm, in particular to recognise its leading applications in management science research. This aim is part of a broad scientific discussion on the phenomenon of algorithmisation in various areas of business operations. The need for even simplified and synthesising research in this area is further supported by its practical value in reducing organisational complexity and accelerating the diffusion of knowledge.

The CHAID (Chi-squared Automatic Interaction Detector) algorithm is used to construct a graphical model (classification tree) that divides a set of observations into maximally homogeneous, disjoint subsets based on a selected categorical dependent variable (Łapczyński, 2010). The most significant independent variable appears in the first node of the classification. The divided set (entire data set) is the so-called parent node (Node_0), while the subsequent subsets are child nodes (Node 1, Node 2).

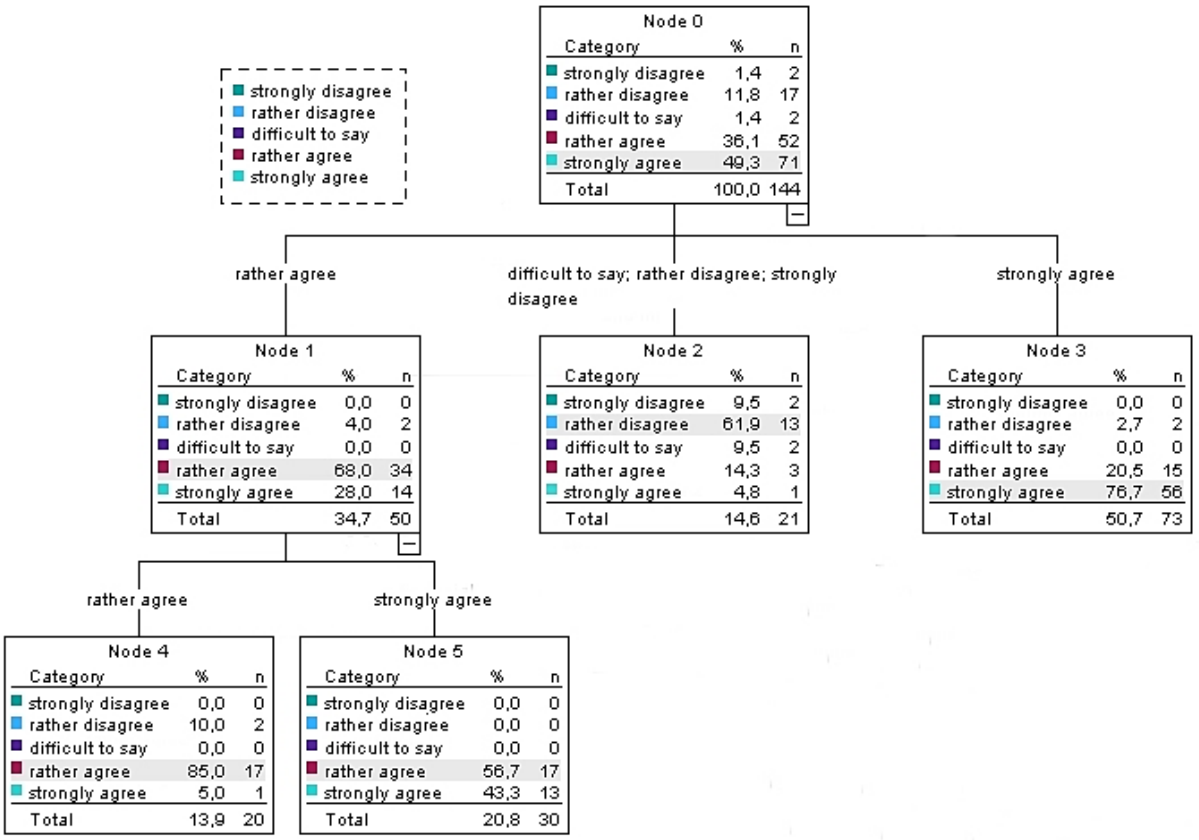

Figure 1. Classification tree CHAID – graphical model.

Source: Author's own study (IBM SPSS Statistics).

Unlike other statistical tools (logistics regression or cluster analysis), CHAID enables optimal splits to be obtained by selecting the predictor that exhibits the most powerful interaction with the endogenous variable (Chang et al., 2023; Higueras-Castillo et al., 2023). The method was popularised by Breiman (2017), Friedman, Olsen and Stone in 1984,

and is described in detail by Gatnar and Walesiak (2004). The use of the algorithm enables decision-makers to provide data, transformed into useful knowledge, in the form of classification or decision rules. Discovering classification rules is one of the most well-known and common tasks in the field of data mining, consisting in finding a way to map data to a set of predefined classes (Smolarek, 2016). This is also where their great usefulness comes from, as managers can directly use these rules to make decisions without the need for complex calculations (Azar et al., 2013).

On the one hand, the ease of interpreting the results obtained is a feature of the algorithm that should prove decisive for its wide application in management sciences. On the other hand, after Choudhury et al., (2020), despite the promise of machine learning in increasing productivity, many firms have encountered significant challenges due to biases in predictions, resulting from either biased training data and/or biased algorithms (Baer, Kamalnath, 2017). Consequently, the above findings further justify the need to assess the tool's application possibilities, especially since contemporary market trends significantly increase the complexity of decision-making processes. The literature points to several key trends that currently create an environment of high uncertainty, including digital transformation, sustainability issues and socio-demographic changes that affect the functioning of modern organisations.

2. Research method

In the context of the formulated objective, the challenge was to identify the most significant and current scientific articles. Using the Systematic Literature Review (SLR) procedure, scientific studies were selected that formed the basis for the development of knowledge about the application potential of the algorithm under study and, indirectly, allowed the opportunities and challenges resulting from the algorithmisation and automation of decision-making processes to be identified.

Bibliometric analysis is a popular and rigorous method for exploring and analysing large volumes of scientific data. It enables the evolutionary nuances of a specific field to be unpacked, while shedding light on the emerging areas in that field (Donthu et al., 2021, p. 285).

A set of publications was extracted from the indexed Web of Science (WoS) database and selected, resulting in the creation of a database that served as a starting point for further inference. The SLR was conducted in stages, in accordance with the approach of Tranfield et al. (2003): (1) planning the review, (2) conducting the review, and (3) reporting.

The starting point was the keywords: "CHAID" and "Chi-squared Automatic Interaction Detector", which were used to identify 1429 records. Then, restrictions were applied based on: document types, Web of Science Categories and English language (Table 1). Finally, the 32 records obtained were subjected to abstract contents and full-text contents analysis.

Table 1.
Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)

Step in searching	The number of records
	WoS
1. Keywords: "CHAID" or "Chi-square Automatic Interaction Detection" (status: 2024.06.30)	1429
2. Search terms applied: a) Document types: journal articles (final and open access) b) Subject area (Web of Science Categories): 'Business' (15) +'Management' (14) + 'Economics' (14) + 'Business Finance' (5) c) Language: English	 614 48 40
3. Total number of records – the result of abstract contents analysis (Records were removed if they were not connected to CHAID)	36
4. Total number of articles - the result of full-text article contents analysis	32
5. Descriptive analysis: Author, year, journal details, industry, and geographical context	

Source: Author's own study.

It was decided not to impose any time restrictions on the records examined in order to obtain a cross-sectional picture of the use of CHAID in scientific research over the years. The final, synthetic report took into account, first and foremost, the most-cited papers, and the individual records were supplemented with an indication of the industry context, the size of the sample studied, the definition of the dependent variable and predictive variables.

3. Results

Based on existing theoretical findings and using the SLR, areas of usefulness of the algorithm in explaining management problems were diagnosed. It was confirmed that the use of the CHAID algorithm enables the transformation of large amounts of data into knowledge and the discovery of patterns in various areas of management of organisations of different types and sizes. A summary of the research (Table 2) confirms the possibility of flexible and multidimensional analysis using the machine learning tool in question.

Analysis of the collected empirical material allowed two categories of conclusions to be formulated. The first group consists of empirical conclusions concerning the identified areas of the algorithm's application to date. The second group consists of methodological conclusions, allowing potential possibilities to be identified for using the algorithm in the future, taking into account its limitations.

Table 2.
Overview of the research designs (dependent variable, predictive variables and samples)

Author/s, Title	Problem	Sample	Predictive variables
Financial management and banking			
Al-Najjar et al. (2022), Machine Learning to Develop Credit Card Customer Churn Prediction	Predicting credit card customer churn in the banking sector	1627 churned customers	Total transaction count, total revolving balance, and changes in transaction count
Náñez Alonso et al. (2025), Evaluating the acceptance of CBDCs: Experimental research with artificial intelligence (AI) generated synthetic response	Identification of factors of acceptance or rejection of central bank digital currencies	663 synthetic responses	Educational level, financial experience, and income level
Durica et al. (2019), Decision tree based model of business failure prediction for Polish companies	Predicting the financial difficulties of companies	28,908 companies	Twelve financial indicators and company size
Popescu, Dragota (2018), What Do Post-Communist Countries Have in Common When Predicting Financial Distress?	Identification of financial distress predictors for five post-communist countries	20,119 firms	Total Assets and Net Assets Turnover (Croatia), Interest Coverage (Czech Republic), and Profit per Employee (Romania)
Yousaf, Dey (2022), Best proxy to determine firm performance using financial ratios: A CHAID approach	Investigation of the best predictor of firm performance among different proxies	287 Czech firms: automobile, construction, manufacturing sectors	Return on assets (ROA) as a robust proxy to predict financial performance
Maličká (2024), Investigating the Determinants of Local Indebtedness in Slovakia: The Machine Learning Approach	Investigation of local indebtedness determinants in a sample of Slovak municipalities	2926 municipalities	Current expenditure, subsidies, size category, and crises on local indebtedness
Human resources management			
Azar et al., (2013), A model for personnel selection with a data mining approach: A case study in a commercial bank	Developing a model for personnel selection	Case study of a commercial bank	Province of employment; education level; exam score; interview score and work experience
Aşkun et al., (2021), Comparative Analysis of Factors Affecting Employee Performance According to Job Performance Measurement Method: The Case of Performing Artists	Defining and comparing factors affecting employee job performance	305 participants from seven countries	Nationality
Botha (2024), Using entrepreneurial competencies and action to profile entrepreneurs: a CHAID analysis approach	Determining which entrepreneurial competencies predict entrepreneurial action for distinct profiles	1150 South African entrepreneurs	Growth mindset
Silva et al. (2019), Managers' leadership style and the commitment of their team members: associating concepts in search of possible relations	Relationships between the leadership style and the components of the commitment.	587 observations	Time working in the current organization
Bach-Pejic et al., (2020), Data mining approach to internal fraud in a project-based organization	Prevention of fraudulent working-hour claims in project-based organizations	300 employees	Customer, origin and level of expertise of the consultant, and cost of the consulting services

Cont. table 2.

Markowska-Przybyła (2023), External Social Capital of Enterprises in Poland	Identification of level of external social capital	650 companies	Age, firm size and the nature of the market
Guevara-Otero (2025), Factors that determine the level of individual entrepreneurial orientation	Segmentation of emerging entrepreneurs	477 students at a public university	Entrepreneurial orientation among young entrepreneurs
Consumer and social research / marketing segmentation			
Adá-Lameiras (2024), The impact of health crisis on sports consumption – A longitudinal study	Identification of sports spending patterns	20,000 households	Gender, age, and socio-cultural level
Bach-Pejic et al. (2021), Churn Management in Telecommunications: Hybrid Approach Using Cluster Analysis and Decision Trees	Improving customer attrition management to explain customer churn in the telecommunications industry	7043 respondents	Bad initial experience, low satisfaction with services
Borsje et al. (2023), Assessing passenger preferences for Bus Rapid Transit characteristics: A discrete choice experiment among current and potential Dutch passengers	Assessing passenger preferences for Bus Rapid Transit characteristics	1019 Dutch civilians	Frequency, service hours, stop spacing and reliability
Chang et al. (2023), Determinants of Subscription Renewal Behavior in Sport Spectatorship Services: A CHAID Decision Tree Modeling Approach	Season ticket renewal prediction model for regular seats	7982 observations	Recent attendance, discounted tickets, and affordability
Cheng (2023), Predicting abnormal trading behavior from internet rumor propagation: a machine learning approach	Predicting abnormal trading behavior from internet rumor propagation	476 over-the-counter stocks listed on TPEx	The rumor propagation outperforms traditional fundamentals in predicting abnormal trading behavior
Diaz-Perez & Bethencourt-Cejas (2016), CHAID algorithm as an appropriate analytical method for tourism market segmentation	Tourism market segmentation	324 tourists	Tourist nationality
Galiano-Coronil et al. (2023), An approach for analysing and segmenting messages about the SDGs on Twitter from the perspective of social marketing	Analysing and segmenting messages about the SDGs on Twitter from the perspective of social marketing	2433 tweets	Impact, social marketing, and happiness
Galiano-Coronil (2023), Social media impact of tourism managers: a decision tree approach in happiness, social marketing and sustainability	Identifying tourism managers' messages with high impact	9458 observations	Emotions, social marketing and topics
Hache et al. (2017), Beyond average energy consumption in the French residential housing market: A household classification approach	Energy consumption habits	42,963 French households	Heating fuel or equipment type, dwelling type and size, and construction date

Cont. table 2.

Higueras-Castillo et al. (2023), Intention to use e-commerce vs physical shopping. Difference between consumers in the post-COVID era	Segmentation of e-commerce consumers vs. physical shoppers in the post-COVID-19 era	491 Portuguese and 345 Spanish users	Habit, trust, and effort expectancy
Higueras-Castillo et al. (2025), The consumer intention to use e-commerce applications in the post-pandemic era: a predictive approach study using a CHAID tree-based algorithm	Customer segmentation (users e-commerce platforms)	836 observations	Trust, habit, effort expectancy
Walker et al., (2017), Antecedents of Retweeting in a (Political) Marketing Context	Retweeting trends in political marketing	42,444 original tweets	Number of followers
Sustainable development, digital transformation, ESG			
Saáry et al., (2022), Profit or less waste? Digitainability in SMEs - a comparison of Hungarian and Slovakian SMEs	Identification of main driving forces in SMEs	210 SMEs	Digitalisation and sustainability
Juma-Michilena et al., (2023) How to increase students' motivation to engage in university initiatives towards environmental sustainability	Establishing a mechanism to predict behaviour and motivation for sustainability	1446 students from nine Latin American universities	Environmental initiatives proposed by the university, specific courses focusing on climate change
Doğan et al., (2025) Credit rating prediction with ESG data using data mining methods	Integration of environmental, social, and governance pillar scores into traditional financial performance indicators in credit risk prediction	6622 firms	Key financial metrics such as leverage (LEV) and return on assets (ROA)

Source: Author's own study.

The results of the research indicate that in the existing literature on the subject, the CHAID algorithm is used for research focusing on four main aspects.

The first of these is financial management and banking, including the analysis of issues such as business failure, predicting the credit rating of companies, internal fraud and managing bank customer loyalty. The latest research (Doğan et al., 2025) concerns predicting the credit rating of companies.

The second area of issues is personnel and HR management, in particular the identification of key variables affecting employee selection and retention, e.g. in the banking sector. Using CHAID, attempts are made to identify groups of employees at the highest risk of churn in order to design dedicated support programmes for them (Bach-Pejic et al., 2021).

The third area covers consumer and social research, in particular concerning the energy consumption and tourism market segments, and the identification of key dimensions of entrepreneurial orientation and financial literacy. This area also examined respondents' satisfaction with transport services and the detection of regularities in the real estate market and the stock exchange. Authors (Cheng et al., 2023) train decision trees to demonstrate the impact of rumour propagation on stock trading behaviour. Sources indicate that CHAID algorithms are

particularly valued in these areas (Díaz-Pérez, 2016) because they allow decision-makers to identify the hierarchy of importance of factors and take precise action.

However, the last identified area of application of the CHAID algorithm deserves special attention, in which researchers address the issues of sustainability, digitalisation (Saáry et al., 2022) or incorporating ESG data into traditional accounting and financial measures (Doğan et al., 2025). It can be assumed that the complexity of issues in these research areas means that detecting useful knowledge patterns using a specific CHAID algorithm is not a common research approach to, for example, the problem of digital transformation. Furthermore, interpreting the classification results obtained can be challenging if it requires taking into account new, complex or previously unrecognised contextual conditions.

The publications listed (Table 2) are largely theoretical and empirical studies. In the introduction, the authors emphasise the usefulness of the tool for the area of problems described. Researchers commonly point to the multidimensional benefits of implementing data mining, including simplicity, computational speed, and ease of application of algorithms, i.e., CHAID. They also suggest the possibility of identifying so-called outliers in the learning process, which are sent to a separate child node without affecting further division. Among others, Díaz-Pérez et al., (2016) suggest that the analysis based on CHAID matches the nature of the tourism market segmentation problem better than those provided by discriminant analysis. Other studies emphasise the possibility of automating processes using the results of CHAID analysis, constructing a unique, customer-tailored product offer, operating in line with the consumer-centric idea, identifying new revenue streams or creating new business solutions in companies from various industries (cf. Hou, 2018).

The tool discussed has great application potential because CHAID is a non-parametric statistical technique (Higueras-Castillo et al., 2023). As emphasised in scientific studies, the ability to visualise the results obtained and map decision paths is also important, as it further improves their clarity and ease of interpretation by potential decision-makers (Yousaf, Dey, 2022). This is a starting point for practitioners and managers to make wider use of the CHAID algorithm as a tool that is appropriate for the individualised needs of their organisations. As shown in Table 2, grouping and classification using the algorithm can be extremely helpful in understanding the changes taking place in various sectors. Knowledge discovery algorithms can be used in almost every area of organisational management. They contribute to the interpretation of economic practice and, as a result, have become part of the scientific debate. Despite the very wide range of research using the CHAID algorithm, there are still areas where its untapped potential can be identified.

4. Discussion

The literature review confirmed the belief in the extensive explanatory and classification capabilities offered by CHAID. The algorithms are gaining recognition among researchers, but the lack of adequate data may be a significant obstacle to the implementation of such solutions. Before data can be transformed into useful knowledge, it must be combined, integrated, cleaned and visualised using tools that enable it to be manipulated (Slinger, Morrison, 2014). Preparing data for grouping and classification eliminates possible inconsistencies in the results. The correctness of the classification rules formulated depends on the quality of the input data used in the tree construction process. Decision trees can be sensitive to changes in the structure of the training set, so a small change in the input data can mean a significant change in the structure of the tree. However, decision trees require less data preparation than many other techniques, e.g. categorical variables are a problem for neural networks and statistical techniques (Hancock, Khoshgoftaar, 2020).

From the perspective of the limitations of CHAID, the fundamental issue is the decision-maker's responsibility for the ultimate consequences of the decisions made. It is impossible to disagree with Sienkiewicz (2018) that, in each case, limiting human involvement and supervision of the work process can lead to serious ethical and managerial challenges. Therefore, the prerequisites for the smooth implementation of new solutions (including those using LM algorithms) are commitment, good practices and adequate human capital resources. Furthermore, it should also be borne in mind that the algorithm is based on retrospective data, which may limit its usefulness in conditions of rapid market changes. At this point, a legitimate doubt may be raised as to the validity of its application to complex, emerging issues or trends. The literature emphasises that research hypotheses determine the choice of research approach, method and tools (Kumar, 2019).

On the one hand, the systematic literature review results may suggest that the increasing complexity of management and economic phenomena requires the use of more sophisticated models and tools as Random Forest, Gradient Boosting Machine (Luo et al., 2025). On the other hand, however, researchers point out that single decision trees are prone to overfitting and may exhibit limited predictive performance on complex datasets (Yin, Pierce, 2024).

5. Conclusion

In conclusion, it can be noted that the issue of the application of algorithms and classification trees is very broad, and this article and the conclusions formulated therein represent only a small part of this issue. As a result of reconstructing the main areas of

application of the algorithm, it can be concluded that, in particular, the ability of the CHAID algorithm to integrate unstructured data (qualitative and quantitative) determines its potential as a modern analytics tool. The algorithm fits perfectly into the digitainability trend (Saáry et al., 2022). Works such as Saáry et al. (2022), prove that this algorithm can be, for example, an effective tool for exploring attitudes towards digitalisation in the context of sustainable development. Moreover, the research results obtained may provide grounds for reorganising business activities. The adequacy and validity of using the CHAID algorithm to study current market changes and detect patterns in this area remains a complex issue. Its high interpretability must always be weighed against the limitations of processing highly volatile data. The recommendation from previous research concerning the need to develop a hybrid framework combining CHAID with advanced ensemble methods still seems relevant. Considering contemporary management trends, the development of appropriate tools adapted to the needs of a dynamic environment remains a serious challenge for business and the scientific community.

The future direction of research is to further verify the effectiveness of the tool in question and the scale of its use in business practices. It can be assumed that the need for real-time knowledge acquisition reported by decision-makers and the opportunities offered by the use of ML tools will stimulate further scientific exploration.

References

1. Adá-Lameiras, A., Antonovica, A., de Esteban Curiel, J., Aydogan, M. (2024). The impact of health crisis on sports consumption – A longitudinal study. *Journal of Retailing and Consumer Services, Vol. 77,* doi:10.1016/j.jretconser.2023.103656
2. Al-Najjar, D., Al-Rousan, N., Al-Najjar, H. (2022). Machine Learning to Develop Credit Card Customer Churn Prediction. *Journal of Theoretical and Applied Electronic Commerce Research*, *Vol. 17(4)*, pp. 1529-1542, doi:10.3390/jtaer17040077
3. Aşkun, V., Çizel, R., Ajanovıc, E. (2021). Comparative Analysis of Factors Affecting Employee Performance According to Job Performance Measurement Method: The Case of Performing Artists. *Ege Academic Review*, *Vol. 21(1),* pp. 29-45, doi:10.21121/eab.874012
4. Azar, A., Sebt, M.V., Ahmadi, P., Rajaeian, A. (2013). A model for personnel selection with a data mining approach: A case study in a commercial bank. *SA Journal of Human Resource Management*, *Vol. 11(1),* pp. 1-10, doi:10.4102/SAJHRM.V11I1.449
5. Bach-Pejic, M., Dumičić, K., Žmuk, B., Ćurlin, T. (2020). Data mining approach to internal fraud in a project-based organization. *International Journal of Information Systems and Project Management, Vol. 8(2),* https://aisel.aisnet.org/ijispm/vol8/iss2/5

6. Bach-Pejic, M., Pivar, J., Jaković, B. (2021). Churn Management in Telecommunications: Hybrid Approach Using Cluster Analysis and Decision Trees. *Journal of Risk and Financial Management, Vol. 14(11),* 544*,* doi:10.3390/jrfm14110544
7. Baer, T., Kamalnath, V. (2017). *Controlling machine-learning algorithms and their biases*. McKinsey & Company.
8. Borsje, R., Hiemstra-van Mastrigt, S., Veeneman, W. (2023). Assessing passenger preferences for Bus Rapid Transit characteristics: A discrete choice experiment among current and potential Dutch passengers. *Research in Transportation Economics, Vol. 100*, 101307, doi:10.1016/j.retrec.2023.101307
9. Botha, M. (2024). Using entrepreneurial competencies and action to profile entrepreneurs: a CHAID analysis approach. *Journal of Research in Marketing and Entrepreneurship*, *Vol. 26(2),* pp. 337-367, doi:10.1108/JRME-07-2022-0091
10. Breiman, L. (2017). *Classification and Regression Trees*. CRC Press.
11. Chang, Y., Warren, C., Katz, M. (2023). Determinants of Subscription Renewal Behavior in Sport Spectatorship Services: A CHAID Decision Tree Modeling Approach. *Sport Marketing Quarterly*, *Vol. 32(2),* pp. 124-136, doi:10.32731/SMQ.322.062023.03
12. Cheng, L.C., Lu, W.T., Yeo, B. (2023). Predicting abnormal trading behavior from internet rumor propagation: a machine learning approach. *Financial Innovation, Vol. 9(1),* pp. 1-23, doi:10.1186/s40854-022-00423-9
13. Choudhury, P., Starr, E., Agarwal, R. (2020). Machine learning and human capital complementarities: Experimental evidence on bias mitigation. *Strategic Management Journal, Vol. 41(8)*, pp. 1381-1411, doi:10.1002/smj.3152
14. Diaz-Perez, F., Bethencourt-Cejas, M. (2016) CHAID Algorithm as an Appropriate Analytical Method for Tourism Market Segmentation. *Journal of Destination Marketing & Management*, *Vol. 5,* pp. 275-282, doi:10.1016/j.jdmm.2016.01.006
15. Doğan, M., Chelery Komath, M.A., Sayilir, O. (2025). Credit rating prediction with ESG data using data mining methods. *Future Business Journal*, *Vol. 11(1)*, pp. 1-14, doi:10.1186/s43093-025-00490-1
16. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., Lim, W.M. (2021). How to conduct bibliometric analysis: An overview and guidelines. *Journal of Business Research, Vol. 133*, pp. 285-296, doi:10.1016/j.jbusres.2021.04.070
17. Durica., M, Frnda, J., Svabova, L. (2019). Decision tree-based model of business failure prediction for Polish companies. *Oeconomia Copernicana, Vol. 10(3),* pp. 453-469, doi:10.24136/oc.2019.022
18. Galiano-Coronil, A., Blanco-Moreno, S., Tobar-Pesantez, L.B., Gutiérrez-Montoya, G.A. (2023). Social media impact of tourism managers: a decision tree approach in happiness, social marketing and sustainability. *Journal of Management Development, Vol. 42(6)*, pp. 436-457, doi:10.1108/JMD-04-2023-0131

19. Galiano-Coronil, A., Ortega-Gil, M., Macías-Varela, B., Ravina-Ripoll, R. (2023). An approach for analysing and segmenting messages about the SDGs on Twitter from the perspective of social marketing. *International Review on Public and Nonprofit Marketing, Vol. 20(3)*, pp. 635-658, doi:10.1007/s12208-023-00369-3
20. Gatnar, I., Walesiak, M. (2004). *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych.* Wrocław: Wydawnictwo Akademii Ekonomicznej we Wrocławiu
21. Gubernat, E., Noga, B., Noga, M. (2025). *Zarządzanie ryzykiem transakcji biznesowych a bezpieczeństwo ekonomiczne jednostek samorządu terytorialnego*. Warszawa: CeDeWu.
22. Guevara-Otero, N., Vargas-Pérez, A.M., Segovia-Pérez, M., Laguna-Sánchez, P. (2025). Factors that determine the level of individual entrepreneurial orientation: profiles of. *International Entrepreneurship and Management Journal*, *Vol. 21,* pp. 1-31, doi:10.1007/s11365-024-01055-4
23. Hache, E., Leboullenger, D., Mignon, V. (2017). Beyond average energy consumption in the French residential housing market: A household classification approach. *Energy Policy, Vol. 107*, pp. 82-95, doi:10.1016/j.enpol.2017.04.038
24. Hancock, J.T., Khoshgoftaar, T.M. (2020). Survey on categorical data for neural networks. *Journal of Big Data, Vol. 7(28),* doi:10.1186/s40537-020-00305-w
25. Higueras-Castillo, E., Alves, H., Liébana-Cabanillas, F., Villarejo-Ramos, Á.F. (2025). The consumer intention to use e-commerce applications in the post-pandemic era: a predictive approach study using a CHAID tree-based algorithm. *European Journal of Management and Business Economics, Vol. 34(1),* pp. 103-119, doi:10.1108/EJMBE-12-2022-0375
26. Higueras-Castillo, E., Liébana-Cabanillas, F., Villarejo Ramos, Á.F. (2023). Intention to use e-commerce vs physical shopping. Difference between consumers in the post-COVID era. *Journal of Business Research*, *Vol. 157,* 113622, doi:10.1016/j.jbusres.2022.113622
27. Hou, Y. (2018) Decision Tree Algorithm for Big Data analysis. *Advances in Intelligent Systems Research, Vol. 161,* pp. 270-274, doi:10.2991/tlicsc-18.2018.43
28. Juma-Michilena, I.J., Ruiz-Molina, M.E., Gi-Saura, I., Belda-Miquel, S. (2023). How to increase students' motivation to engage in university initiatives towards environmental sustainability. *Journal of Consumer Affairs*, *Vol. 57(3),* pp. 1304-1323, doi:10.1111/joca.12527
29. Kumar, R. (2019). *Research Methodology: A Step-by-Step Guide for Beginners*. London: SAGE Publications.
30. Łapczyński, M. (2010). *Drzewa klasyfikacyjne i regresyjne w badaniach marketingowych*. Kraków: Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
31. Luo, G., Arshad, M.A., Luo, G. (2025). Decision Trees for Strategic Choice of Augmenting Management Intuition with Machine Learning. Symmetry*, Vol. 17(7),* 976*,* doi:10.3390/sym17070976

32. Maličká, L. (2024). Investigating the Determinants of Local Indebtedness in Slovakia: The Machine Learning Approach. *Journal of Economics*, *Vol. 72,* doi:10.31577/ekoncas.2024.07-08.02
33. Markowska-Przybyła, U. (2023). External Social Capital of Enterprises in Poland. *Polish Journal of Management Studies, Vol. 27(1)*, pp. 202-220, doi:10.17512/pjms.2023.27.1.12
34. Náñez Alonso, S.L., Ozili, P.K., Sastre Hernández, B.M., Pacheco, L.M. (2025). Evaluating the acceptance of CBDCs: Experimental research with artificial intelligence (AI) generated synthetic response. *Quantitative Finance and Economics, Vol. 9(1),* pp. 242-273, doi:10.3934/QFE.2025008
35. Neiroukh, S., Aljuhmani, H.J., Alnajdawi, S. (2024). In the Era of Emerging Technologies: Discovering the Impact of Artificial Intelligence Capabilities on Timely Decision-Making and Business Performance. *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS*). Manama, Bahrain, pp. 1-6, doi: 10.1109/ICETSIS61505.2024.10459356
36. Paramesha, M., Rane, N., Rane, J. (2024). Big Data Analytics, Artificial Intelligence, Machine Learning, Internet of Things, and Blockchain for Enhanced Business Intelligence. *Partners Universal Multidisciplinary Research Journal, Vol. 1(02),* pp. 110-133*,* doi:10.2139/ssrn.4855856
37. Popescu, M.E., Dragotă, V. (2018). What Do Post-Communist Countries Have in Common When Predicting Financial Distress? *Prague Economic Papers*, *Prague University of Economics and Business, Vol. 6,* pp. 637-653, doi:10.18267/j.pep.664
38. Saáry, R., Kárpáti-Daróczi, J., Tick, A. (2022). Profit or less waste? Digitainability in SMEs - a comparison of Hungarian and Slovakian SMEs. *Serbian Journal of Management, Vol. 17,* pp. 33-49, doi:10.5937/sjm17-36437
39. Shamim, M. (2024). Artificial Intelligence in Project Management: Enhancing Efficiency and Decision-Making. *International Journal of Management Information Systems and Data Science, Vol. 1(1),* pp. 1-6, doi:10.62304/ijmisds.v1i1.107
40. Sienkiewicz, Ł. (2018). *Optymalizacja użyteczności kapitału ludzkiego*. *Perspektywa przedsiębiorstwa*. Warszawa: Oficyna SGH.
41. Silva, P.L., Nunes, S.C., Andrade, D.F. (2019). Managers' leadership style and the commitment of their team members: associating concepts in search of possible relations. *Review of Business Management, Vol. 21(2),* pp. 291-311, doi:10.7819/rbgn.v21i2.3975
42. Slinger, G., Morrison, R. (2014). Will Organization Design Be Affected By Big Data? *Journal of Organization Design, Vol. 3(3),* pp. 17-26, doi:10.7146/jod.9729
43. Smolarek, M. (2016) *Technologie informatyczne w zarządzaniu kapitałem ludzkim w organizacji.* Katowice: Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.
44. Toffler, A. (1984). *Future Shock*. New York: Bantam Books.

45. Tranfield, D., Denyer, D., Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management, Vol. 14(3),* pp. 207-222, doi:10.1111/1467-8551.00375
46. Walker, L.W., Baines, P.R. (2017). Antecedents of Retweeting in a (Political) Marketing Context. *Psychology and Marketing*, *Vol. 34(3),* pp. 275-293, doi:10.1002/mar.20988
47. Wieczorkowski, J., Chomiak-Orsa, I., Pawełoszek, I. (2021). *Big Data w zarządzaniu*. Warszawa: Polskie Towarzystwo Ekonomiczne.
48. Yin, R., Pierce, B.G. (2024). Evaluation of AlphaFold antibody–antigen modeling with implications for improving predictive accuracy. *Protein Science, Vol. 33(1),* doi:10.1002/pro.4865
49. Yousaf, M., Dey, S.K. (2022). Best proxy to determine firm performance using financial ratios: A CHAID approach. *Review of Economic Perspectives*, *Vol. 22(3),* pp. 219-239, doi:10.2478/revecp-2022-0010