# An Adaptive Heterogeneous Ensemble Learning Method for Multi-dimensional Company Performance Decision-Making

Yi Feng[1,4], Mohammad Zoynul Abedin [2], Yunqiang Yin [3], Dujuan Wang[1,*], Edwin Tai Chiu Cheng[4], Kristof Coussement[5]

[1] Business School, Sichuan University, Chengdu 610064, China

[2] Department of Accounting and Finance, School of Management, Swansea University, SA1 8EN Swansea, UK

[3] School of Economics and Management, University of Electronic Science and Technology of China, Chengdu 610064, China

[4] Department of Logistics and Maritime Studies, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

[5] IÉSEG School of Management, LEM-CNRS UMR 9221, 3 Rue de la Digue, Lille, France

**Abstract:** Evaluating company performance is central to strategic decision-making and sustainable development. Given the limited development of comprehensive investigations into the impact of multidimensional variables on company performance within existing research, we pioneer the use of machine learning methods to explore this issue. To accurately predict company performance in terms of operating income, net profit, and total assets, we propose a novel adaptive heterogeneous ensemble learning (AHEL) method that adaptively outputs the heterogeneous ensemble learning model with the best predictive performance for each predicted aspect. Experimental results on real-life data from 740 Chinese listed companies over the period 2010-2020 demonstrate that AHEL outperforms several state-of-the-art machine learning methods for multi-dimensional company performance prediction and leads to better organizational decisions. We also examine the relative importance of the features of the predicted aspect and interpret the correlations between the important feature values and the prediction results of AHEL. The findings reveal that the features

[*]Corresponding author: Dujuan Wang. E-mail address: djwang@scu.edu.cn (D. WANG)

‘social responsibility’, ‘shareholder responsibility’ and ‘R&D expenditure’ all positively impact the predicted results, ‘government subsidy’ has a threshold effect on the predicted results, and ‘digital transformation’ and ‘innovation’ have a mixed impact on the predicted results. These prescriptive insights enhance researchers’ understanding of multi-dimensional company performance prediction that benefits their future work.

**Keywords:** heterogeneous ensemble learning; company performance; model interpretability; machine learning

## 1. Introduction

The main objective of a business company is to earn profits and increase shareholders' wealth. Company performance has played a crucial role in long-term industrial upgrading and economic growth for over two decades (Katila et al. 2012). In pursuit of this objective, it is important to investigate the variables that influence company performance, which is a major concern for companies and researchers from theoretical and practical perspectives (Busch and Friede 2018). From the company's standpoint, given that companies have limited resources and information, understanding the variables that influence company performance allows them to make informed strategic decisions, allocate resources efficiently, and enhance company performance (Truant et al. 2021). Researchers contribute by providing empirical evidence and prediction methods that advance theory development and support practice (Ben Lahouel et al. 2022).

Identifying what variables influence business performance can help a company better adapt to market changes and technological advances, and there is a large body of literature on the variables influencing company performance. Previous studies based on econometric methods have demonstrated that corporate social responsibility (CSR) (Horváthová 2011; Saeidi et al. 2015; Ben Lahouel et al. 2022), research and development (R&D) expenditure ( Sueyoshi and Goto 2009; James and McGuire 2016; Alam et al. 2020), government subsidy (Zhang et al. 2014; Chu et al. 2017; Wang et al. 2021), and innovation (Gopalakrishnan 2000; Huang and Huarng 2015; Aastvedt et al. 2021) are important determinants of company performance. Recently, the rapid growth of digital technologies, such as artificial intelligence, big data, cloud computing, cryptocurrencies, and blockchain, has significantly transformed companies in terms of strategy, structure, and process (Chouaibi et al. 2022; Truant et al. 2021; Wamba et al. 2017). An increasing number of CEOs are trying to learn how digitalization influences their business decision-making (Chouaibi et al. 2022; Truant et al. 2021; Wamba et al. 2017). Several researchers have studied the influence of digitalization on company performance using econometric methods (Truant et al. 2021).

To the best of our knowledge, few empirically estimated models simultaneously consider the effects of the aforementioned variables together on company performance. Furthermore, the vast majority of past studies use multiple linear regression for analysis, yielding different results on the importance and impacts of the investigated variables (Ben Lahouel et al. 2022; Franceschelli et al. 2019).

There are often non-linear behaviours between the investigated variables and company performance due to the complexity of company performance and the non-linear threshold effect (Yang et al. 2019; Ben Lahouel et al. 2022). So the application of the linear model on company performance might impede theory development and not allow the cohabitation of contradictory theories (Ben Lahouel et al. 2022). Furthermore, when the investigated variables are associated with company performance in a non-linear way, estimates from linear models can lead to misleading interpretations and flawed management suggestions (Ben Lahouel et al. 2022; Latan et al. 2018), which are difficult to provide decision support for managers. Applying new alternative methods to explore complex data and overcome the limitations of traditional models is worth exploring.

Machine learning, an essential part of Industry 4.0, provides a means to deal with complex problems, especially non-linear problems, by utilizing big data analytics and algorithms to build predictive models based on historical datasets of related problems (Jabeur et al. 2021; Wu et al. 2022; Yu et al. 2022). The predictive model makes predictions based on future data. Currently, machine learning has been shown to improve decision-making in company management (Cielen 2004; Olson et al. 2012; Jabeur et al. 2021; Yu et al. 2022). Despite this, few studies have developed machine learning models for company performance prediction. Furthermore, machine learning techniques that lack interpretability and understandability are criticized as black-box models (Jabeur et al. 2021; Coussement and Benoit 2021; Wang et al. 2022), which hinders the application of machine learning techniques. In practice, executives generally not only wonder about the key predictors and predictive performance of the optimal model, but also pay attention to the impact of the key predictors on the optimal model's predictive results. In this way, executives can better trust the model's predictions and make informed decisions based on the interpretable information. However, the interpretability of predicted results given by machine learning models is under-studied.

Setting out to fill this gap, this study proposes a novel adaptive heterogeneous ensemble learning (AHEL) method for multi-dimensional company performance decision-making. AHEL adaptively outputs and interprets the best heterogeneous ensemble learning model tailored to each company performance metric. Using a dataset from 740 Chinese listed companies spanning 2010 to 2020, our experiments demonstrate the adaptability and superior performance, and interpretability of AHEL. And our study contributes to machine learning-driven company management in three ways:

1) Our work extends the scope of the existing literature on company management driven by machine learning. Most of the current studies rely on linear regression to examine the impact of a single variable on company performance (Ben Lahouel et al. 2022; Bond and Guceri 2018; Busch and Friede 2018; Chouaibi et al. 2022; Truant et al. 2021; Wamba et al. 2017; Wang et al. 2017). However, the combined effects of multiple variables and company performance remain underexplored, and linear regression is also limited in capturing these complex relationships. Our work is a pioneering study on multi-dimensional company performance prediction, introducing a novel machine learning method that offers valuable insights into developing multi-dimensional decision-making strategies for company performance.
2) Our research advances the development of heterogeneous ensemble learning methods on multi-dimensional outputs. Currently, several excellent heterogeneous ensemble learning methods have been proposed, but their limitation is that these methods can only achieve optimal performance on one-dimensional outcome variables (Papouskova and Hajek, 2019, Cui et al. 2021, Wang et al. 2022, Hou et al. 2023). The novelty of our proposed AHEL method lies in its ability to adaptively output the heterogeneous ensemble learning model that performs best for multi-dimensional outcome variables, enabling accurate multi-dimensional predictions. And we provide robust evidence that AHEL performs better than several state-of-the-art machine learning methods on

multi-dimensional outcome variables.

3) Our work provides valuable insights into the important features of the proposed AHEL method for multi-dimensional company performance prediction, interpreting how these feature values influence the AHEL's predicted outcomes through interpretable data science methods (Coussement and Benoit 2021). This sets our study apart from previous research, which has predominantly focused on the predictive performance of models (Papouskova and Hajek, 2019; Baradaran et al. 2022; Hou et al. 2023). Our approach provides interpretable insights at both global and local levels, providing a comprehensive understanding of how multi-dimensional variables impact various aspects of company performance prediction. These understandings also build decision-makers' confidence in applying the AHEL method and support more informed decision-making tailored to their companies' specific circumstances.

We organize the rest of the paper as follows: In Section 2 we briefly review the related literature on company performance and the ensemble learning method. In Section 3 we present the data and feature engineering, propose the AHEL method, and discuss the experimental design. We analyze the results of statistical descriptions, model performance evaluation, and model interpretability in Section 4. In Section 5 we discuss the predictive results and performance of AHEL's predictions of a company's operating income, net profit, and total assets from theoretical and managerial perspectives. Finally, in Section 6, we conclude the paper and suggest future research directions.

## 2. Literature review

### 2.1 Company performance

The impacts of CSR (Ben Lahouel et al. 2022; Dixon-Fowler et al. 2013; Endrikat et al. 2014; Saeidi et al. 2015), digital transformation (Chouaibi et al. 2022; Truant et al. 2021; Wamba et al. 2017), R&D expenditure (Alam et al. 2020; Bond and Guceri 2018), government subsidies (Luo et al. 2021; Wang et al. 2021), and innovation (Huang and Huarng 2015; Aastvedt et al. 2021) on company performance is an important research topic. We summarize in Table 1 the major findings in the related literature.

First, most of the studies (shown in bold in Table 1) focus on investigating the influence of a single variable on company performance, and few studies consider the impacts of multiple variables on company performance from the macro perspective. Second, most of the extant literature performs quantitative analyses using multiple linear regression (Aastvedt et al. 2021; Chouaibi et al. 2022; Chu et al. 2017; Horváthová 2012; Ehie and Olibe 2010; Franceschelli et al. 2019; Gonenc and Scholtens 2017; Huang and Huarng 2015; James and McGuire 2016; Luo et al. 2021; Wang et al. 2017; Zhou et al. 2021). However, it is difficult for linear models to truly reflect the non-linear relationships between the variables (e.g., CSR, digital transformation, etc.) and company performance, which causes problems such as poor predictive performance ( Ben Lahouel et al. 2022). Despite existing studies that have found that machine learning methods can achieve more accurate predictions in business and corporate management (Jabeur et al. 2021; Wang et al. 2022; Wu et al. 2022), research on company performance prediction using machine learning methods has not been carried out. Finally, the impacts of CSR, digital transformation, R&D expenditure, government subsidy, and innovation on company performance are still considered ambiguous, mixed, and contradictory based on the findings of past studies.

**Table 1.** Findings on the impacts of CSR, digital transformation, R&D expenditure, government subsidy, and innovation on company performance.

| Aspect | Finding | Literature |
|---|---|---|
| Corporate social responsibility | Positive influence | ***Busch and Friedel*** (2018) [2,3]; ***Dixon-Fowler et al.*** (2013) [2,3]; ***Endrikat et al.*** (2014) [2,3]; **Horváthová** (2011) [s]; **Franceschelli et al.** (2019) [2,3]; **Saeidi et al.** (2015) [1, 2,3]; Velte (2017) |
| | Negative influence | **Gonenc and Scholtens** (2017) [s]; **Horváthová** (2012) [2,3] |
| | Non-linear influence | **Ben Lahouel et al.** (2022) [s] |
| | Neutral influence | **Becchetti and Ciciretti** (1978) [s]; **Nelling and Webb** (2009) [2,3] |
| Digital transformation | Positive influence | **Truant et al.** (2021) [s]; Ribeiro-Navarrete et al. (2021) [1,2,3]; **Chouaibi et al.** (2022) [s] |
| | Negative influence | **Wamba et al.** (2017) [s] |
| | Non-linear influence | Kohtamki et al. (2020) [2,3] |
| | Neutral influence | Zhou et al. (2021) [2,3] |
| R&D expenditure | Positive influence | Ehie and Olibe (2010)[1]; Wang et al. (2017)[2]; Patel et al. (2018)[1]; James and McGuire (2016)[2,3]; Sueyoshi and Goto (2009) [1] |
| | Negative influence | Sueyoshi and Goto (2009)[1] |
| | Neutral influence | **Bond and Guceri** (2018) [1]; Alam et al. (2020) [2,3] |
| Government subsidy | Positive influence | Luo et al. (2021) [2,3]; **Wang et al.** (2021) [2,3] |

| | | |
|---|---|---|
| | Negative influence | **Chu et al.** (2017) [2,3] |
| | Neutral influence | Zhang et al. (2014) [s] |
| Innovation | Positive influence | Yang et al. (2022) [s]; Huang and Huanrng (2015)[1]; **Aastvedt et al.** (2021) [2,3]; Luo et al. (2021) [s]; Liao (2018)[2,3] |
| | Negative influence | **Albuquenrquer et al.** (2018)[1,2,3] |
| | Non-linear influence | **Hatzikian (2013)**[1] |

Note: The superscripts [1], [2], [3,] and [s] represent the company performance and are defined as operating income, net profit, total assets, and scale measurement respectively; italics denote that the study is a review or meta-analysis.

## 2.2 Ensemble learning methods

Ensemble learning methods, a branch of machine learning methods, are the most popular prediction methods in recent years (Baradaran et al. 2022). Compared with traditional machine learning methods that train only one learner, such as support vector regression (SVR) (Cortes and Vapnik 1995) and decision tree (DT), ensemble learning integrates multiple base learners through certain integration strategies to achieve better predictive effects (Jabeur et al. 2021). Ensemble learning can be further divided into homogeneous and heterogeneous ensemble learning methods according to whether the base learner adopts the same type of learning algorithm (Papouskova and Hajek 2019; Wang et al. 2022). Figure 1 presents the differences in structure of the homogeneous and heterogeneous ensemble learning models.

For the homogeneous ensemble learning model, the base learners consist of the same algorithm, and for the heterogeneous ensemble learning model, the base learners consist of different algorithms. Random forest (RF), extreme gradient boosting (XGBoost) (Chen and Gustrin, 2016), and gradient boosting decision tree (GBDT) are typical homogeneous ensemble learning methods as their base learners all consist of decision trees. Heterogeneous ensemble learning methods allow different types of algorithms as the base learners, so their structures are very flexible. Currently, numerous studies on company management have applied homogeneous ensemble learning methods to predict problems such as corporate credit scores and corporate bankruptcies. For example, Jabeur et al. (2021) proposed a novel approach for classifying categorical data using gradient boosting and categorical features (CatBoost) for corporate failure predictions. The experimental results demonstrate that CatBoost effectively increases the power of classification performance compared with other advanced classifications. Yu et al. (2022)

used various machine learning methods, including classification and regression trees, artificial neural networks, RF, and support vector machine classifiers, to predict the credit ratings of Eurozone firms. The results show that RF can be used together with regression trees to predict near-default ratings.

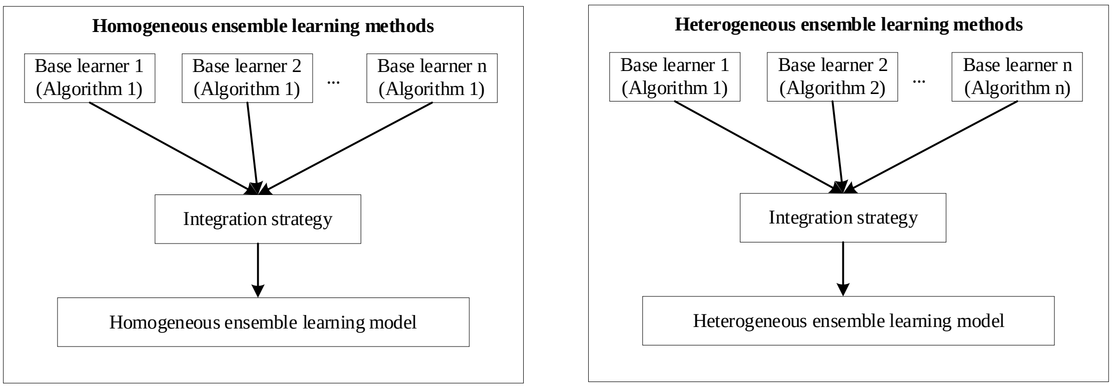



**Figure 1.** The structures of homogenous ensemble learning methods and heterogeneous ensemble learning methods.

Furthermore, several theoretical studies have suggested that any successful ensemble learning model is closely associated with the diversity of integration members and strategies (Papouskova and Hajek 2019; Hou et al. 2023). Diversity refers to the fact that the errors made in the predictions of each base learner are not coincident, whereas the integration strategy aggregates the predicted results of multiple base learners into one final predicted result. For the homogeneous ensemble learning approach, the base learners are formed by one type of algorithm, resulting in relatively poor diversity among the base learners and hence a high degree of coincidence in the errors they predict (Papouskova and Hajek 2019; Hou et al. 2023). Contrarily, heterogeneous ensemble learning methods can increase the diversity between base learners by using different types of algorithms as base learners. This is because the principles of different types of algorithms are quite different, which significantly reduces the possibility of coincidence in the errors they predict.

In summary, despite extensive research on company performance and the development of ensemble learning methods, several gaps persist. First, existing research focuses on the influence of a single variable on company performance (Ben Lahouel et al. 2022; Bond and Guceri 2018; Busch and Friede 2018; Chouaibi et al. 2022; Gonenc and Scholtens 2017), overlooking the combined effects of multiple variables. Our study fills this gap by generating and incorporating five-dimensional variables derived from diverse data sources, moving beyond the single-variable focus of prior research. Second, prior studies primarily rely on linear regression to examine company performance issues, which limits their ability to capture the complex relationships between variables (Aastvedt et al. 2021; Chouaibi et al. 2022). The proposed AHEL method addresses this limitation, which utilizes heterogeneous ensemble learning method methods to model these non-linear relationships more effectively. Finally, heterogeneous ensemble learning methods have demonstrated the potential in enhancing the predictive accuracy of single-dimensional prediction in various fields (Wang et al. 2022; Wu et al. 2022). However, there is a gap in their development to company management, particularly in predicting and interpreting multi-dimensional company performance. Our study develops the AHEL method specifically designed for this purpose. We demonstrate that AHEL can adaptively output the best model for each predicted aspect. All in all, AHEL not only improves prediction accuracy but also enhances interpretability, leading to better organizational decisions across multiple dimensions of company performance.

## 3. Methodology

### 3.1 Data and feature engineering

To improve the predictive performance and interpretability of the model, we collect research data based on three criteria of data availability, data integrity, and research relevance. First, the collected data from various open websites and available sources (e.g., Hexun website, China Stock Market & Accounting Research (CSMAR), Wind Dataset, Chinese Research Data Services Platform (CNRDS), etc.) on Chinese listed companies. Second, companies with financial or other anomalies, and companies that were delisted during that period are excluded following Cui et al. (2019) and Wang et al. (2021). Finally, as summarized in Table 1, we first consider the information of Chinese listed companies in terms of CSR, digital transformation, R&D expenditure, government subsidy, and innovation. Furthermore, the company's basic information, such as the company age, the age of listing, etc., were considered in the model construction referring to prior literature (Hu et al. 2023; Wang et al. 2021).

Accordingly, all information on Chinese listed companies includes six aspects: CSR, digital transformation, expenditure and subsidy, innovation, basic information, and company performance, and we obtain all Chinese listed companies that publicly and fully disclosed the above six aspects from 2010 to 2020. Finally, 740 Chinese companies listed on the Shanghai and Shenzhen Stock Exchanges (commonly known as A-share listed firms in China) from 2010 to 2020 satisfy the above criteria and are taken as the research data.

**Table 2.** The origins and definitions of features.

| Aspect | Feature | Source | Definition |
|---|---|---|---|
| Corporate social responsibility information | Shareholder responsibility | Hexun website | Measured by profit score, debt service score, reward score, credit approval score and innovation score |
| | Environmental responsibility | | Environment score |
| | Social responsibility | | Contribution value score |
| Digital transformation information | AI technology | The annual report of the listed company | The level of AI technology |
| | Block chain technology | | The level of block chain technology |
| | Cloud computing technology | | The level of cloud computing technology |
| | Big data technology | | The level of big data technology |
| | Extended application of digital technology | | The level of extended application of digital technology |
| Expenditure and subsidy information | R&D expenditure | Hexun website, CSMAR dataset | Research and development expenditures |
| | Environmental investment | Additional information in the annual reports of the listed company | Environmental investment |
| | Government subsidy | Wind dataset | Government subsidy |
| Innovation information | Green patent application | CSMAR dataset, CNRDS dataset, Innovation patent database | Number of green patent applications |
| | Green patent grant | | Number of green patents grant |
| | Other patent application | | Number of other types of patent applications (excluding green patents) |
| Basic information | Heavy polluting company | CSMAR dataset | Whether it is a heavily polluting company |
| | Number of directors | CSMAR dataset | Number of directors |
| | Number of independent directors | CSMAR dataset | Number of independent directors |

| | | | |
|---|---|---|---|
| | Nature of shareholding | CSMAR dataset | Nature of shareholding （Local state-owned enterprises, public enterprises, collective enterprises, private enterprises, other enterprises, foreign-funded enterprises, central state-owned 8140enterprises） |
| | Company age | CSMAR dataset | Company age （Calculated from the year of company registration） |
| | Listing age | CSMAR dataset | Age of listing （Calculated from the year the company was listed） |
| | SOE | CSMAR dataset | Whether the company is a state-ownership status for the top one shareholder of the listed companies |
| | Big4 | CSMAR dataset | Audited by the Big Four (PwC, Deloitte, KPMG, Ernst & Young) |
| Company performance | Operating income | Wind dataset, Hexun website | Operating income |
| | Net profit | CSMAR dataset | Net profit |
| | Total assets | CSMAR dataset | Total assets |

Table 2 lists the sources and definitions of the relevant features in the six aspects of Chinese listed companies covered in this study, and Table A1 in the appendix shows the descriptions of all features. Specifically:

*Corporate social responsibility information*. As reviewed in Section 2.1, CSR has now emerged as a force for companies that can have a mixed influence on their businesses by participating in CSR activities. China has stated that the listed companies, the most important part of the national economy, should actively fulfill their social responsibility and become exemplary, which not only meets the expectations and requirements of stakeholders but also improves the efficiency and efficacy of management, cultivates a competitive advantage, and establishes a good image, thus achieving sustainable development (Busch and Friede 2018; Dixon-Fowler et al. 2013; Endrikat et al. 2014; Franceschelli et al. 2019; Saeidi et al. 2015). The CSR scores of all the A-share listed firms in China are published on the Hexun website, a third-party rating system in China that has provided technical information about CSR since 2010. We collected data on the CSR performance of Chinese listed companies on a third-party rating system in China, i.e., the Hexun website. The definitions of each dimension are detailed in Table 2.

*Digital transformation information*. The rapid development of digital technologies, such as artificial intelligence, machine learning, big data, cloud computing, cryptocurrencies, and blockchain, has led to an increasing number of businesses facing digital transformation issues and attempting to take initiatives to explore new digital technologies. Despite the many challenges companies face in digital transformation, such as integrating and utilizing new digital technologies, they still invest heavily in digital transformation to ensure survival and sustainable development (Chouaibi et al. 2022; Ribeiro-Navarrete et al. 2021; Truant et al. 2021; Wamba et al. 2017). In our study, a company's digital transformation is measured in terms of five technologies: AI, blockchain, cloud computing, big data, and digital technology application referring to Hu et al. (2023). Specifically, the annual reports of Chinese listed companies were collected using a crawler, and Java PDFbox was used to extract the textual content. Next, the relevant feature words for each of the five techniques were summarized based on Hu et al. (2023), with each technique including multiple feature words, presented in Table A2 in the Appendix. Feature word search and frequency statistics calculations were conducted on the extracted text content according to Table A2. Finally, the frequency of feature words in the annual reports of listed companies for each technology was taken as the degree of transformation of that technology.

*Expenditure and subsidy information.* This part mainly relates to expenditure on R&D and financial subsidies received from the Chinese government, as well as environmental protection by Chinese listed companies. First, R&D expenditure has always been considered an essential component of economic growth (Alam et al. 2020). At the company level, R&D expenditure is an important determinant of differences in profitability because companies may face a barrier or may also be on the rise when they increase their investment in R&D.

Second, government and public policies generally support companies' sustainable development through financial subsidies. Government subsidy plays an important role in the performance of companies in risky environments, such as economic difficulties, new technologies, and emerging industries. Differences in the form and amount of subsidy will result in a more significant gap in the financial performance of listed companies.

Third, environmental investment refers to a certain amount of funds paid by companies for pollution prevention, protection, and improvement of the ecological environment, as well as the activities associated with it, intending to promote the coordinated development of economic construction and environmental protection (McWilliams and Siegel 2001), which can lead to a loss of economies of scale and influence the company's operations. Consequently, R&D expenditure, government subsidy, and environmental investment should all be considered in the construction of the company's business performance predictive model.

*Innovation information*. Innovation, a dominant concept in business, is the primary competitive lever for improving economic performance (Huang and Huarng 2015). Patent data remain one of the most widely used and robust measures of a company's innovation performance in existing studies (Aastvedt et al. 2021; Huang and Huarng 2015; Liao 2018; Luo et al. 2021; Yang et al. 2022). Accordingly, we consider the number of green patent applications, the number of green patents granted, and the number of applications for other types of patents (except green patents) to measure the company's innovation performance. Meanwhile, it is worth noting that the patents granted by the Chinese State Intellectual Property Office include invention and utility patents. Therefore, we combine the number of applications and grants for both types of patents and consider the total number of patent applications in the model construction.

*Company performance*. Prior research has identified company business performance by measuring the overall level of the organization, using common measures such as return on assets (ROA), return on equity (ROE), total shareholder return rate, and sales growth rate (Nirino et al. 2021; Zhou et al. 2021). Many studies have measured the business performance of companies in a more intuitive way using accounting metrics, including operating income, net profit, and operating profit (Cui et al. 2019; Huang and Huarng 2015). To reflect the prediction model's results more intuitively on the business performance of listed companies, we consider the company's accounting net profit and total assets as measures of company business performance, considering that the definition of ROA is the ratio of net profit to total assets. In addition, operating income is considered to further measure the level of operations to reflect company performance. Thus, operating income, net profit, and total assets are labels for the predictive models.

*Company's base information*. We also take into account the same basic information on Chinese listed companies by referring to existing studies (Aastvedt et al. 2021; Bond and Guceri 2018; Busch and Friede 2018; Chu et al. 2017; Huang and Huarng 2015; Saeidi et al. 2015; Truant et al. 2021; Wamba et al. 2017), including "heavy polluting company", "number of directors", "number of independent directors", "nature of shareholding", "number of independent directors", "level", "company age", "listing age", "SOE", "Big4", "size". Table 2 provides a detailed definition of each feature.

**3.2 Adaptive heterogeneous ensemble learning method**

We aim to make regression predictions of the multi-dimensional performance of listed Chinese companies, including operating income, net profit, and total assets. Given the existence of multiple prediction tasks within this study, it is difficult to achieve optimal predictive performance across various prediction tasks using the same ensemble learning model. Consequently, the most effective configuration of the heterogeneous ensemble learning model (comprising base learners and strategies) could vary across the distinct prediction tasks.

To accommodate this variability and secure the optimal models for each task, we develop an adaptive heterogeneous ensemble learning (AHEL) method for multi-dimensional company performance, which focuses on reducing the errors between the predicted and actual results. Finally, AHEL can automatically output three heterogeneous ensemble learning models with the best predictive performance for three company performance predictions. Furthermore, to visualize and interpret the predictions of the model, we extract the important predictors and capture the trend of change between predictors and predicted responses of AHEL to support executive decisions drawing on interpretable technologies. Fundamentally, building a superior AHEL method requires consideration of five aspects, namely generation of the base learner pool, choice of integration strategy, evaluation of model performance, the output of the best heterogeneous ensemble learning models, and *post hoc* interpretation for the best heterogeneous ensemble learning models. Figure 2 presents the framework of AHEL, and we discuss the five key parts of the proposed AHEL method in detail below.

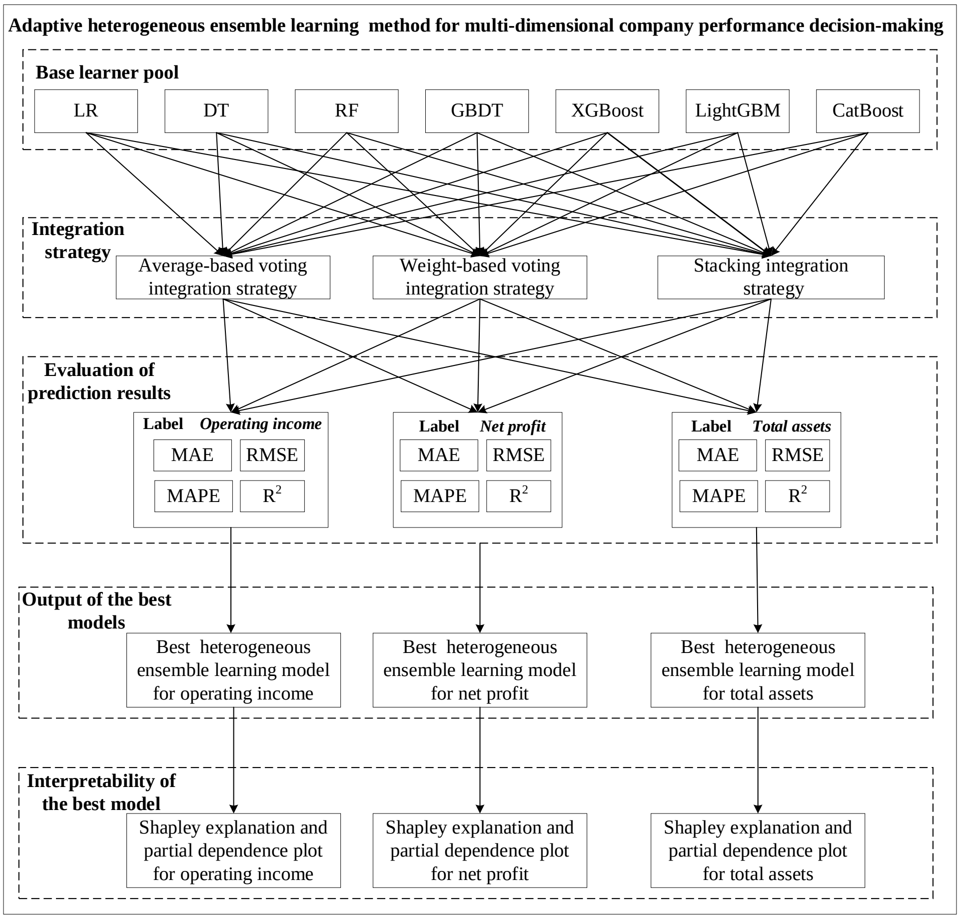



**Figure 2.** The framework of the proposed method.

### 3.2.1 Base learner pool

Considering that the fusion of heterogeneous base learners can achieve better generalization performance, we adopt seven different algorithms, namely linear regression (LR), DT (Breiman 2001), RF (Breiman 2001), GBDT (Friedman 2001), XGBoost (Chen and Guestrin 2016), light gradient boosting machine (LightGBM) (Ke et al. 2017), and CatBoost (Prokhorenkova et al. 2018) to generate the base learner pool. Our three reasons for choosing the above algorithms for base learners are as follows: First, these algorithms are popular for their outstanding performance in the corporate management field (Truant et al. 2021; Jabeur et al. 2021; Jeon et al. 2020). Second, the diversity of base learners in the base learner pool forms the basis for building a high-performance ensemble learning model. Theoretical differentiation among heterogeneous base learners increases their diversity, which allows for better generalization of unseen observations (Belhadi et al. 2021). Finally, we did not consider more complex algorithms, such as neural networks, due to their typically higher computational costs compared to the aforementioned algorithms. We are keen to reduce the computational burden and improve the efficiency of the AHEL, thus we did not consider it. We next introduce each algorithm in the base learner pool as follows:

*Linear regression.* LR is widely used to investigate variables that influence company performance (Kohtamki et al. 2020; Truant et al. 2021; Zhou et al. 2021). It aims to identify the relationships between a set of features and labels using the ordinary least squares (OLS) method.

*Decision tree.* DT is a classic supervised machine learning method with a tree structure, which is popular in business performance prediction owing to its ease of use and speed (Wang et al. 2021), and we use the class and regress tree in our study.

*Random forest.* RF is a popular ensemble learning method that achieves excellent performance for business and company management (Jabeur et al. 2021; Belhadi et al. 2021). RF is composed of multiple DTs, where DT is the base learner. Because RF is composed of multiple DTs, it has good interpretability.

*Gradient boosting decision tree.* GBDT, a boosting ensemble method, is one of the most popular ensemble learning models used in many empirical studies (Feng et al. 2022; Baradara et al. 2022). Unlike bagging ensemble learning techniques, such as RF, the final predictive result is computed in GBDT in a typical forward-stage fashion, where the newly generated base learner should be maximally correlated with the negative gradient of the loss function. In each iteration, GBDT seeks to construct a new regression tree to reduce the errors generated in the previous iteration. GBDT allows for not only flexible handling of a wide range of data but also higher prediction accuracy with relatively short parameter adjustment times.

*Extreme gradient boosting.* Traditional machine learning methods, such as DT and RF, have dominated the finance and company management fields for a long time. An advanced ensemble learning method, XGBoost, has recently received significant attention in business (Jeon et al. 2020). The basic idea of XGBoost is to develop a new decision tree in the gradient direction of the residuals to minimize the loss function. XGBoost supports both row and column sampling, introduces a second-order Taylor expansion for the loss function, and uses second-order partial derivatives in training, making XGBoost converge faster.

*Light Gradient Boosting Machine*. LightGBM has drawn the attention of researchers in finance and business because it performs accurately in practical regression tasks (Jabeur et al. 2021). LightGBM improves the two problems of XGBoost. First, XGBoost needs to traverse all leaf nodes, which leads to the problem of high cost and low efficiency whereas XGBoost performs information gain calculations. To solve this problem, LightGBM uses gradient-based one-side sampling (GOSS) to eliminate the step of traversing all the leaf nodes during each iteration. Second, XGBoost traverses all the samples during each iteration, making it memory- and time-consuming. To solve this problem, LightGBM reduces the calculation complexity and calculation cost through the exclusive feature bundling (EFB) strategy. Research shows that it is significantly better than XGBoost in terms of performance, efficiency, and running speed.

*Gradient boosting and categorical features*. Recently, Catboost has been successful for company management (Jabeur et al. 2021). CatBoost is a variation of the gradient boosting algorithm. Essentially, CatBoost has three main advantages. First, CatBoost effectively solves the problem of target leakage using ordered boosting. Second, CatBoost can handle classification features, making it possible to train and test various data types and formats. Third, CatBoost overcomes overfitting caused by traditional gradient boosting algorithms by performing random permutations in the selection of tree structures to estimate leaf values.

### 3.2.2 Integration strategy

The choice of the integration strategy is an integral part for the ensemble learning model. In this study, we consider three advanced integration strategies in AHEL, namely average-based voting, weight-based voting, and stacking integration. We discuss the principles of each integration strategy below.

*Average-based voting integration strategy.* Let a base learner be $b_t(x) \in B$ , $BL = \{b_t(x): t = LR, DT, RF, GBDT, XGBoost, LightGBM, Catboost\}$, and $BL$ denotes the base learner pool. The idea behind the average-based voting integration strategy is to integrate all the heterogeneous base learners in $BL$ and return an average predictive value, which is useful for a set of equally well-performing models to balance their respective weaknesses. Finally, an enhanced composite regression $H(x)$ is generated, and the formula for the average-based voting integration strategy is as follows:

$$H(x) = F\left[\sum_{t=1}^{T} \frac{1}{T} b_t(x)\right], \tag{1}$$

where $T$ is the number of base learners in $BL$.

*Weighted-based voting integration strategy.* Average-based voting integration strategy considers each base learner's decision equally and ignores the impacts of base learners with relatively poor performance. Thus, an average-based voting integration strategy may limit the effectiveness of the integration of the base learners. To overcome this shortcoming, we propose a weight-based voting integration strategy and embed it in AHEL. The weight-based voting integration strategy considers the $RMSE$ value of the base learners, and the base learners with lower $RMSE$ are assigned higher weights, and vice versa. The formula for the weight-based voting integration strategy is as follows:

$$H(x) = F[\sum_{t=1}^{T} w_t\, b_t(x)], \tag{2}$$

where $T$ is the number of base learners in $BL$, $w_t$ is the weight of the $t$th base learner, which can be calculated as $w_t = \frac{1/RMSE_{b_t(x)}}{\sum_{t=1}^{T} 1/RMSE_{b_t(x)}}, b_t(x) \in BL$, and $\sum_{t=1}^{T} w_t = 1$.

*Stacking integration strategy:* Stacking integration strategy, a novel integration strategy in ensemble learning methods, stacks the predictions of each base learner together and uses them as input to a final meta-learner to compute the final prediction. The stacking integration strategy includes two stages, the first of which includes a variety of base learners and the second consists of a meta-learner. All the base learners are trained and tested on the corresponding set in the first stage, and the meta-learner is trained based on the outputs of all base learners in the second stage, and makes predictions in the test set. In our study, we consider heterogeneous base learners in the first stage and chose linear regression as the meta-learner in the second stage.

**3.2.3 Model evaluation**

To comprehensively evaluate the predictive performance of the adaptive heterogeneous ensemble learning method and other advanced machine learning methods for company performance prediction, we employ four machine learning metrics, namely the mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and $R$-squared ($R^2$), which are computed, respectively, as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|, \tag{3}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}, \tag{4}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\left(\frac{y_i-\hat{y}_i}{y_i}\right)\right| \times 100\%, \tag{5}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N-1}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{N-1}(y_i-\bar{y}_i)^2}, \tag{6}$$

where $\hat{y}_i$ is the predicted value of the $i$th sample, $y_i$ is the actual value of the $i$th sample, and $N$ is the number of samples.

**3.2.4 Output of the best heterogeneous ensemble learning models**

AHEL considers seven heterogeneous base learners and three different integration strategies. Thus, there are rich heterogeneous ensemble learning models in AHEL, consisting of different base learner pools and integration strategies. Depending on the required prediction labels, AHEL incorporates the objective by maximizing $R^2$ in the ensemble selection stage under each integration strategy.

To facilitate the construction of a heterogeneous ensemble learning model, AHE utilizes recursion to increase the number of various candidate base learners in heterogeneous ensemble models, and the pseudocode of the best heterogeneous ensemble learning acquisition for each integration strategy is shown in Table 3. AHEL first trains and assesses the best single candidate model from the base learner pool as the original model. On this basis, a new base learner is continuously added from the base learner pool to improve the performance of the heterogeneous ensemble learning model. Subsequently, AHEL compares the performance of the best heterogeneous ensemble learning models of the three integration strategies and outputs the heterogeneous ensemble learning model with the best performance as the best predictive model for that label. Accordingly, the heterogeneous ensemble learning model with the best performance can be obtained in AHEL, including the best base learner pool and integration strategy information for each predictive task. In this study, we have three predictive labels, namely operating income, net profit, and total assets. Hence, the heterogeneous ensemble learning model with the best predictive performance for each predictive label is output, supporting the subsequent *post hoc* interpretation of the best heterogeneous ensemble learning model for each label.

**Table 3.** Pseudocode of the best heterogeneous ensemble learning acquisition.

**Input:** the original model and the evaluation value; the base learner pool; integration strategy
**Output:** the heterogeneous ensemble learning model with the best performance for the integration strategy
**For** *base learner* in the base learner pool:
    Integrate the original model and *base learner* as the new heterogeneous ensemble learning model using the integration strategy
    Evaluate the performance of the new heterogeneous ensemble learning model
Compare and select the heterogeneous ensemble learning model with the best performance
**If** the performance of the new heterogeneous ensemble learning model is superior to the performance of the original model
    Update the original model and evaluation value
**End**

### 3.2.5 Interpretability for the best heterogeneous ensemble learning models

The high predictive performance model is insufficient to support executive decisions, as the lack of model interpretability may lead to skepticism among executives regarding the model's predicted results. Accordingly, elucidating the model's predicted results is necessary and valuable, which can offer executives a plausible explanation for the predicted results, thereby enhancing their confidence in relying on the model's predictions to guide business decisions (Jabeur et al. 2021; Coussement and Benoit 2021; Wang et al. 2022). In our study, we utilize two interpretability methods, namely Shapley additive explanations and partial dependence plot, to conduct research and interpret which features play an important role in the prediction process of the models with the best performance for operating income, net profit, and total assets, as well as how the model makes its predictive response. Specifically, Shapley additive explanations represent the feature importance of AHEL's predictions of operating income, net profit, and total assets, which can help us understand which features are important. However, simply understanding which features are important is not enough to allow the researcher to visualize the response output predicted by the model since it cannot interpret the association between the values of the important features and the predicted responses of AHEL. Thus, we address this problem by leveraging a partial dependence plot. Overall, we combine both techniques to enhance the interpretability and understandability of AHEL's predictions of operating income, net profit, and total assets.

*Shapley additive explanations*. To explain the internal mechanism of the predictive model with the best performance, we apply *Shapley additive explanations* to the model, which provides insight into model behaviour and the contribution of each feature to a real observation (Schlembach et al. 2022). *Shapley additive explanations* is a model-agnostic approach based on explainable artificial intelligence, which derives its basis from game theory. Specifically, the significant values calculated are the Shapley values in game theory and the coefficients of the local linear regression, and the loss function of the kernel Shapley is the follows:

$$L(f,g,\pi_x) = \sum_{d\in D}\left[f\left(h_x(d)\right) - g(d)\right]^2 \pi_x(d) \quad , \tag{7}$$

where $D$ denotes the training set, $g(d) = \theta_0 + \sum_{i=1}^{M} \theta_i d_i$, and $\pi_x(d) = \frac{M-1}{\binom{M}{|d|} d(M-|d|)}$, where $\theta_i$ represents the Shapley value, $M$ denotes the maximum coalition size, and $|d|$ is the total number of features.

*Partial dependence plot.* To explore how important features make predictive responses in predicting different dimensions of company performance, we use the partial dependency graph-based method to explain the prediction models with the best performance for operating income, net profit, and total assets. Valuable information is also extracted to support our further discussion. Specifically, the partial dependence plot displays the marginal effect of one or two features on the outcome of a predictive model. First, we calculate the average marginal effect value of important feature *A* on the model's predicted outcome according to the marginal average function of Eq. (8). Second, we draw a line graph of the predicted outcome of important feature *A* based on the calculated average marginal effect value, reflecting the impacts of different values of important feature *A* on the predicted response of the model through graphical visualization. Finally, we analyze the partial dependence plot of the model on important feature $A$ as follows:

$$g_{X_S}(X_S) = \frac{1}{n}\sum_{i=1}^{n} g(X_S, X_T^{(i)}), \tag{8}$$

where $g$ represents the model with the best performance, $n$ represents the number of samples in the test set, $X_S$ denotes the value of important feature $A$, and $X_T^{(i)}$ is the actual feature value from the test set for features in which we are not interested.

### 3.3 Experimental design

In the experimental design, our systematic analysis of multi-dimensional company performance predictions, including operating income, net profit, and total assets, mainly focuses on three aspects. First, to explain and evaluate the structure and performance of the model for each label obtained from AHEL, we compare and evaluate the best performance of the heterogeneous ensemble learning models developed under different integration strategies. We provide the details of the best learner pool and best integration strategy extracted from AHEL for each label. Furthermore, we compare AHEL with mainstream machine-learning methods in terms of MAE, RMSE, MAPE, and $R^2$, including single regressors (DT and LR) and homogeneous ensemble learning regressors (RF, GBDT, Boost, LightGBM, and Catboost). We use the Bayesian optimization method to tune their hyperparameters based on the value ranges found in the literature and list them in Table 4, which also contains the optimal parameters per machine-learning method (Shahriari et al. 2016). We also further compare the performance of AHEL with four econometric models (lasso regression, ridge regression, ARIMA (autoregressive integrated moving average model), SARIMA(seasonal auto regressive integrated moving average) and four latest heterogeneous ensemble learning models proposed by Papouskova and Hajek (2019), Cui et al. (2021), Wang et al. (2023) and Hou et al. (2023). Finally, to analyze the impact of important features on the predicted results on multi-dimensional company performance, we interpret the model with the best predictive performance for each company performance from AHEL, following the discussion in Subsection 3.2.4.

In our study, we try to train the models using the company's historical data, and predict multi-dimensional company performance based on the company's test data, so the traditional random cross validation (e.g., *k*-fold) is not suitable for this study (Albrecht et al. 2021). To model validation, we perform cross-validation with an expanding rolling window referring to (Albrecht et al. 2021), as shown in Figure 3. We perform five different lengths of time windows from five years to nine years for the training set, and the length of the time window for the test data set is one year. Taking as an example of the length of the time window for the training set to 5 years, the initial model is fitted with its optimized parameters using the observations of five years from 2010 to 2014. We then predict multi-dimensional company performance on the observation of 2015 and evaluate the predictive performance of the models. For the next iteration, we roll the training data one year forward, re-optimize the parameters of the models, predict one year further and re-evaluate the predictive performance of the models. The same ways are performed for the training set for other window length settings. The average performance for each model in terms of performance metrics is reported and used for comparative analysis in Section 4.2. All the training and testing experiments were performed using Python software and Python third-party libraries (numpy, pandas, scikit-learn, etc.).

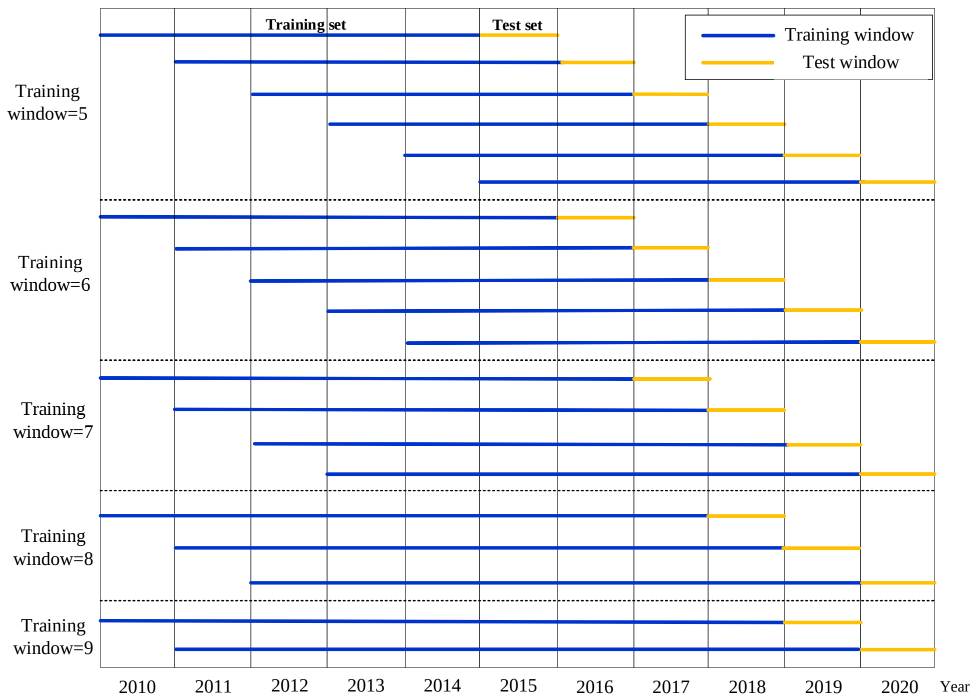



**Figure 3.** Division of the training set and test set.

**Table 4.** Hyperparameter settings in AHEL.

| Algorithm | Parameter | Setting | Value | Algorithm | Parameter | Setting | Value |
|---|---|---|---|---|---|---|---|
| LR | intercept for this model | True | True | XGBoost | number of trees | [10,300] | 100 |
| DT | maximum depth of the tree | [1,12] | 6 | | maximum depth of the tree | [1,12] | 9 |
| RF | number of trees | [10,300] | 125 | | learning rate | [0.01,0.3] | 0.1 |
| | maximum depth of the tree | [1,15] | 6 | LightGBM | number of trees | [10,300] | 70 |
| GBDT | number of trees | [10,300] | 60 | | maximum depth of the tree | [1,12] | 4 |
| | maximum depth of the tree | [1,12] | 7 | | learning rate | [0.01,0.3] | 0.1 |
| | learning rate | [0.01,0.3] | 0.1 | Catboost | maximum iterations | [10,200] | 120 |
| | | | | | learning rate | [0.01,0.3] | 0.1 |

## 4. Results

### 4.1 Model performance evaluation

We now compare the performance of the best heterogeneous ensemble learning models under different strategies obtained from AHEL. We list the results in Table 5 in terms of the four machine learning metrics introduced in Subsection 3.2.3.

For operating income prediction, we see from Table 5 that the heterogeneous ensemble learning model constructed under the stacking integration strategy for the base learners LR, DT, RF, GBDT, and XGBoost, and meta-learner LR has the best performance in terms of RMSE and $R^2$, with values of 162.0037 and 0.4961, respectively. The heterogeneous ensemble learning models constructed under the average-based voting integration strategy achieve the best performance in terms of the MAE and MAPE. For net profit prediction, the heterogeneous ensemble learning model is constructed when the base learner consists of DT and GBDT, and the integration strategy is a weight-based voting integration strategy with the best performance compared with the other two heterogeneous ensemble learning models. For total assets prediction, we see that the model with the stacking-based integration strategy for heterogeneous integration of the base learners LR, DT, and RF and the meta-learner for LR has the best performance in terms of RMSE and $R^2$. In addition, the model with the average-based voting integration strategy for the base learners RF and CatBoost, and the meta-learner for LR has the best performance in terms of RMSE and $R^2$. RF and CatBoost for heterogeneous integration exhibit the best performance in terms of MAE and MAPE.

To compare the performance of the heterogeneous ensemble learning model output by AHEL with that of existing mainstream machine learning models, we select the best-performing model in terms of RMSE and $R^2$ for each label from Table 5 as the best heterogeneous ensemble learning model for that label. Prior studies on corporate management based on machine learning have extensively applied LR, DT, RF, SVR, GBDT, XGBoost, LightGBM, and CatBoost to conduct research because of their better predictive performance (Herrera et al. 2022; Jabeur et al. 2021; Schlembach et al. 2022; Yu et al. 2022). Therefore, we also compare and evaluate the predictive performance of AHEL with the above eight popular regressors, and we show the results in Table 6.

From Table 6, AHEL achieves 162.0037 and 0.4961 for operating income prediction, 5.1000 and 0.7677 for net profit prediction, and 89.6920 and 0.8282 for total assets prediction in terms of RMSE and $R^2$, respectively. AHEL performs better than various homogeneous ensemble learning models in terms of RMSE and $R^2$ for each predictive label, including RF, GBDT, XGBoost, LightGBM, and CatBoost. In addition, we also use the paired *t*-test to test whether the predictive performance of AHEL is statistically different from other models in terms of RMSE, MAE, MAPE, and $R^2$, and we show the results in Table 6. From Table 6, the *p*-value for AHEL and the comparison methods are less than 0.05 in general, which reveals that we should reject that AHEL is not statistically significantly different from the comparison methods.

**Table 5.** Performance evaluation of heterogeneous ensemble learning models with different labels.

| Label | Base learner pool | Integration strategy | MAE | RMSE | MAPE | $R^2$ |
|---|---|---|---|---|---|---|
| Operating income | RF+GBDT+XGBoost | Average-based voting integration strategy | **29.9237** | 165.5413 | **49.7899** | 0.4700 |
| | RF+XGBoost+LightGBM | Weight-based voting integration strategy | 30.9986 | 164.5433 | 49.9305 | 0.4712 |
| | F: LR+DT+RF+GBDT+XGBoost; S: LR | Stacking integration strategy | 30.4106 | **162.0037** | 49.9385 | **0.4961** |
| Net profit | DT+RF+GBDT | Average-based voting integration strategy | **1.7930** | 5.2883 | 12.4858 | **0.7677** |
| | DT+GBDT | Weight-based voting integration strategy | 1.8690 | **5.1000** | 12.4707 | **0.7677** |
| | F: LR+DT+RF+GBDT; S: LR | Stacking integration strategy | 1.8686 | 5.4216 | **12.1928** | 0.7558 |
| Total assets | RF+CatBoost | Average-based voting integration strategy | **27.1690** | 93.7021 | 64.5356 | 0.8094 |
| | RF+CatBoost | Weight-based voting integration strategy | 30.4458 | 92.2731 | **63.9985** | 0.81111 |
| | F: LR+DT+RF; S: LR | Stacking integration strategy | 32.2469 | **89.6920** | 65.1430 | **0.8282** |

Note: The units for operating income, net profit, and total assets are in hundreds of millions of RMB. *F* denotes the first stage; *S* denotes the second stage.

**Table 6.** Performance overview of the different machine learning methods.

| | Label: Operating income | | | | Label: Net profit | | | | Label: Total asset | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model | MAE | RMSE | MAPE | $R^2$ | MAE | RMSE | MAPE | $R^2$ | MAE | RMSE | MAPE | $R^2$ |
| LR | 50.5543$^{***}$ | 182.1377$^{***}$ | **48.2176**$^{**}$ | 0.3631$^{***}$ | 3.2689$^{***}$ | 8.2109$^{***}$ | **8.0906**$^{***}$ | 0.4399$^{***}$ | 66.9134$^{***}$ | 179.4753$^{***}$ | 64.6864$^{**}$ | 0.3122$^{***}$ |
| SVR | 49.5324$^{***}$ | 181.2253$^{***}$ | 50.2673$^{**}$ | 0.4012$^{***}$ | 2.7342$^{***}$ | 10.0138$^{***}$ | 8.2059$^{***}$ | 0.5770$^{***}$ | 43.2356$^{***}$ | 166.6633$^{***}$ | 64.9251$^{***}$ | 0.3412$^{***}$ |
| DT | 32.0337$^{**}$ | 197.8285$^{***}$ | 49.5152$^{*}$ | 0.2486$^{***}$ | 2.3985$^{***}$ | 6.5756$^{***}$ | 13.0487$^{***}$ | 0.6408$^{***}$ | 37.0864$^{***}$ | 189.4363$^{***}$ | 64.8005$^{***}$ | 0.2338$^{***}$ |
| RF | 31.6422$^{***}$ | 167.7107$^{**}$ | 50.0732$^{***}$ | 0.4600$^{***}$ | **1.7944**$^{**}$ | 5.9222$^{***}$ | 12.5285$^{***}$ | 0.7086$^{***}$ | **27.7695**$^{***}$ | 95.3567$^{***}$ | 64.7404$^{***}$ | 0.8059$^{**}$ |
| GBDT | 32.0702$^{***}$ | 170.8406$^{***}$ | 49.7313$^{**}$ | 0.4396$^{***}$ | 1.9108$^{**}$ | 5.5153$^{*}$ | 12.2823$^{**}$ | 0.7473$^{***}$ | 31.9744$^{***}$ | 105.9337$^{***}$ | 64.7549$^{***}$ | 0.7604$^{***}$ |
| XGBoost | 31.1112$^{***}$ | 168.4650$^{***}$ | 49.7881$^{**}$ | 0.4551$^{***}$ | 1.8555$^{**}$ | 5.7807$^{**}$ | 12.1777$^{*}$ | 0.7224$^{***}$ | 30.0647$^{***}$ | 172.6517$^{***}$ | 64.6788$^{**}$ | 0.3635$^{***}$ |
| LightGBM | 36.6872$^{***}$ | 178.4379$^{***}$ | 50.2316$^{**}$ | 0.3887$^{***}$ | 1.9327$^{**}$ | 6.4231$^{***}$ | 12.3015 | 0.6573$^{***}$ | 36.7903$^{***}$ | 118.6370$^{***}$ | 64.9442$^{***}$ | 0.6995$^{***}$ |
| CatBoost | 31.1118$^{***}$ | 180.6653$^{***}$ | 49.7156$^{**}$ | 0.3733$^{***}$ | 1.8285$^{**}$ | 6.5936$^{***}$ | 12.2451$^{*}$ | 0.6388$^{***}$ | 29.5560$^{***}$ | 107.9571$^{***}$ | **64.3899**$^{***}$ | 0.7512$^{***}$ |
| AHEL | 30.4106 | **162.0037** | 49.9385 | **0.4961** | 1.8690 | **5.1000** | 12.4707 | **0.7677** | 32.2469 | **89.6920** | 65.1430 | **0.8282** |

Note: The units for operating income, net profit, and total assets are in hundreds of millions of RMB. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

To further verify the superiority of AHEL, we compare the performance of AHEL and four econometric models (lasso regression, ridge regression, ARIMA, SARIMA), and four latest heterogeneous ensemble learning models proposed by Papouskova and Hajek (2019), Cui et al. (2021), Wang et al. (2022) and Hou et al. (2023), and the experimental results are listed in Table 7. As we can see from Table 7, AHEL all achieves the best performance in terms of RMSE and $R^2$ compared with four econometric models and four latest heterogeneous ensemble learning models in multidimensional company performance and exists significantly different from them in terms of *p*-value. As such, there are good reasons to believe that AHEL is superior to other popular machine-learning methods in predicting multi-dimensional company performance.

**Table 7.** Comparison of the econometric models and latest heterogeneous ensemble learning models and AHEL.

| Prediction | Model | MAE | RMSE | MAPE | $R^2$ |
|---|---|---|---|---|---|
| Operating income | Lasso | 49.162$^{***}$ | 185.9388$^{***}$ | 51.3217$^{***}$ | 0.3341$^{***}$ |
| | Ridge | 48.1449$^{***}$ | 184.2691$^{***}$ | 51.0483$^{***}$ | 0.3452$^{***}$ |
| | ARIMA | 38.5615$^{***}$ | 175.3452$^{***}$ | **49.6474**$^{*}$ | 0.36601$^{***}$ |
| | SARIMA | 36.7111$^{***}$ | 173.588$^{***}$ | 49.9753 | 0.3694$^{***}$ |
| | Papouskova and Hajek (2019) | 31.0723$^{***}$ | 164.0231$^{***}$ | 50.1004$^{***}$ | 0.4671$^{***}$ |
| | Cui et al. (2021) | **29.9901**$^{***}$ | 163.3212$^{***}$ | 49.7483$^{**}$ | 0.4713$^{***}$ |
| | Wang et al. (2022) | 31.0100$^{*}$ | 163.9563$^{***}$ | 49.9174$^{*}$ | 0.4659$^{***}$ |
| | Hou et al. (2023) | 32.6751$^{***}$ | 164.9878$^{***}$ | 50.2362$^{***}$ | 0.4535$^{***}$ |
| | AHEL | 30.4106 | **162.0037** | 49.9385 | **0.4961** |
| Net profit | Lasso | 3.0355$^{***}$ | 9.102$^{***}$ | **9.2584**$^{***}$ | 0.4213$^{***}$ |
| | Ridge | 3.0652$^{***}$ | 9.1287$^{***}$ | 10.0273$^{***}$ | 0.4236$^{***}$ |
| | ARIMA | 3.0452$^{***}$ | 9.1473$^{***}$ | 10.0014$^{***}$ | 0.4201$^{***}$ |
| | SARIMA | 2.7849$^{***}$ | 9.2545$^{***}$ | 11.973$^{***}$ | 0.4338$^{***}$ |
| | Papouskova and Hajek (2019) | 1.8943$^{**}$ | 5.8817$^{***}$ | 12.5103$^{**}$ | 0.7401$^{***}$ |
| | Cui et al. (2021) | 1.8788$^{**}$ | 5.6699$^{***}$ | 12.9635$^{***}$ | 0.7453$^{***}$ |
| | Wang et al. (2022) | **1.8335** | 5.7839$^{***}$ | 12.8451$^{***}$ | 0.7501$^{***}$ |
| | Hou et al. (2023) | 1.9057$^{***}$ | 5.9312$^{***}$ | 13.0100$^{***}$ | 0.7393$^{***}$ |
| | AHEL | 1.8690 | **5.1000** | 12.4707 | **0.7677** |
| Total asset | Lasso | 68.2423$^{***}$ | 171.3443$^{***}$ | 64.3219$^{***}$ | 0.3793$^{***}$ |
| | Ridge | 68.176$^{***}$ | 171.0331$^{***}$ | **65.0503**$^{***}$ | 0.3742$^{***}$ |
| | ARIMA | 45.1305$^{***}$ | 154.8982$^{***}$ | 64.7855$^{***}$ | 0.4512$^{***}$ |
| | SARIMA | 43.5847$^{***}$ | 152.4578$^{***}$ | 64.6713$^{***}$ | 0.4545$^{***}$ |
| | Papouskova and Hajek (2019) | **27.9954**$^{***}$ | 92.0003$^{***}$ | 65.2395$^{***}$ | 0.8123$^{***}$ |
| | Cui et al. (2021) | 28.5463$^{***}$ | 91.0125$^{***}$ | 64.2449$^{***}$ | 0.8105$^{***}$ |
| | Wang et al. (2022) | 30.7716$^{***}$ | 91.3457$^{***}$ | 64.3281$^{*}$ | 0.8136$^{***}$ |
| | Hou et al. (2023) | 31.9258$^{***}$ | 92.2026$^{***}$ | 65.2567$^{***}$ | 0.8089$^{***}$ |

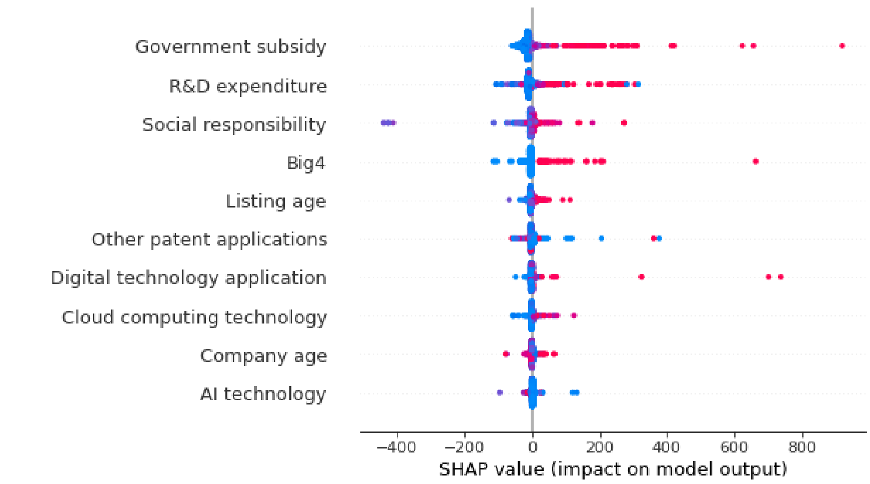
| | | | | |
|---|---|---|---|---|
| AHEL | 32.2469 | **89.6920** | 65.143 | **0.8282** |

Note: The units for operating income, net profit, and total assets are in hundreds of millions of RMB. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

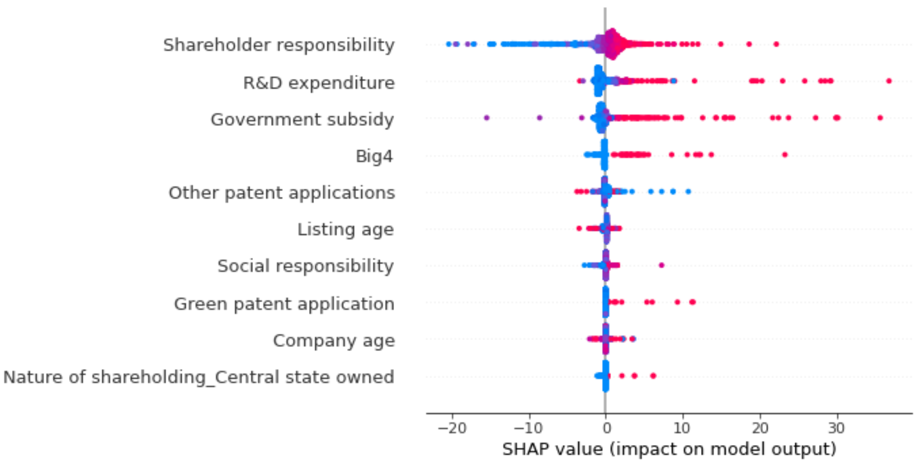
### 4.2 Model interpretability

In practice, managers generally wonder not only the most important features in predictions, but also how these features affect the predicted response so as to make better-informed decisions. This subsection interprets AHEL's predictions of operating income, net profit, and total assets. First, we analyze the feature importance of operating income, net profit, and total assets predictions using the Shapley value. The Shapley value of a feature of a given prediction indicates the degree of change in AHEL's prediction when we observe that feature (Schlembach et al. 2022). Figure 4 summarizes the ten most important features in terms of operating income, net profit, and total assets predictions and reflects all the Shapley values for a single feature, where the $x$-axis represents the Shapley values, the red dots represent high feature values, and the blue dots represent low feature values.

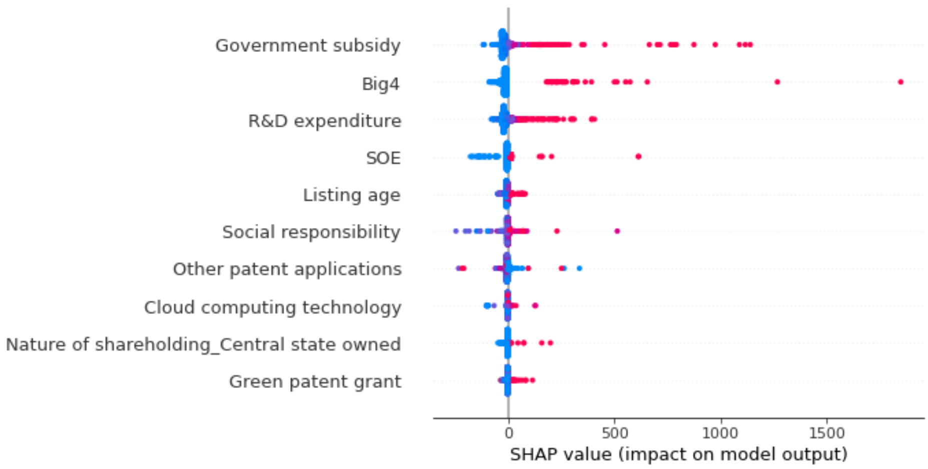
Figure 4(a) suggests that "government subsidy", "R&D expenditure", and "social responsibility" are the top three most important features of AHEL's operating income prediction. Among the first ten important features, it can be observed that the red dots in rows "government subsidy", "R&D expenditure", "social responsibility", "Big4", "listing age", and "digital technology application" tend to appear on the right-hand side, which means that the high values of these features lead to a higher predicted operating income. Thus, managers should pay attention to company performance in these features, and avoid poor performance in these features. From Figure 4(b), managers can learn that "shareholder responsibility", "R&D expenditure", and "government subsidy" are the top three most important features of AHEL's net profit prediction. The blue and red dots in rows "shareholder responsibility" tend to appear on the left-hand and right-hand sides, respectively, which reveals that the low (high) values of these features lead to a lower (higher) predicted net profit. This finding means that managers should pay attention to shareholder responsibility, and understand low shareholder responsibility is critical to yield a high net profit. In addition, from Figure 4(c), managers should focus on the performance of "government subsidy", "R&D expenditure", and "Big4", because they are the top three most important features of total assets prediction. For these three features, a higher value drives AHEL to predict a higher value of total assets.




(a) Feature importance for operating income prediction.




(b) Feature importance for net profit prediction.




(c) Feature importance for total assets prediction.

**Figure 4.** Feature importance for multi-dimensional company performance.

Subsequently, drawing on the partial dependence plot, we also visualize the general influence trends of smoothing curves (blue line) and raw curves (gray line) for the important features of operating income, net profit, and total assets predictions based on the analysis in Figure 4, which are shown in Figures 5, 6, and 7, respectively.

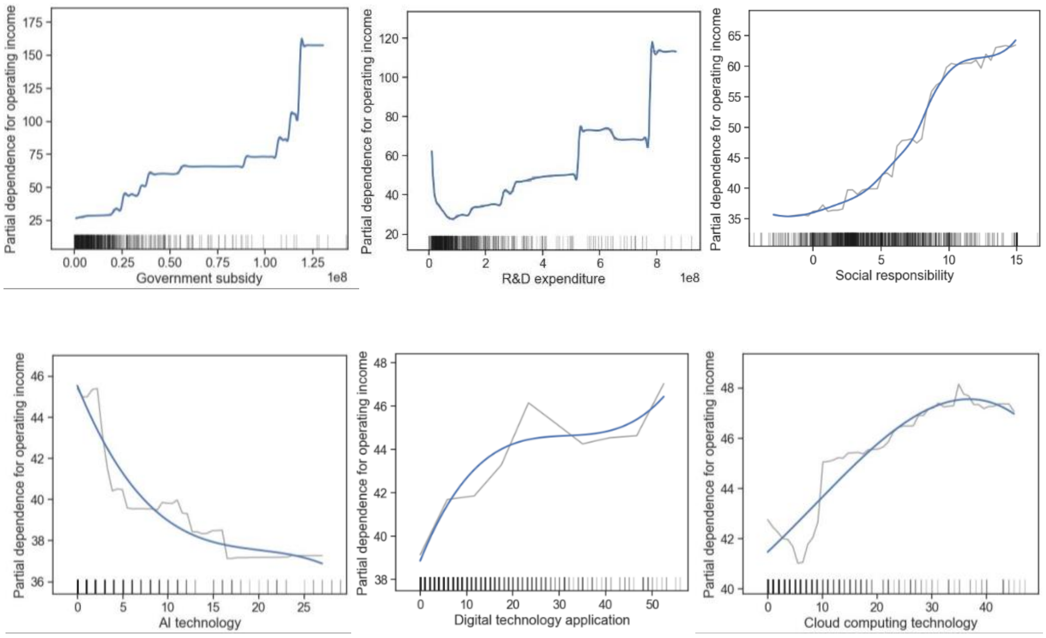



**Figure 5.** Partial dependence plots for the six most influential features on operating income prediction.

From Figure 5, managers should be aware that the larger the values of "government subsidy", "R&D expenditure", "social responsibility", "digital technology application", and "cloud computing application" are, the higher the operating income in the general trend, while "AI technology" has the opposite trend. These findings support the arguments that companies with good social responsibility generally have higher performance and sustainability (Dixon-Fowler et al. 2013; Endrikat et al. 2014; Franceschelli et al. 2019), and that companies tend to benefit from government subsidies as well as new knowledge gained through R&D expenditure (Patel et al. 2018; Luo et al. 2021).

As we can observe from Figure 6, similar to the trend of "government subsidy" and "R&D expenditure" on operating income, these features also have a positive impact on the predicted net profit. Furthermore, "shareholder responsibility" and "social responsibility" positively influence the predicted net profit in an approximately linear way. The influence of shareholder responsibility on company performance is typically in line with the influence of social responsibility on company performance (Dixon-Fowler et al. 2013; Franceschelli et al. 2019; Saeidi et al. 2015), and fostering a stronger sense of responsibility is crucial for companies. Furthermore, the number of fewer green patent applications is negatively correlated with company performance, and only when the number of green patent applications exceeds a certain threshold does it positively affect company performance. This non-linear relationship between green patent applications and company performance, emphasizes the significance of innovation in driving positive outcomes beyond a certain critical point (Hatzikian, 2015).

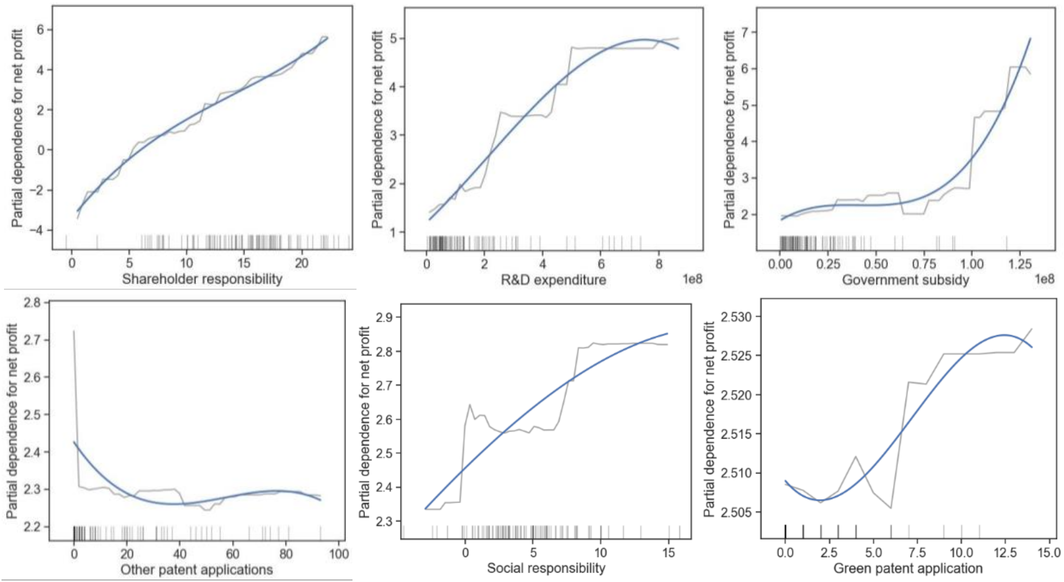
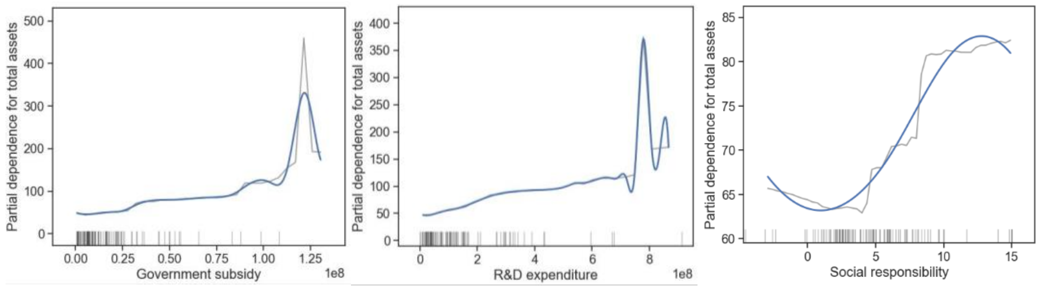



**Figure 6.** Partial dependence plots for the six most influential features on net profit prediction.

As seen in Figure 7, the predicted total assets fluctuate significantly when the values of "government subsidy" and "R&D expenditures" are large, thus conservative companies should be cautious about investing too much R&D expenditure or applying for too much government subsidy (Chu et al. 2017; Sueyoshi and Goto 2009). In addition, for features "cloud computing application" and "green patent grant, when their values reach a certain threshold, the predicted total assets of the company tend to stabilize. These features have a negative impact on firm performance when their values are low, and changes in their lower values do not bring about positive changes in predicted total assets. However, we suggest that companies should persist in their development because as their values increase, company performance is moving 'in the positive direction', which is in line with Kohtamki et al. (2020), Albuquenrquer et al. (2018) and Ben Lahouel et al. (2022).



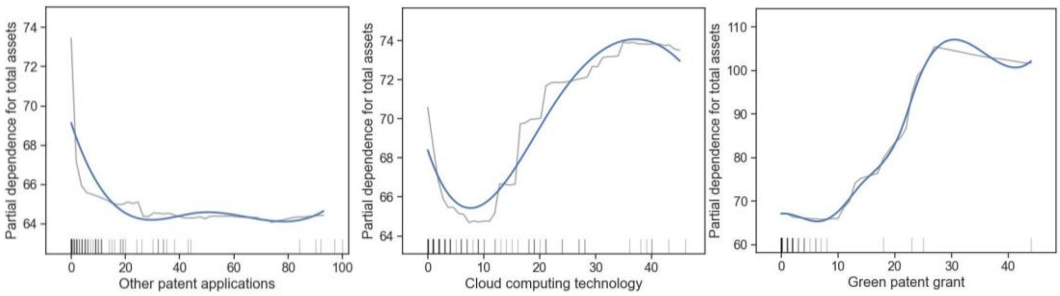



**Figure 7.** Partial dependence plots for the six most influential features on total assets prediction.

## 5. Discussions

This section discusses several important implications based on these comparisons and interpretations. First, our experimental results provide strong evidence that heterogeneous ensemble learning models outperform single models and homogeneous ensemble learning models in terms of predicted performance, which corroborates the findings of previous studies (Wang et al. 2022; Hou et al. 2023). The diversity between the base learners and the integration strategy are two of the most critical variables affecting the performance of the ensemble learning model (Feng et al. 2022). The superiority of AHEL is attributed to these variables. First, it considers a wide range of base learners with different structures and principles, which enhances the diversity of the base learners. Second, it embeds three different advanced integration strategies that help the base learners achieve the best integration effect.

Second, our interpreted results suggest that managers should pay more attention to "social responsibility" and "shareholder responsibility", because they all have a trend of a positive influence on the predicted operating income, net profit, and total assets. Previous studies have investigated the influence of CSR on company financial performance using econometric methods, but the results are inconsistent (Ben Lahouel et al. 2022; Busch and Friede 2018; Dixon-Fowler et al. 2013; Gonenc and Scholtens 2017). Our findings support the positive influence from the viewpoint of model interpretability in general trends, which is in line with the findings of the positive influence of CSR (Busch and Friede 2018; Dixon-Fowler et al. 2013; Endrikat et al. 2014; Franceschelli et al. 2019; Saeidi et al. 2015). Although it requires much effort to maintain a company's CSR image, such as engaging in activism, philanthropy, and other social activities, CSR performance is, in principle, consistent with the overall governance performance. It is worthwhile for companies to put their effort into attaining good CSR, and companies with good CSR generally also have good overall governance, which is central to achieving strong performance and a sustainable business model (Busch and Friede 2018; Dixon-Fowler et al. 2013; Endrikat et al. 2014; Franceschelli et al. 2019; Saeidi et al. 2015).

Third, our study supports the idea that R&D expenditure positively influences the predicted results of operating income, net profit, and total assets from AHEL. Because R&D expenditure involves high levels of information asymmetry, which is risky and long-term for the company, such as an innovative product or service that does not meet market demand, the company performance brought about by R&D expenditure is uncertain, as shown in Table 1. Although there is debate as to whether R&D expenditure contributes to business performance, we endorse the view that it can improve business performance by reducing production costs and introducing new products, which is in line with the findings of (James and McGuire 2016; Patel et al. 2018; Wang et al. 2017). In addition, R&D activities can be beneficial to a company's long-term performance by generating new knowledge, thereby expanding the company's knowledge base and improving the absorption and integration of existing knowledge (James and McGuire 2016; Patel et al. 2018; Wang et al. 2017).

Fourth, we provide new evidence on the influence of digital transformation on company performance. Our results indicate that digital transformation has mixed impacts, both positive and negative, on AHEL's operating income and total assets predictions. An increase in the number of AI technologies leads to a drop in the predicted operating income of AHEL, whereas an increase in the number of cloud computing technologies leads to an overall upward trend in the predicted operating income and total assets. On the one hand, the digitalization drive has been demonstrated to increase the value of its product-services offerings and the productivity of its people, resulting in improved profitability and financial performance of the company (Chouaibi et al. 2022; Ribeiro-Navarrete et al. 2021; Truant et al. 2021). On the other hand, traditional service channels with non-digital service solutions can directly and effectively support customer access, which causes a negative link between the digital and business performance of the company, especially when digital technologies do not meet the needs of existing customers (Zhou et al. 2021).

Fifth, our study yields new insights into the influence of government subsidy on multi-dimensional company performance. Previous studies have argued that companies can use subsidy to invest in technological upgrades, resulting in lower production costs, higher output, and fixed asset growth rates (Luo et al. 2021; Wang et al. 2021). Meanwhile, other studies suggest that subsidy can make companies less efficient, making them less likely to negatively influence company performance (Chu et al. 2017). Our study also shows that increased government subsidy leads to an overall upward trend in company performance; this is not a linear increase. The contribution of government subsidy to company performance is reinforced when they are greater than a threshold value, supporting the findings of (Yang et al. 2019).

Finally, we believe that the influence of innovation on the net profit and total assets of a company, as predicted by AHEL, is mixed. The predicted net profit and total assets show a trend of an initially slight decrease, then an increase, and finally, stabilization as the number of green patents increases. In contrast, for the other types of patents, the predicted net profits and total assets fall steeply with the number of patents, then stabilize around a value. Green innovation can reduce the consumption of raw materials and energy, improve the speed of resource acquisition, and differentiate itself from competitors by providing innovative products, thus positively influencing corporate performance and reputation (Liao 2018). However, it has to be accepted that companies that allocate resources to innovation efforts may face increased operating costs due to the complexity and risks involved in the innovation process (Albuquerquer et al. 2018).

*Management implications*. This study has three practical implications for managers and CEOs. First, we enhance the applications of machine learning methods drawing on interpretable technologies, and mitigate the 'black box' nature of machine learning methods. By interpreting the predicted response of AHEL in a visual way, managers can see how the features they are interested in influence the trend of company performance predictions, thus trusting the predicted results of the models. Second, the results still emphasize that CSR, R&D expenditure, and government subsidy positively influence multi-dimensional company performance in general. Especially shareholder and social responsibility, they influence the net profit in an approximately linear way. Thus, it is worthwhile for companies to put their effort into attaining good CSR, and companies with good CSR generally also have good overall governance, which is central to achieving strong performance and a sustainable business model (Busch and Friede 2018; Dixon-Fowler et al. 2013; Endrikat et al. 2014; Franceschelli et al. 2019; Saeidi et al. 2015). Furthermore, drawing on the information extracted from *post hoc* interpretations, our study provides guidance on how companies should develop and make decisions in terms of government subsidy, R&D expenditure, digital transformation, and innovation to achieve greater company performance. We emphasize that it is worthwhile to pursue development in innovation and digital transformation for Chinese listed companies. Despite innovation and digital transformation may not yield a significant positive effect on company performance, they will evolve in a positive direction with their persistence.

**6. Conclusion**

Our study advances data-driven company management by introducing a novel adaptive heterogeneous ensemble learning (AHEL) for multiple-dimensional company performance prediction and decision-making. We address key gaps in the literature by extracting five-dimensional variables through feature engineering and developing AHEL based on these variables, thus providing a more accurate prediction and comprehensive decision-making of multiple-dimensional company performance, including operating income, net profit, and total assets (Ben Lahouel et al. 2022; Aastvedt et al. 2021). The AHEL method, which can automatically output the best heterogeneous ensemble learning models for each predicted aspect, demonstrates superior predictive accuracy compared to several state-of-the-art methods using Chinese listed companies' data. In addition to its predictive prowess, AHEL also offers valuable interpretability, providing a detailed analysis of feature importance and their impact on the prediction outcomes utilizing shapley additive explanations and partial dependence plots. Overall, this study is a pioneering attempt to predict and interpret multi-dimensional company performance, paving the way for future studies in the company performance prediction field.

Our study also contributes to the literature by addressing three critical research gaps: 1) the limited exploration of multi-dimensional variables impacts on company performance (Ben Lahouel et al. 2022; Chouaibi et al. 2022); 2) the inadequacies of linear regressions in capturing complex relationships between investigated variables and company performance (Aastvedt et al. 2021; Chouaibi et al. 2022); 3) the insufficient development of heterogeneous ensemble learning methods in multi-dimensional performance prediction and decision-making (Wang et al. 2022; Wu et al. 2022). From a managerial perspective, AHEL enhances decision-makers' understanding of the relationships between multi-dimensional variables and company performance—spanning CSR, digital transformation, R&D expenditure, government subsidies, and innovation—and supports better organizational decisions.

However, our work also has three limitations that warrant further exploration in future research. First, we demonstrate AHEL's superiority in multi-dimensional company performance from Chinese stock-listed companies. As a result, the findings are context-specific and may not be generalizable to companies in other regions. A valuable extension of our study would be to apply the AHEL method to different types of companies and geographic regions to assess its broader applicability.

Second, while we extracted key variables from sources such as Hexun, CSMAR, and CNRDS, there may still be relevant variables not included in our analysis. Future research could enhance the robustness of predictions by incorporating additional data sources, such as customer and supplier information, as well as CEO characteristics (Hopp et al. 2023). This would provide a more comprehensive understanding of the factors influencing company performance.

Finally, we worked with structured data, which typically favors machine learning over deep learning. For image and text datasets, deep learning models generally outperform machine learning in predictive tasks (De

Caigny et al. 2020; Potrawa and Tetereva 2022). Future research could explore company performance prediction by combining multiple techniques based on multimodal data.

**Compliance with Ethical Standards**

The authors declare that they have no conflict of interest, and that this study does not contain any studies with human participants or animals performed by any of the authors.

**References**


Aastvedt, T. M., Behmiri, N. B., & Li, L. (2021). Does green innovation damage financial performance of oil and gas companies? *Resources Policy*, *73*(4), 102235.

Alam, A., Uddin, M., Yazdifar, H., Shafique, S., & Lartey, T. (2020). R&D investment, firm performance and moderating role of system and safeguard: Evidence from emerging markets. *Journal of Business Research*, *106*, 94–105.

Albrecht, T., Rausch, T. M., & Derra, N. D. (2021). Call me maybe: Methods and practical implementation of artificial intelligence in call center arrivals' forecasting. *Journal of Business Research*, *123*. https://doi.org/10.1016/j.jbusres.2020.09.033

Albuquerquer, S., Basso, L. C., Kimura, H., & Sobreiro, V. A. (2018). Innovation and financial performance of companies doing business in Brazil. *International Journal of Innovation Studies*, *2*(4), 153–164. https://doi.org/10.1016/j.ijis.2019.03.001

Baradaran R., H., Amjadian, A., Sebt, M. V., Askari, R., & Gharaei, A. (2022). An ensemble method of the machine learning to prognosticate the gastric cancer. *Annals of Operations Research*. https://doi.org/10.1007/s10479-022-04964-1

Becchetti, L., & Ciciretti, R. (1978). Corporate social responsibility and stock market performance. *Academy of Management Journal*, *21*(3), 479–486.

Belhadi, A., Kamble, S. S., Mani, V., Benkhati, I., & Touriki, F. E. (2021). An ensemble machine learning approach for forecasting credit risk of agricultural SMEs' investments in agriculture 4.0 through supply chain finance. *Annals of Operations Research*.

Ben Lahouel, B., Ben Zaied, Y., Managi, S., & Taleb, L. (2022). Re-thinking about U: The relevance of regime-switching model in the relationship between environmental corporate social responsibility and financial performance. *Journal of Business Research*, *140*, 498–519.

Bond, S. R., & Guceri, I. (2018). R&D and productivity: evidence from large UK establishments with substantial R&D activities. *Economics of Innovation and New Technology*, 1–13.

Breiman, L. (2001). Random forest. *Machine Learning*, 45, 5–32.

Busch, T., & Friede, G. (2018). The robustness of the corporate social and financial performance relation: A second-order meta-analysis. *Corporate Social Responsibility and Environmental Management*, *25*(4), 583–608. https://doi.org/10.1002/csr.1480

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*, 785-794.

Chouaibi, S., Festa, G., Quaglia, R., & Rossi, M. (2022). The risky impact of digital transformation on organizational performance–evidence from Tunisia. *Technological Forecasting and Social Change*, *178*, 121571.

Chu, Y. L., Wang, J., & Cheng, C. Z. (2017). China's "Mercantilist" government subsidies, the cost of debt and firm performance. *Journal of Banking & Finance*, *86*, 37–52.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273–297.

Coussement, K., & Benoit, D. F. (2021). Interpretable data science for decision making. *Decision Support Systems*, *150*, 113664.

Cui, S., Wang, Y., Wang, D., Sai, Q., Huang, Z., & Cheng, T. C. E. (2021). A two-layer nested heterogeneous ensemble learning predictive method for COVID-19 mortality. *Applied Soft Computing*, *113*, 107946

Cui, Y., Zhang, Y., Guo, J., Hu, H., & Meng, H. (2019). Top management team knowledge heterogeneity, ownership structure and financial performance: Evidence from Chinese IT listed companies. *Technological Forecasting and Social Change*, *140*, 14–

21.
Dixon-Fowler, H. R., Slater, D. J., Johnson, J. L., Ellstrand, A. E., & Romi, A. M. (2013). Beyond '"does it pay to be green?"' A meta- analysis of moderators of the CEP–CFP relationship. *Journal of Business Ethics*, *112*(2), 353–366.
E Horváthová. (2011). Does environmental performance affect financial performance? A meta-analysis. *Ecological Economics*, *70*(1), 52–59.
E Horváthová. (2012). The impact of environmental performance on firm performance: Short-term costs and long-term benefits? *Ecological Economics*, *84*, 91–97.
Ehie, I. C., & Olibe, K. (2010). The effect of R&D investment on firm value: An examination of US manufacturing and service industries. *International Journal of Production Economics*, *128*(1), 127–135.
Endrikat, J., Guenther, E., & Hoppe, H. (2014). Making sense of conflicting empirical findings: A meta-analytic review of the relationship between corporate environmental and financial performance. *European Management Journal*, *32*(5), 735–751.
Feng, Y., Yin, Y., Wang, D., & Dhamotharan, L. (2022). A dynamic ensemble selection method for bank telemarketing sales prediction. *Journal of Business Research*, *139*, 368–382.
Franceschelli, M. V., Santoro, G., Giacosa, E., & Quaglia, R. (2019). Assessing the determinants of performance in the recycling business: Evidence from the Italian context. *Corporate Social Responsibility and Environmental Management*, *26*(11).
Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232.
Gonenc, H., & Scholtens, B. (2017). Environmental and financial performance of fossil fuel firms: A closer inspection of their interaction. *Ecological Economics*, *132*(Complete), 307–328. https://doi.org/10.1016/j.ecolecon.2016.10.004
Gopalakrishnan S (2000). Unraveling the links between dimensions of innovation and organizational performance. *The Journal of High Technology Management Research*, *11*(1): 137-153.
Hatzikian, Y. (2015). Exploring the link between innovation and firm performance. *Journal of the Knowledge Economy*, *6*(4), 749–768.
Herrera, R., Climent, F., Carmona, P., & Momparler, A. (2022). The manipulation of Euribor: An analysis with machine learning classification techniques. *Technological Forecasting and Social Change*, *176*, 121466.
Hopp, C., Wentzel, D., & Rose, S. (2023). Chief executive officers' appearance predicts company performance, or does it? A replication study and extension focusing on CEO successions. *The Leadership Quarterly*, 34(4), 101437.
Hou, H., Liu, C., Wei, R., He, H., Wang, L., & Li, W. (2023). Outage duration prediction under typhoon disaster with stacking ensemble learning. *Reliability Engineering & System Safety*, *237*, 109398.
Hu, Y., Che, D., Wu, F., & Chang, X. (2023). Corporate maturity mismatch and enterprise digital transformation: Evidence from China. *Finance Research Letters*, *53*, 103677.
Huang, C.-W., & Huarng, K.-H. (2015). Evaluating the performance of biotechnology companies by causal recipes. *Journal of Business Research*, 6.
Jabeur, S. B., Gharib, C., Mefteh-Wali, S., & Arfi, W. B. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change*, *166*, 120658.
James, B. E., & McGuire, J. B. (2016). Transactional-institutional fit: Corporate governance of R&D investment in different institutional contexts. *Journal of Business Research*, *69*(9), 3478–3486.
Jeon, H., Seo, W., Park, E., & Choi, S. (2020). Hybrid machine learning approach for popularity prediction of newly released contents of online video streaming services. *Technological Forecasting and Social Change*, *161*, 120303.
Katila, R., Chen, E. L., & Piezunka, H. (2012). All the right moves: How entrepreneurial firms compete effectively. *Strategic Entrepreneurship Journal*, *6*(2), 116–132.
Ke, G., Meng, Q., Finely, T., Wang, T., Chen, W., Ma, W., et al. (2017). LightGBM: A highly efficient gradient boosting decision tree. *In Advances in Neural Information Processing Systems 30* (NIP 2017).
Kohtamki, M., Parida, V., Patel, P. C., & Gebauer, H. (2020). The relationship between digitalization and servitization: The role of servitization in capturing the financial potential of digitalization. *Technological Forecasting and Social Change*, *151*, 119804.
Latan, H., Chiappetta Jabbour, C. J., Lopes de Sousa Jabbour, A. B., Renwick, D. W. S., Wamba, S. F., & Shahbaz, M. (2018). 'Too-much-of-a-good-thing'? The role of advanced eco-learning and contingency factors on the relationship between corporate

environmental and financial performance. *Journal of Environmental Management*, *220*, 163-172.

Liao, Z. (2018). Corporate culture, environmental innovation and financial performance. *Business Strategy and the Environment*, *27*(4), 368–1375.

Luo, G., Liu, Y., Zhang, L., Xu, X., & Guo, Y. (2021). Do governmental subsidies improve the financial performance of China's new energy power generation enterprises? *Energy*, *227*, 120432.

McWilliams, A., & Siegel, D. (2001). Corporate social responsibility: A theory of the firm perspective. *Academy of Management Review*, *26*(1), 117–127.

Nelling, E., & Webb, E. (2009). Corporate social responsibility and financial performance: the "virtuous circle" revisited. *Review of Quantitative Finance & Accounting*, *32*(2), 197–209.

Nirino, N., Santoro, G., Miglietta, N., & Quaglia, R. (2021). Corporate controversies and company's financial performance: Exploring the moderating role of ESG practices. *Technological Forecasting and Social Change*, *162*, 120341.

Papouskova, M., & Hajek, P. (2019). Two-stage consumer credit risk modelling using heterogeneous ensemble learning. *Decision Support Systems*, *118*, 33–45.

Patel, P. C., Guedes, M. J., Soares, N., & da Conceição Gonçalves, V. (2018). Strength of the association between R&D volatility and firm growth: The roles of corporate governance and tangible asset volatility. *Journal of Business Research*, *88*, 282–288.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features (pp. 6638–6648). *Advances in neural information processing systems*, *31*: 6638-6648 .

Ribeiro-Navarrete, S., Botella-Carrubi, D., D Palacios-Marqués, & Orero-Blat, M. (2021). The effect of digitalization on business performance: An applied study of KIBS. *Journal of Business Research*, *126*, 319–326.

Saeidi, S. P., Sofian, S., Saeidi, P., Saeidi, S., & Saaeidi, S. (2015). How does corporate social responsibility contribute to firm financial performance? The mediating role of competitive advantage, reputation, and customer satisfaction. *Journal of Business Research*, *68*, 341–350.

Schlembach, C., Schmidt, S. L., Schreyer, D., & Wunderlich, L. (2022). Forecasting the Olympic medal distribution – A socioeconomic machine learning model. *Technological Forecasting and Social Change*, *175*, 121314.

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., & de Freitas, N. (2016). Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, *104*(1), 148–175. Presented at the Proceedings of the IEEE.

Sueyoshi, T., & Goto, M. (2009). Can R&D expenditure avoid corporate bankruptcy? Comparison between Japanese machinery and electric equipment industries using DEA–discriminant analysis. *European Journal of Operational Research*, *196*(1), 289–311.

Truant, E., Broccardo, L., & Dana, L. P. (2021). Digitalisation boosts company performance: an overview of Italian listed companies. *Technological Forecasting and Social Change*, *173*, 121173.

Velte, P. (2017). Does ESG performance have an impact on financial performance? Evidence from Germany. *Journal of Global Responsibility*, *8*(2): 169-178.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356–365.

Wang, L., Jia, F., Chen, L., & Xu, Q. (2022). Forecasting SMEs' credit risk in supply chain finance with a sampling strategy based on machine learning techniques. *Annals of Operations Research*. https://doi.org/10.1007/s10479-022-04518-5

Wang, R., Wang, F., Xu, L., & Yuan, C. (2017). R&D expenditures, ultimate ownership and future performance: Evidence from China. *Journal of Business Research*, *71*, 47–54.

Wang, N., Xu, Y., & Wang, S. (2022). Interpretable boosting tree ensemble method for multisource building fire loss prediction. *Reliability Engineering & System Safety*, *225*, 108587.

Wang, X., Li, Z., Shaikh, R., Ranjha, A. R., & Batala, L. K. (2021). Do government subsidies promote financial performance? Fresh evidence from China's new energy vehicle industry. *Sustainable Production and Consumption*, *28*, 142–153.

Wu, Z., Jing, L., Wu, B., & Jin, L. (2022). A PCA-AdaBoost model for E-commerce customer churn prediction. *Annals of Operations Research*. https://doi.org/10.1007/s10479-022-04526-5

Yang, M., Sulaiman, R., Yin, Y., Mallamaci, V., & Alrabaiah, H. (2022). The effect of business intelligence, organizational learning and innovation on the financial performance of innovative companies located in Science Park. *Information Processing &*

*Management*, *59*(2), 102852.
Yang, X., He, L., Xia, Y., & Chen, Y. (2019). Effect of government subsidies on renewable energy investments: The threshold effect. *Energy Policy*, *132*, 156–166.
Yu, B., Li, C., Mirza, N., & Umar, M. (2022). Forecasting credit ratings of decarbonized firms: Comparative assessment of machine learning models. *Technological Forecasting and Social Change*, *174*, 121255.
Zhang, H., Li, L., Zhou, D., & Zhou, P. (2014). Political connections, government subsidies and firm financial performance: Evidence from renewable energy manufacturing in China. *Renewable Energy*, *63*, 330-336.
Zhou, D., Kautonen, M., Dai, W., & Zhang, H. (2021). Exploring how digitalization influences incumbents in financial services: The role of entrepreneurial orientation, firm assets, and organizational legitimacy. *Technological Forecasting and Social Change*, *173*, 121120.

## Appendix

**Table A1.** Feature description.

| Feature | Description | Type |
|---|---|---|
| Operating income | Operating income (Target variable) | Floating-Point |
| Net profit | Net profit (Target variable) | Floating-Point |
| Total assets | Total assets (Target variable) | Floating-Point |
| Shareholder responsibility | Shareholder responsibility score | Integer |
| Environmental responsibility | Environmental governance score | Integer |
| Social responsibility | Contribution value score | Integer |
| AI technology | The level of AI technology | Integer |
| Block chain technology | The level of block chain technology | Integer |
| Cloud computing technology | The level of cloud computing technology | Integer |
| Big data technology | The level of big data technology | Integer |
| Digital technology application | The level of digital technology application | Integer |
| R&D expenditure | Research and development expenditures | Floating-Point |
| Environmental investment | Environmental investment | Floating-Point |
| Government subsidy | Government subsidy | Floating-Point |
| Green patent application | Number of green patent applications | Integer |
| Green patent grant | Number of green patents grant | Integer |
| Other patent application | Number of other types of patent applications (excluding green patents) | Integer |
| Heavy polluting company | Whether it is a heavily polluting company | Bool |
| Number of directors | Number of directors | Integer |
| Number of independent directors | Number of independent directors | Integer |
| Nature of shareholding | Nature of shareholding （Local state-owned enterprises, public enterprises, collective enterprises, private enterprises, other enterprises, | Floating-Point |

| | | |
|---|---|---|
| | foreign-funded enterprises, central state-owned 8140enterprises） | |
| Company age | Company age （Calculated from the year of company registration） | Integer |
| Listing age | Age of listing （Calculated from the year the company was listed） | Integer |
| SOE | Whether the company is a state-ownership status for the top one shareholder of the listed companies | Bool |
| Big4 | Audited by the Big Four (PwC, Deloitte, KPMG, Ernst & Young) | Bool |

**Table A2.** Summary of digital transformation features.

| Digital transformation | Feature |
|---|---|
| AI technology | Artificial intelligence, Business intelligence, Image understanding, Investment decision support systems, Intelligent data analytics, Intelligent robotics, Machine learning, Deep learning, Semantic search, Biometric recognition, Face recognition, Speech recognition, Identity verification, Autonomous driving, Natural language processing |
| Block chain technology | Blockchain, Digital currency, Distributed computing, Differential privacy technologies, Smart financial contracts |
| Cloud computing technology | Cloud computing, Streaming computing, Graph computing, In-memory computing, Multi-party secure computing, Brain-like computing, Green computing, Cognitive computing, Converged architectures, Billion dollar concurrency, Exabyte-level storage, Internet of things, Information physical systems |
| Big data technology | Big data, Data mining, Text mining, Data visualization, Heterogeneous data, Credit collection, Augmented reality, Mixed reality, Virtual reality |
| Digital technology application | Mobile internet, Industrial internet, Mobile internet, Internet healthcare, E-commerce, Mobile payment, Third party payment, NFC payment, Smart energy, B2B, B2C, C2B, C2C, O2O, Non-banking Institute Payment, Smart wear, Smart agriculture, Smart transportation, Smart healthcare, Smart customer service, Smart investment, Smart cultural tourism, Smart environmental protection, Smart grid, Smart marketing, Digital marketing, Unmanned retail, Internet finance, Digital finance, Fintech, Quantitative finance, Open banking |