Volume 9, Issue 2, pp. 83-89, 2025.
International Journal of Scientific Engineering and Science
ISSN (Online): 2456-7361

Application of Artificial Intelligence in Business Analytics: A Case Study of Decision Tree Analysis

Vu Thi Thu Huong[1]
[1]Department of Mathematical Economics, Thuongmai University, Hanoi, Vietnam, 100000
Email: huong.vtt@tmu.edu.vn

***Abstract**: This article discusses the application of artificial intelligence in business analytics, focusing on decision tree analysis for innovation based on research and development investment and Information Technology application in enterprises. The study uses data from the 2021 Vietnam Enterprise Survey, which includes 811,092 active businesses. The decision tree model is designed to predict technological innovation, with the output variable being whether a company has introduced or improved products or processes. The input variables include internet use, websites, automated systems, software, and Research and Development investment. The article outlines the process of applying Artificial Intelligence in Decision Tree Analysis, including writing Python code on Google Colab to draw the decision tree based on data from a CSV file. The model helps businesses classify data, predict trends, and optimize strategies based on historical information. However, challenges such as overfitting, sensitivity to input data, and high computational requirements are also discussed. Despite these challenges, the article emphasizes the value of decision tree analysis in data-driven decision-making and how Artificial Intelligence, combined with advanced technologies like deep learning and Explainable Artificial Intelligence, can improve forecasting accuracy and enhance business performance optimization.*

***Keywords**: Artificial intelligence (AI), business analytics (BA), decision tree, technological innovation.*

I. Introduction

In the digital era, artificial intelligence (AI) has emerged as an essential component in business analytics (BA), playing a pivotal role in optimizing decision-making processes, forecasting trends, and enhancing operational performance (Davenport & Harris, 2007). The rapid advancements in big data, machine learning (ML), and cloud computing have facilitated the widespread integration of AI in BA, propelling digital transformation and boosting competitiveness (Brynjolfsson & McAfee, 2014).

Business analytics encompasses four key areas: descriptive, diagnostic, predictive, and prescriptive analysis, where AI significantly contributes by automating and enhancing the efficiency of data processing (Davenport & Harris, 2007). Technologies such as ML, deep learning (DL), and natural language processing (NLP) enable the analysis of structured and unstructured data, delivering valuable insights that drive business decisions.

AI plays a critical role in forecasting by recognizing patterns in historical data to predict future outcomes. Machine learning models, including decision trees, neural networks, and regression analysis, help businesses accurately predict consumer behavior, market trends, and financial risks. These AI-driven models provide improved forecasting accuracy, leading to better decision-making.

AI-powered decision support systems enable organizations to make data-driven decisions more swiftly and accurately. Decision trees, for instance, are user-friendly predictive tools that help businesses classify and forecast outcomes based on historical data. AI also facilitates real-time decision-making in finance, healthcare, and supply chain management, allowing companies to respond to changing conditions promptly.

AI plays a key role in automating Business Intelligence (BI) processes, reducing the reliance on human intervention in data analysis and reporting (Biecek & Burzykowski, 2021). BI platforms use NLP and ML algorithms to generate insights for decision-makers through dashboards and automated reports, enabling businesses to adjust to market dynamics and optimize operational strategies.

Despite its numerous advantages, AI in BA faces challenges, such as concerns over data security, algorithmic bias, and the need for high-quality data. Additionally, businesses must cultivate a skilled workforce to develop, manage, and interpret AI-driven models (Witten et al., 2011). As data from digital platforms continues to grow, organizations increasingly require AI to process and extract valuable insights from structured and unstructured data, enabling the analysis of social media, financial transactions, and customer behaviors.

AI algorithms enhance the accuracy of revenue forecasts, risk analysis, and market trend predictions. These models not only rely on historical data but also adapt to evolving factors over time. Key applications of AI include forecasting stock price volatility, assessing credit risk, predicting product demand, and optimizing supply chains.

Furthermore, AI automates business decision-making by analyzing real-time data and recommending optimal actions without human intervention. For example, AI-powered customer relationship management (CRM) systems can automatically identify potential clients, suggest suitable products, and personalize user experiences.

AI is also instrumental in fraud detection, securing data, and meeting regulatory requirements. In the finance industry, AI aids in detecting fraudulent transactions, preventing cyberattacks, and safeguarding customer privacy. As AI becomes increasingly integral to Business Analytics, it helps businesses leverage Big Data, streamline decision-making, and improve competitiveness. In the future, AI will continue to evolve, enhancing predictive analytics, supporting real-time decision-making, and fostering transparency.

To fully harness the potential of AI in BA, organizations must invest in robust data infrastructure, upskill their workforce, and ensure the protection of data security.

This research concentrates on investigating the opportunities and challenges of applying AI in BA through a comprehensive review. The application of AI in decision tree analysis is examined, particularly in analyzing the transformation of processes in Vietnamese enterprises, which depends on decisions regarding R&D investment and the application of information technology (IT).

II. Applications of Artificial Intelligence in Business Analytics

The integration of AI into BA is revolutionizing various aspects of data-driven decision-making. AI technologies, particularly large language models (LLMs), are being increasingly utilized across different types of analytics - prescriptive, predictive, and descriptive. The following studies illustrate some key applications of AI in BA, showcasing both the potential and the challenges that come with these advancements.

Firstly, Amarasinghe et al. (2023) developed an AI framework to automate problem formulation for business optimization, specifically within prescriptive analytics. Their framework leverages a code-generating LLM to assist in production scheduling, reducing the need for human expertise in problem structuring. This results in significant productivity gains for Business Analysts. However, the study highlights challenges such as overfitting, data token limitations, and the complexity of defining problem constraints, which can influence the model's effectiveness and success rate.

In the realm of Business Intelligence (BI), Liu et al. (2024) introduced the BIBench framework to evaluate the data analysis capabilities of 41 LLMs. This study focused on assessing the models' proficiency across three dimensions of BI: foundational knowledge, knowledge application, and technical skills. The results showed that models like BIChat and GPT-4 significantly outperformed others in generating insights and analyzing data. However, the study also pointed out the limitations of LLMs, particularly their inability to generate questions or perform complex tasks without domain-specific fine-tuning, underscoring the need for further customization in business applications.

Xing (2024) examined the use of LLMs for financial sentiment analysis (FSA), where AI models analyze unstructured data such as annual reports and social media posts to forecast market trends. Xing developed a unique framework that incorporates heterogeneous LLM agents, each simulating different mental functions based on Minsky's theory of mind and emotions. This approach improved the accuracy of sentiment predictions by considering various emotional cues in financial data. However, the study found that not all agents contributed equally to the performance, with certain agents (such as mood and rhetoric) proving to be more influential than others.

In handling complex data, Hu et al. (2024) proposed a modular framework called PyTorch Frame, which focuses on efficiently processing multi-modal tabular data. This framework employs a tensor-based approach to transform raw data into a format easily integrated with external foundation models, including LLMs for processing textual data. Their research demonstrated the utility of this framework in deep learning applications for tabular data analysis, showing how it can manage complex columns and improve data processing efficiency.

Wang et al. (2024) introduced the HARMONIC framework, which uses LLMs to synthesize tabular data while ensuring privacy protection. This framework addresses concerns related to data confidentiality by generating synthetic data that mimics real-world data without compromising sensitive information. This is especially useful in industries like healthcare and finance, where sharing real data is often prohibited by privacy regulations. The HARMONIC framework enables businesses to continue their data analysis processes using synthetic data while maintaining privacy compliance.

Additionally, Feuerriegel et al. (2024) explored the transformative potential of generative AI (GenAI) in Business Analytics, especially in generating new content such as text, images, and audio. Their study focused on how GenAI is reshaping industries like content creation, IT help desks, and healthcare. They highlighted the opportunities GenAI provides, such as the ability to enhance productivity and creativity, but also pointed out significant challenges, including bias and misinformation. The study emphasized the importance of developing new governance frameworks to manage the ethical implications of GenAI in business environments.

Moreover, Davenport (2018) discussed the transition from traditional BA to AI-driven analytics, emphasizing that AI can build on existing analytical capabilities. His research outlined the evolution of analytics through four stages: Analytics 1.0, 2.0, 3.0, and AI-powered Analytics 4.0. Davenport argued that AI's role in enhancing product development, business processes, and customer engagement offers vast potential. However, businesses must assess their analytical strengths, develop a structured AI strategy, and ensure that talent acquisition and partnerships are aligned for successful AI adoption.

Finally, Salazar and Kunc (2025) explored the integration of Generative AI (GenAI) with BA, particularly in research processes. The authors highlighted how GenAI tools can be used for literature review, summarization, and thematic analysis, accelerating academic writing and research. This study underlined the shift from traditional discriminative models to AI-driven generative models, offering businesses a path to enhance their analytical capabilities incrementally. Despite these advantages, the paper warned about the risks of over-reliance on AI, especially in terms of data quality, security, and model reliability.

In conclusion, these studies highlight the significant role AI plays in enhancing decision-making processes within BA across various domains. AI, particularly LLMs, brings automation, efficiency, and improved accuracy to analytics tasks. However, challenges such as the need for domain-specific fine-tuning, model explainability, and data privacy must be carefully managed. As the field of AI in BA continues to evolve, businesses need to overcome these challenges while utilizing AI's capabilities to achieve enhanced business results.

III. Challenges and Opportunities of Applications of AI in BA

A. Opportunities

The application of AI in BA is opening up tremendous opportunities for businesses to leverage data and enhance their competitive edge. One of the key benefits is improving the ability to forecast market trends, helping companies make quicker and more accurate decisions. AI models can analyze vast amounts of data from various sources, such as transactions, customer behavior, and macroeconomic data, to predict consumer demand and market trends with high accuracy (Davenport & Harris, 2007). This is especially valuable in sectors like retail, finance, and e-commerce, where timely decisions can provide a significant competitive advantage.

Additionally, AI helps businesses optimize operations and reduce costs. Previously, companies relied on traditional methods of analysis, which consumed significant time and resources to compile data and make decisions. With AI, automated systems can analyze data in real time, minimizing human errors and optimizing processes in production, supply chain management, and financial planning. For example, AI can forecast product demand, helping businesses optimize inventory levels, reduce warehousing costs, and ensure adequate product availability for customers.

AI also provides opportunities to personalize customer experiences, increasing conversion rates and revenue. AI systems can analyze shopping behaviors, personal preferences, and consumption habits to offer product recommendations or design personalized marketing campaigns. E-commerce platforms such as Amazon, Shopee, and Lazada use AI to suggest products based on users' purchases and search history, enhancing the customer experience and boosting sales (Biecek & Burzykowski, 2021).

Moreover, AI can improve risk management and fraud detection in business. In the financial sector, AI is used to analyze unusual transaction patterns, detect fraud signals, and issue timely alerts. Banks and financial institutions are also utilizing AI to evaluate credit risk and assess customers' ability to repay based on past data and pertinent economic indicators (Russell & Norvig, 2021). This not only helps reduce the risk of bad debts but also optimizes credit decisions, benefiting both businesses and customers.

Another significant opportunity is the integration of AI into decision-making strategies and business management. AI systems can analyze real-time data and provide dynamic reports that offer business leaders a comprehensive view of operational performance (Davenport & Harris, 2007). AI also helps managers evaluate employee performance, optimize resources, and suggest sustainable development strategies, improving operational efficiency in an increasingly competitive business environment.

Furthermore, with the development of Explainable AI (XAI), businesses now have more opportunities to apply AI while ensuring transparency and interpretability in business decisions. Previously, a major barrier to AI adoption was the "black-box" nature of machine learning models, making it difficult for businesses to explain why AI made a particular decision. However, XAI is helping overcome this issue by providing clear explanations of how AI analyzes data and arrives at conclusions, thereby increasing trust in AI systems.

In addition to operational benefits, AI generates opportunities for product and service innovation, helping businesses develop new business models. For example, many tech companies are leveraging AI to create automated chatbots, smart virtual assistants, and 24/7 customer support systems, which improve customer satisfaction and reduce the workload on human staff. In the manufacturing sector, AI is applied to develop predictive maintenance systems, enabling businesses to detect and resolve machine issues before they disrupt production.

Finally, with the rise of AI-as-a-Service (AIaaS), even small and medium-sized enterprises (SMEs) can access powerful AI tools without requiring significant investments in technology infrastructure. Cloud-based AI services make it easier for businesses to deploy AI, reduce initial costs, and take advantage of advanced machine learning models without requiring an in-house expert team. This allows businesses of all sizes to harness the power of AI to improve operations and enhance their competitive position in the market.

B. Challenges

Although AI offers many benefits for BA, the implementation of AI in this field still faces significant challenges. One of the biggest obstacles is the requirement for data quality. AI relies on data to function effectively, but many businesses struggle with collecting, storing, and cleaning data. If the input data is inaccurate, inconsistent, or biased, AI models may produce inaccurate results, affecting the quality of business decisions. Additionally, in some cases, data may contain biases due to uneven distribution across customer groups or market regions, reducing the objectivity of AI predictions (Witten et al., 2011).

In addition to data quality, security, and privacy are major challenges when applying AI in business analytics. Collecting and processing large amounts of personal and business data presents security risks, especially when AI is deployed on cloud platforms or uses data from various sources. Strict privacy regulations, such as the European General Data Protection Regulation (GDPR), require businesses to comply with personal data protection principles, which can slow down the AI implementation process or demand large investments in security systems (Biecek & Burzykowski, 2021). In the financial sector, AI systems can be exploited by hackers to manipulate market forecasts or carry out illegal transactions, necessitating organizations to improve and strengthen their security protocols regularly.

Another challenge when using AI in BA is the complexity of AI models and the ability to explain their results. Advanced machine learning models like deep learning often function as "black boxes," making it difficult for businesses to understand how AI makes decisions. This reduces managers' trust in AI

Volume 9, Issue 2, pp. 83-89, 2025.
ISSN (Online): 2456-7361

models and can hinder their adoption in critical decision-making processes (Russell & Norvig, 2021). To address this issue, businesses need to integrate Explainable AI (XAI) – an approach that enhances the explainability and transparency of AI models. However, developing XAI comes with high costs and requires deep technical expertise, making it difficult for many small and medium-sized enterprises (SMEs) to implement.

Moreover, the shortage of high-quality AI talent is a significant obstacle for businesses. AI requires teams of experts in data science, machine learning, and business analytics, but this workforce remains scarce, particularly in emerging markets like Vietnam. Businesses not only need to compete for talent but also must invest in internal training to enhance employees' skills. The lack of AI talent can prevent businesses from fully leveraging advanced analytics technologies and slow down digital transformation efforts (Witten et al., 2011).

Finally, the high initial investment cost is a major challenge when implementing AI in business analytics. Building an AI system requires robust infrastructure, advanced software, and high-quality data, leading to significant upfront costs. Additionally, businesses need to continuously maintain and update systems to ensure AI operates efficiently over time. For small and medium-sized enterprises (SMEs), the high costs can be a barrier, making them hesitant to adopt AI in business analytics. Some businesses have addressed this issue by using AI-as-a-Service (AIaaS) – cloud-based AI services that reduce initial investment costs, but still present risks related to data security and dependence on external providers.

Despite facing numerous challenges, AI remains an essential technology in business analytics. Businesses need a suitable strategy to overcome these barriers, including investing in data management, ensuring security, training human resources, and choosing AI models with better explainability. Only by addressing these challenges can businesses fully unlock the potential of AI and maintain a competitive advantage in an increasingly complex business environment.

IV. APPLICATION AI IN DECISION TREE ANALYSIS

A. Advantages and disadvantages of the decision tree in BA

Decision Tree Analysis (DTA) is one of the most popular methods in BA due to its visual nature, ease of understanding, and ability to make quick decisions. Decision tree models help businesses classify data, predict trends, and optimize business strategies based on historical data. With the development of AI, decision tree algorithms have been continuously improved to handle Big Data, minimize errors, and enhance model interpretability (Quinlan, 1996). As a result, DTA has become an essential tool in fields such as finance, marketing, supply chain management, and healthcare (Davenport & Harris, 2007).

One of the greatest advantages of decision trees in BA is their ability to visualize the decision-making process. Unlike complex machine learning models, DTA is easy to understand even for non-technical individuals, helping managers and business experts interpret results clearly (Biecek & Burzykowski, 2021). For example, in banking credit, decision trees can be used to classify customers by credit risk, enabling banks to make more accurate lending decisions without relying on complex statistical models. In addition to its simplicity, another advantage of DTA is its ability to work with both categorical and quantitative variables. While some statistical models like linear regression struggle with categorical variables, decision trees can effectively handle mixed data types (Witten et al., 2011). As a result, DTA is widely applied in customer behavior analysis, helping e-commerce businesses identify potential customer segments and optimize marketing campaigns.

However, despite its many advantages, the applications of DTA in BA also face several disadvantages. One of the biggest limitations of decision trees is the risk of overfitting. Without proper control, the model can become overly complex, overfitting the training data and losing the ability to generalize to new data. To address this issue, data scientists often apply pruning techniques or use ensemble models such as Random Forest or Gradient Boosting to improve model performance (Russell & Norvig, 2021).

Another challenge of DTA is its sensitivity to input data. If the training data has imbalanced groups or contains outliers, decision trees can become biased, leading to inaccurate predictions. For example, in healthcare, if the dataset contains only a small proportion of patients with severe illnesses, the decision tree model might underestimate health risks, affecting diagnostic decisions.

Furthermore, while decision trees provide easily understandable results, implementing them in large business systems can be difficult due to the high computational requirements as data size increases. Large decision tree models can become difficult to manage and require significant computational resources to handle Big Data (Gandomi & Haider, 2015). In such cases, businesses can use variations C4.5, CART, or ensemble models like Random Forest to improve performance while maintaining model transparency (Biecek & Burzykowski, 2021).

Although these challenges exist, decision tree analysis applications in BA continue to provide substantial value in data-driven decision-making. When combined with advanced technologies like AI and Cloud Computing, DTA can help businesses improve forecasting accuracy, optimize operational processes, and enhance customer experience (Davenport & Harris, 2007). In the future, decision trees will continue to be optimized with hybrid models, combined with deep learning and Explainable AI (XAI) to increase transparency and minimize risks in business analytics (Russell & Norvig, 2021).

B. The application of AI in decision tree analysis regarding the technological innovation status of enterprises depends on R&D investment and IT application.

Data

This study uses data from the 2021 Vietnam Enterprise Survey conducted by the General Statistics Office of Vietnam. The sample consists of 811,092 active enterprises as of

Volume 9, Issue 2, pp. 83-89, 2025.
International Journal of Scientific Engineering and Science
ISSN (Online): 2456-7361

December 31, 2020, excluding corporations and state-owned companies under the Ministry of Defense.

Variables

TABLE 1. Explanation of Input and Output Variables for the Decision Tree

Variable Symbol	**Variable Symbol**	**Variable Symbol**
Output variable		
Tech	Technological Innovation	Receives a value of 1 if the company introduced or improved products or processes in the year, otherwise 0.
Input variables		
Internet	Use of Internet	Receives a value of 1 if the company uses the internet in its business activities, otherwise 0.
web	Use of Website	Receives a value of 1 if the company uses a website in its business activities, otherwise 0.
Auto	Use of Automated Systems	Receives a value of 1 if the company uses automated systems in its business activities, otherwise 0.
soft	Use of Software	Receives a value of 1 if the company uses management software in its business activities, otherwise 0.
RD	Investment in R&D	Receives a value of 1 if the company invests in R&D, otherwise 0.

Source: Author's research

Steps to draw a decision tree with AI support

Step 1: Describe the problem and request ChatGPT or Gemini integration on
https://colab.research.google.com/ to write Python code.
Request to draw a decision tree based on data from a CSV file, with Tech as the output variable (with two states: innovation = 1, no innovation = 0). The input variables include the following columns:
(i) Internet (two states: used = 1, not used = 0);
(ii) web (two states: used = 1, not used = 0);
(iii) Auto (two states: used = 1, not used = 0);
(iv) soft (two states: used = 1, not used = 0);
(v) RD (two states: invested = 1, not invested = 0).

Step 2: Copy the code into the Colab window to run Python online in
https://colab.research.google.com/
Note: Upload the data file to the content section and copy the file link to paste into the code as instructed.

Step 3: Correct errors and formatting according to Gemini's instructions.
Correct errors until the output results meet the desired requirements.
You can adjust the tree format, select the tree depth, modify the image size, change the font size on the tree, and choose between the entropy or Gini criterion.
The code for drawing the decision tree is in Appendix 2.
Note: Copy the decision tree to Paint, then copy from Paint to Word.

Step 4: Guide on interpreting the results and making predictions.

V. RESEARCH RESULTS AND DISCUSSION

Discuss the output of the decision tree and the dependence of technological innovation (*Tech*) on input variables (*internet, soft, web, Auto, RD*)

Decision tree model:
Target variable (y): Tech (representing technological innovation, likely with two classes: "innovation", and "no innovation").
Input variables (X): internet, web, Auto, soft, RD. These represent various factors potentially related to technological innovation.
Criterion: Entropy (we used criterion='entropy').
Maximum depth: 3 (we used max depth=3).
Visualization: We have been using plot tree to visualize the resulting decision tree.
The decision tree describes the technological innovation status of Vietnamese enterprises based on R&D investment status and information technology application status in Figure 1.

Interpreting from the decision tree:

(1) Key drivers of innovation

The RD (Research & Development) variable is the most significant factor in determining technological innovation. Companies with higher *RD* investments (*RD > 0.5*) are more likely to be involved in innovation, reflecting the critical role of research and development in driving technological progress.

Auto (Automation) is another important predictor. Companies utilizing automation technologies (*Auto > 0.5*) are more inclined to engage in innovation, likely due to the efficiency, scalability, and new technological possibilities that automation offers.

Web and *Soft* are secondary factors. While they help refine predictions, they are less decisive than *RD* and *Auto*. For instance, companies with higher reliance on *Web* and *Soft* technologies are more likely to be innovative when combined with *Auto > 0.5*.

(2) Clear separation between innovation and non-innovation

The decision tree demonstrates a clear division between companies involved in innovation and those not. If *RD > 0.5*, the model strongly favors innovation, regardless of other factors. In contrast, if *RD ≤ 0.5,* companies are predominantly classified as non-innovative unless they exhibit significant use of automation or web technologies.

(3) Certainty and predictability

The entropy values across different decision nodes show a high degree of certainty in predictions at certain splits, particularly for *RD ≤ 0.5* and *Auto > 0.5*. The tree provides clear-cut decisions when *RD > 0.5*, with innovation being the consistent outcome. However, higher entropy values (such as 0.992) in certain splits, particularly when considering *Web ≤ 0.5*, suggest some uncertainty in predicting innovation, indicating that innovation can still occur in complex scenarios with less obvious indicators.

(4) Decision-making for business strategy

For businesses, this decision tree analysis can serve as a strategic tool for identifying key areas to focus on if they want to enhance or promote technological innovation. Investing in R&D and automation technologies appears pivotal in fostering innovation, while factors like *Web* and *Soft* technologies can provide additional support.

Additionally, this model suggests that companies with low *RD* investments should consider focusing on *Auto* or *Web*

Volume 9, Issue 2, pp. 83-89, 2025.
International Journal of Scientific Engineering and Science
ISSN (Online): 2456-7361

technologies to boost innovation, especially if they plan to stay competitive in rapidly evolving industries.

VI. Conclusions

This paper examines the application of Artificial Intelligence (AI) in Business Analytics (BA) through a case study focused on decision tree analysis of technological innovation. AI in BA presents significant opportunities such as improved predictive accuracy, real-time decision-making, and automation of routine tasks, all of which enhance efficiency and reduce costs. AI also enables personalized services at scale and uncovers hidden patterns in large datasets, providing businesses with a competitive edge. However, there are challenges to overcome, including the need for high-quality data, the complexity of AI implementation, and the risks of overfitting or bias in models. Integrating AI into existing systems can also be challenging, and ethical concerns about fairness and transparency remain. To maximize AI's potential, businesses must address these challenges with proper governance and continuous monitoring.

The study specifically explores how Research and Development (RD) investment and Information Technology applications impact a company's likelihood of engaging in technological innovation. The findings indicate that:

(i) *RD* Investment is the most significant driver of technological innovation. Companies with higher *RD* investments (*RD > 0.5*) are more likely to innovate.
(ii) Automation (*Auto*) plays a crucial role, with companies investing in automation technologies showing a higher likelihood of innovation.
(iii) While *Web* and *Soft* technologies offer additional insights, they are secondary to *RD* and *Auto* predicting innovation.
(iv) The decision tree model segments innovative and non-innovative companies, with *RD* and *Auto* being the key differentiators.
(iv) The model also highlights areas of uncertainty (e.g., Web ≤ 0.5), suggesting that innovation can still occur in complex situations where relationships are less clear.

Overall, the paper demonstrates how AI, particularly decision tree analysis, can help businesses identify the factors driving technological innovation. To foster innovation, companies should prioritize RD investments and automation technologies, while secondary IT applications like Web and Soft can further support these efforts. This study offers valuable insights for business leaders and decision-makers looking to leverage AI and data analytics for strategic planning and innovation.

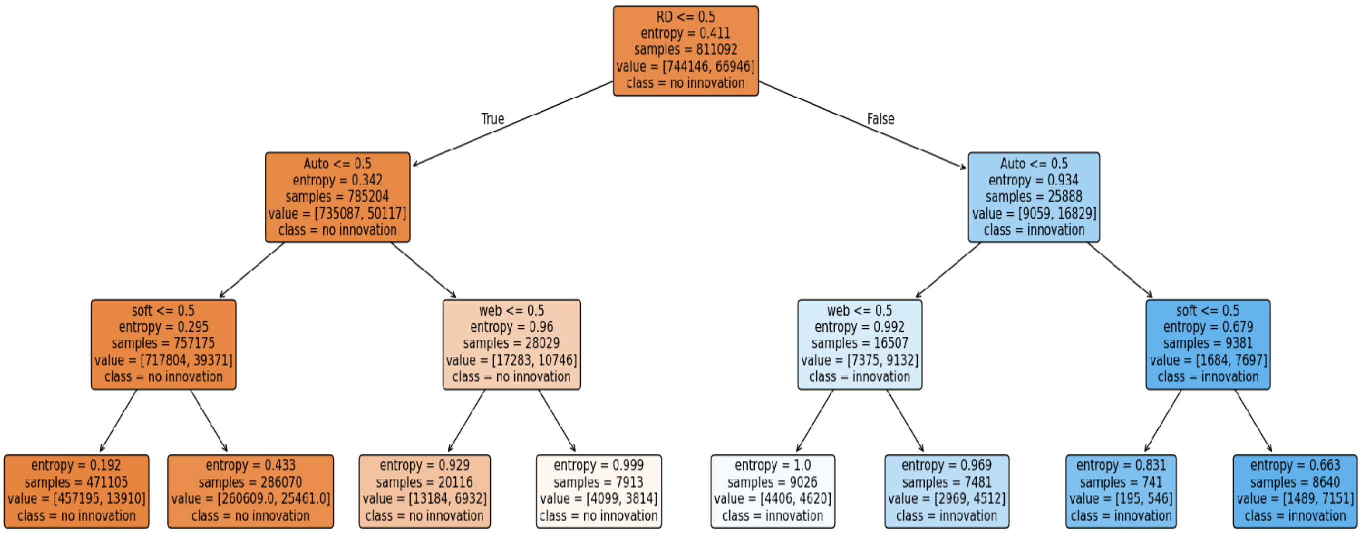

Fig. 1. Decision tree describing technological innovation based on R&D investment status and IT application
Source: Author's research results using Python

Appendix

Appendix 1. Steps to draw a decision tree in Python
Install Necessary Libraries: Ensure you have the required libraries installed (pandas, scikit-learn, and matplotlib).
Load Data: Load your CSV file into a pandas Data Frame.
Prepare Data:
Define the input features (X) and the target variable (y).
Make sure these variables are set to the correct columns.

Create and Train the Decision Tree:
Instantiate a DecisionTreeClassifier from scikit-learn.
Fit the model to your data.
Visualize the Tree:
Use plot_tree to draw the decision tree.
Customize the plot with feature names and class labels.

Appendix 2: Code in Python for decision tree

```
      # Install necessary libraries
  !pip install pandas scikit-learn matplotlib

# Import libraries
import pandas as pd
from sklearn.tree import
DecisionTreeClassifier, plot_tree
import matplotlib.pyplot as plt
```

```
import os
# Check if the file exists
file_path = '/content/Data DMST.csv'
if not os.path.exists(file_path):
    print(f"Error: File '{file_path}' not
found. Please check the file path.")
    exit()  # Exit if the file is not found
# Load the data
try:
    df = pd.read_csv(file_path)
except Exception as e:
    print(f"An error occurred while reading
the file: {e}")
    exit()
# Check the data
print(df.head())
# Define input features (X) and target
variable (y)
X = df[['internet', 'web', 'Auto', 'soft',
'RD']]
y = df['Tech']  # Assuming "Tech" column
represents Inno
# Create and train the decision tree model
with a maximum depth of 3
model =
DecisionTreeClassifier(criterion='entropy',
max_depth=3)
model.fit(X, y)
# Visualize the decision tree with larger font
size
plt.figure(figsize=(25,8))  # Adjust the image
size to be suitable for A4 paper
plot_tree(model, feature_names=['internet',
'web', 'Auto', 'soft', 'RD'],
            class_names=['no innovation',
'innovation'], filled=True, rounded=True,
fontsize=11) # Increased fontsize
plt.show()
```

REFERENCES

[1] Amarasinghe, P. T., Nguyen, S., Sun, Y., Alahakoon, D, “Ai-copilot for business optimisation: A framework and a case study in production scheduling”,. *arXiv preprint, 2023*. https://doi.org/10.48550/arXiv.2309.13218

[2] Biecek, P., & Burzykowski, T, “Explanatory Model Analysis: Explore, Explain, and Examine Predictive Models” (1st ed.). *Chapman and Hall/CRC, 2021*. https://doi.org/10.1201/9780429027192

[3] Brynjolfsson, E., & McAfee, A, “The second machine age: Work, progress, and prosperity in a time of brilliant technologies”, *WW Norton & company*, 2014.

[4] Davenport, T. H, “From analytics to artificial intelligence”, *Journal of Business Analytics*, 1(2), 73-80, 2018. https://doi.org/10.1080/2573234X.2018.1543535

[5] Davenport, T. H., & Harris, J. G, “*Competing on Analytics: The New Science of Winning”,* Harvard Business Press, 2007.

[6] Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P, “Generative AI: Transforming business analytics”, *Business & Information Systems Engineering*, 66(1), 111-126, 2024 https://doi.org/10.1007/s12599-023-00834-7

[7] Gandomi, A., & Haider, M, “Beyond the hype: Big data concepts, methods, and analytics”, *International Journal of Information Management*, Volume 35, Issue 2, April, Pages 137-144, 2015. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

[8] Hu, L., Chen, Z., & Zhang, X, “PyTorch Frame: A modular framework for multi-modal tabular data processing”, *Data Science and Analytics Journal*, 5(2), 33-48, 2024. https://doi.org/10.1016/j.dsaj.2024.05.007

[9] Kumar, R., & Singh, M, "AI-Based Predictive Analytics in Business: A Systematic Review", *Expert Systems with Applications, 213*, 119005, 2023.

[10] Liu, S., Zhao, S., Jia, C., Zhuang, X., Long, Z., & Lan, M, “BIBench: Benchmarking data analysis knowledge of large language models”, *arXiv Preprint*, arXiv:2401.02982, 2024. https://doi.org/10.48550/arXiv.2401.02982

[11] Quinlan, J. R, "Improved Use of Continuous Attributes in C4.5." *Journal of Artificial Intelligence Research*, 4, 77-90, 1996.

[12] Russell, S. J., & Norvig, P, “*Artificial Intelligence: A Modern Approach” (4th ed.)*. Pearson, 2021.

[13] Salazar, A., & Kunc, M, “The contribution of GenAI to business analytics”, *Journal of Business Analytics*, 1-14, 2025. https://doi.org/10.1080/2573234X.2024.2435835

[14] Wang, Y., Feng, D., Dai, Y., Chen, Z., Huang, J., Ananiadou, S., & Wang, H, “HARMONIC: Harnessing LLMs for tabular data synthesis and privacy protection”, *arXiv Preprint*, arXiv:2408.02, 2024. https://arxiv.org/abs/2408.02

[15] Witten, I. H., Frank, E., & Hall, M. A, “*Data Mining: Practical Machine Learning Tools and Techniques”* (3rd ed.). Morgan Kaufmann, Burlington, MA, 2011.

[16] Xing, F, “Designing heterogeneous LLM agents for financial sentiment analysis”, *arXiv Preprint*, arXiv:2401.05799, 2024. https://doi.org/10.1145/3688399