https://doi.org/10.62836/iaet.v3i1.312

Global Science Publishing

IAET

Innovations in Applied Engineering and Technology
https://ojs.sgsci.org/journals/iaet

Article

Stacked Generalization for Robust Prediction of Trust and Private Equity on Financial Performances

Kaixian Xu [1], Yunxiang Gan [2,] and Alan Wilson [3,*]

[1] *Risk & Quant Analytics, BlackRock, Jersey City, NJ 07310, USA*

[2] *Moloco, Redwood City, CA 94063, USA*

[3] *Intact Financial Corporation, Toronto, ON M5G 1Z5, Canada*

Abstract: Predicting financial performance, particularly in the Trust and Private Equity sectors, is a critical challenge for investors and analysts, as it directly impacts decision-making and risk management. This study addresses this problem by proposing a novel solution using stacked generalization, an ensemble learning technique, to improve the accuracy and robustness of financial performance predictions. We utilized a dataset consisting of key financial metrics, including Net Income, EBITDA, Market Cap, Gross Profit, Current Ratio, and Debt/Equity Ratio, to predict Net Income as a proxy for Trust and Private Equity performance. The proposed solution employs Random Forest (RF), Linear Regression (LR), and Support Vector Machine (SVM) as the base model and the Neural network as meta-model to enhance prediction precision. The results indicate that stacked generalization significantly outperforms traditional single-model approaches, yielding higher accuracy and reliability in forecasting financial performance. This study not only offers an effective tool for financial prediction but also provides a foundation for future work, where additional features, more advanced models, and sector-specific applications can be explored to improve predictive capabilities.

Keywords: stacked generalization; financial performance; trust; private equity

1. Introduction

In the financial sector, the ability to predict a company's performance with high accuracy is crucial for informed decision-making, especially in areas like private equity and trust management [1]. Financial metrics such as Net Income, EBITDA, Market Cap, Gross Profit, Current Ratio, and Debt/Equity Ratio are essential indicators of a company's health and are frequently used for financial forecasting. However, despite the abundance of financial data, accurately predicting these outcomes remains challenging due to the complex interrelationships between variables and the non-linear nature of financial markets. As investors in private equity and trust management sectors rely heavily on the accuracy of financial forecasts to guide their strategies, improving the precision of predictive models is critical [2]. End-to-end stock recommendation algorithms based on time-frequency consistency capture market patterns, offering a new framework to improve short-term financial forecasting accuracy.[3]

This study aims to enhance prediction accuracy in financial performance forecasting by using advanced machine learning techniques, specifically Stacked Generalization. By integrating multiple base models,

Received: 18 November 2024; Accepted: 6 December 2024.

* Corresponding: Alan Wilson (alan.wilson@intact.net)

including Random Forests (RF), Linear Regression (LR), and Support Vector Machines (SVM), and combining them with a Neural Network (NN) as the meta-model, this research proposes a robust approach to predicting financial outcomes such as Net Income. The goal is to provide a more accurate tool for predicting financial performance, thereby assisting investors and analysts in making more informed decisions, managing risks better, and developing effective investment strategies. Research focused on Chinese A-share listed companies has demonstrated that digital transformation significantly enhances their equity financing capabilities, thereby improving the efficiency of capital operations amid the wave of informatization [4,5].

The problem of predicting financial performance is inherently complex due to the interdependencies and non-linear relationships present in financial data and machine learning [6–11]. Financial variables such as Net Income, EBITDA, Market Cap, Gross Profit, Current Ratio, and Debt/Equity Ratio are not independent; instead, they subtly influence each other [12]. The difficulty in accurately predicting these variables lies in their intricate connections and the challenges posed by noisy, missing, or incomplete data. Predicting metrics like Net Income is vital in private equity and trust management as it informs decisions about investments, risk management, and strategic planning [13].

In recent years, machine learning models such as has emerged as a powerful tool for financial prediction, enabling analysts to model complex relationships and make data-driven decisions [14]. Similarly, fingerprint image generation using attention-based deep generative adversarial networks demonstrates the potential of complex generative models in ensuring data integrity and financial information security [15,16]. In the field of e-commerce, predictive methods combining multi-empirical mode decomposition with deep learning models have been proven to achieve higher accuracy and robustness when handling large-scale and complex data [17, 18]. Traditional statistical models such as RF, SVM, and LR have been widely used for financial forecasting. However, they struggle to handle high-dimensional, noisy, and non-linear data. For instance, RF leverages ensemble learning to perform well with complex and high-dimensional datasets but is prone to overfitting when the data contains substantial noise [19]. SVMs, with their ability to model complex patterns using appropriate kernels, are effective for high-dimensional datasets but are sensitive to noisy data and computationally demanding [20]. LR provides an accessible approach with a clear interpretability advantage, yet it is less suitable for non-linear relationships [21]. While each model offers value, their limitations highlight the need for more comprehensive approaches. Stacked Generalization, often referred to as Stacking, addresses these gaps by integrating the outputs of diverse base models into a meta-model, leading to enhanced prediction accuracy and reliability [22]. The meta-model, typically a Neural Network (NN), learns how to combine the outputs of the base models best to generate a more accurate final prediction. Although Stacked Generalization has been successfully applied in various domains, its application in financial performance prediction-especially in predicting Net Income and Trust/Private Equity performance—has been underexplored. This research seeks to fill that gap by applying Stacked Generalization to financial data and evaluating its effectiveness in improving prediction accuracy.

Most existing studies focus on simpler ensemble methods or the combination of tree-based models, such as Random Forests and Gradient Boosting. Few have explored the benefits of combining these particular models for financial prediction, especially in predicting Net Income and assessing its impact on Trust and Private Equity. Additionally, many of the existing models do not adequately address the challenges posed by missing, noisy, or incomplete data, which are common in financial datasets. While Random Forests and SVMs are relatively robust to noise, their performance can still be compromised in certain conditions. This study seeks to address these research gaps by applying Stacked Generalization to financial prediction, thus offering a more robust and accurate approach for forecasting financial performance and improving decision-making in Trust and Private Equity. Notably, with the deepening of digitalization, both the financial and industrial sectors have witnessed innovative applications centered on multi-model fusion technology, offering more flexible and diverse solutions for enhancing corporate performance and risk management [18,19].

By using a Neural Network as the meta-model, this approach leverages the NN's ability to learn non-linear patterns and fine-tune the predictions made by the base models. The performance of the Stacked Generalization model was evaluated using several financial metrics, including Mean Absolute Error (MAE), Root Mean

Squared Error (RMSE), and R-squared. The study aims to demonstrate that Stacked Generalization based on RF, LR, and SVM with a Neural Network meta-model offers a more accurate and robust solution for predicting financial outcomes such as Net Income in the context of Trust and Private Equity.

The primary objective of this study is to develop a robust predictive model for financial performance using Stacked Generalization. Specifically, the research focuses on predicting Net Income and analyzing the impact of key financial metrics on Trust and Private Equity performance. The specific objectives of this study are:

1. To apply Stacked Generalization by using Random Forests (RF), Linear Regression (LR), and Support Vector Machines (SVM) as base models and a Neural Network (NN) as the meta-model to predict Net Income.
2. To compare the performance of the Stacked Generalization model with each base model to assess improvements in prediction accuracy.
3. To investigate the relationships between financial metrics such as Market Cap and Debt/Equity Ratio and their impact on the performance of Trust and Private Equity investments.
4. To evaluate the robustness of the Stacked Generalization model in handling incomplete and non-linear financial data.

This study is structured into distinct sections to explore the utility and implementation of Stacked Generalization for financial performance predictions in the Trust and Private Equity sectors. Section 2, Methodology, details the data collection, preprocessing strategies, and the selection of the target variable while also outlining the theoretical foundation of the applied machine learning models. Section 3, Model Selection and Implementation, describes the integration of base models—Random Forest, Linear Regression, and Support Vector Machines—with a Neural Network as the meta-model. Section 4, Results analysis, evaluates the performance of the proposed model against traditional methods using statistical metrics such as Mean Absolute Error, Root Mean Squared Error, and R-squared. Section 5, Conclusion, synthesizes our findings, discusses the implications for financial analytics in the Trust and Private Equity sectors, and proposes directions for future research.

2. Methodology

2.1. Data Collection and Preprocessing

The dataset employed in this study is an open-access resource available on Kaggle, comprising financial performance metrics from a range of companies between 2009 and 2023, with 32 indexes. We selected six financial indicators crucial for predicting the performance in the Trust and Private Equity sectors, Net Income, EBITDA, Market Cap, Gross Profit, Current Ratio, and Debt/Equity Ratio. It encompasses data from 161 companies, collected over several years, providing a robust basis for training and testing predictive models.

The dataset is split into training and testing sets to evaluate the performance of the machine learning models accurately. The training set consists of 70% of the data used to train the models. The remaining 30% constitutes the testing set, which is used to assess the models' predictive accuracy and generalization capability outside of the training data. This division ensures that the evaluation of the models is both rigorous and indicative of their performance in real-world scenarios. Before model implementation, the data underwent the following preprocessing steps:

- Handling Missing Values: Missing data points were imputed using the most appropriate method, either mean imputation for continuous variables or mode imputation for categorical variables, ensuring no significant data loss.
- Normalization/Standardization: Continuous features were standardized to ensure all features were on the same scale, which is essential for models like Linear Regression and Neural Networks.
- Feature Engineering: Key features influencing equity performance were identified, such as net income, EBITDA, Market Cap, Gross Profit, Current Ratio, and Debt/Equity Ratio. These features were transformed where necessary to improve model interpretation.

2.2. Heatmap Trends in Correlation Analysis

A correlation heatmap was generated to visually analyze the relationships between the target variable (Net Income) and the potential predictor variables and among the predictor variables. The correlation heatmap provides a clear representation of the degree of association between each pair of variables in the dataset, where:

- Positive correlations are indicated by values close to +1 (strong positive correlation),
- Negative correlations are shown with values close to −1 (strong negative correlation),
- No correlation is represented by values near 0.

The heatmap also uses colors to represent correlation strengths. Typically, a gradient color scale is applied, where darker hues represent stronger correlations, and lighter hues represent weaker correlations.

2.3. Model Selection

A stacking generalization approach was employed for predicting Net Income, utilizing multiple base models and a meta-model to improve prediction accuracy and robustness. The following base models were selected based on their strengths in handling various aspects of the data:

- **Random Forest (RF):** An ensemble learning method that creates multiple decision trees and aggregates their outputs to improve prediction accuracy. RF was chosen due to its ability to model non-linear relationships and interactions in the data.
- **Linear Regression (LR):** A linear regression model with **L2 regularization** to prevent overfitting and handle multicollinearity, making it suitable for the financial dataset with potential feature interdependencies.
- **Support Vector Machine (SVM):** A kernel-based method for regression (SVR) that is highly effective at capturing complex, non-linear relationships between features and the target variable.

The Neural Network (NN) was selected as the meta-model for stacking Generalization. Neural Networks are particularly suited for combining the predictions of base models, as they can learn complex interactions and dependencies between base model outputs to enhance overall prediction accuracy.

2.4. Stacking Generalization Approach

The stacking generalization method follows a two-level model structure:

- **Base Level:** The base models (RF, RM, SVM) are trained independently on the dataset. Each model generates predictions that serve as input features for the meta-model.
  - For each base model, hyperparameters were optimized using **grid search** with **cross-validation** to ensure optimal performance. A training set and validation set were used to train the base models. In contrast, a separate test set was reserved for the final model performance evaluation. **Meta-Level (Neural Network):** The predictions from the base models (RF, RM, SVM) are passed as input features to the Neural Network, which learns the optimal weights for combining these inputs into a final prediction. The Neural Network acts as the meta-model and is trained to improve the overall prediction accuracy based on the outputs of the base models (as shown in Figure 1).

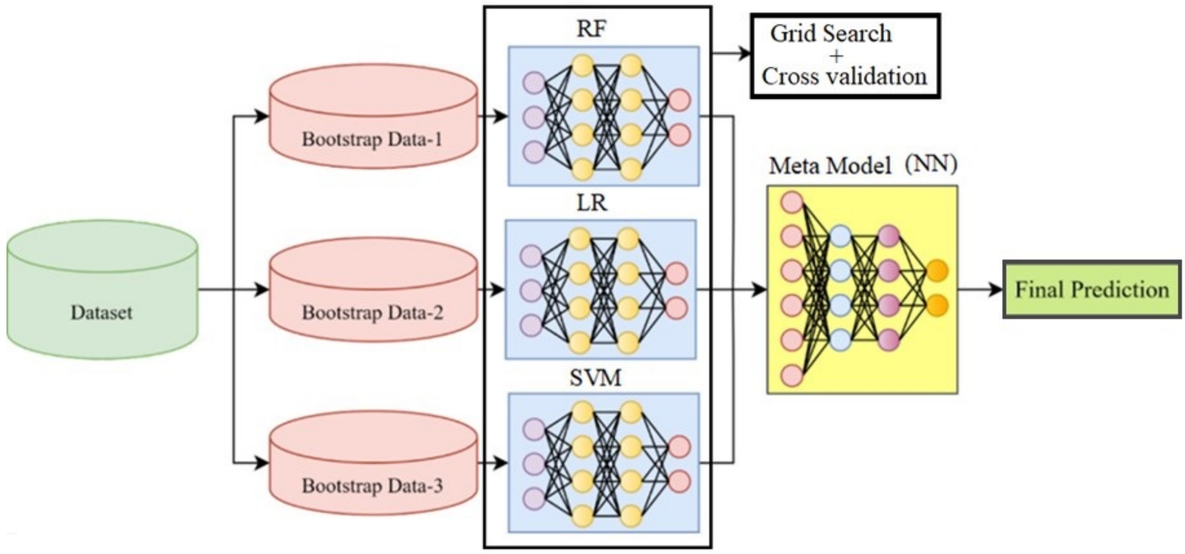

Figure 1. Structure diagram of stack generalization for robust prediction.

2.5. Performance Evaluation

To assess the performance of the models, the following metrics were calculated:

- Mean Absolute Error (MAE): Measures the average magnitude of errors in the model's predictions, interpreting model accuracy.
- Root Mean Squared Error (RMSE): Highlights the model's error magnitude by giving higher weight to larger errors, making it useful in financial predictions.
- R-squared (R^2): Indicates the proportion of variance in the target variable (Net Income) explained by the model. A higher R^2 value implies better model fit and predictive power.

2.6. Model Tuning and Optimization

Hyperparameter optimization was carried out for all models (base models and meta-model) to ensure the best performance:

- Random Forest: Parameters such as n_estimators (number of trees) and max_depth were tuned.
- Linear Regression: linear regression's parameters, such as regularization terms are adjusted to mitigate overfitting and handle multicollinearity among features.
- Support Vector Machine: The C parameter (penalty term) and kernel type (RBF or linear) were adjusted to improve Generalization.
- Neural Network: The meta-model's hyperparameters were also tuned, including number of layers, neurons per layer, learning rate, activation function (e.g., ReLU, Sigmoid), and batch size.

Hyperparameter optimization was performed using grid search with cross-validation to ensure that the best configurations were selected.

3. Results Analysis

This section presents the findings from the analysis of financial performance metrics, focusing on the relationships between selected variables (Net Income, EBITDA, Market Cap, Gross Profit, Current ratio, and Debit/Equity ratio), model performance, and prediction accuracy. The results are visualized through heatmaps, line plots, and scatter plots, offering a clear view of how these variables interact over time. Additionally, we evaluate the performance of multiple predictive models, including Random Forest, Linear Regression, Support Vector Machine, and a Meta-Model (Neural Network), comparing their ability to predict Net Income. The analysis also thoroughly examines residuals, offering insight into the models' error distributions and their overall calibration. These results provide a comprehensive understanding of the relationships between financial metrics and the predictive capabilities of different machine learning models for estimating Net Income.

Figure 2 illustrates the correlations between key financial variables in the dataset, with color intensity signifying the strength of the relationships. The strong positive correlations between Net Income, EBITDA, Market Cap, and Gross Profit are immediately noticeable. The dark red shading of these variables suggests that they tend to move in the same direction, with fluctuations in EBITDA, Market Cap, and Gross Profit typically aligning with changes in Net Income. This reinforces the idea that profitability, as measured by Net Income, is influenced by operational performance (EBITDA) and overall company size (Market Cap).

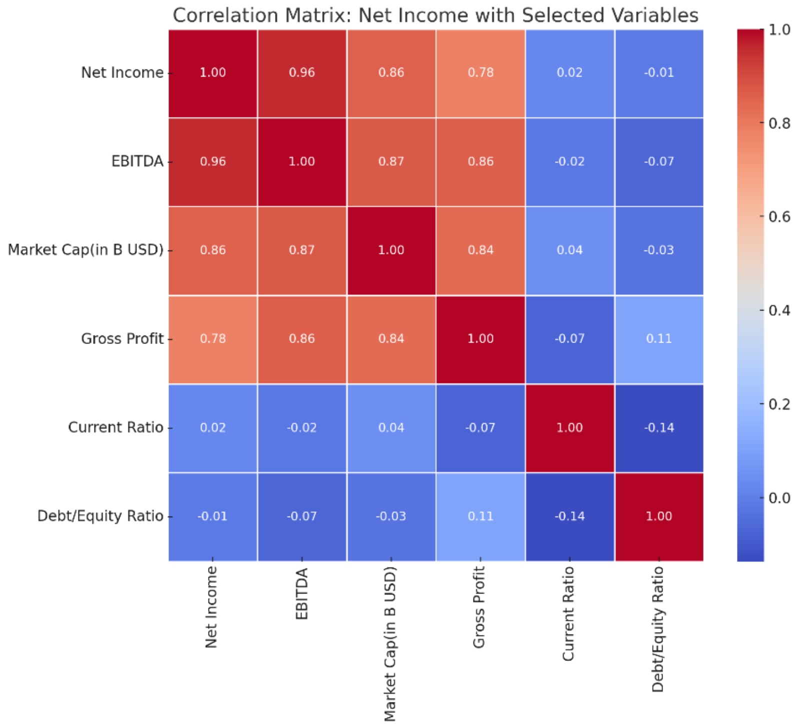

Figure 2. Heat map of correlations between key financial variables.

Figure 3 shows the relationship between Net Income and key financial metrics: EBITDA, Market Cap, and Gross Profit. The first plot shows a positive correlation between Net Income and EBITDA, suggesting that companies with higher EBITDA tend to have higher net income. The second plot illustrates the connection between Net Income and Market Cap, where a moderate correlation is observed, indicating that larger companies by market capitalization generally report higher net income, although some outliers are present. Lastly, the third plot demonstrates a strong positive correlation between Net Income and Gross Profit, implying that companies with higher gross profits also tend to achieve higher net income. Together, these plots provide valuable insights into how different financial indicators relate to a company's profitability and overall financial performance.

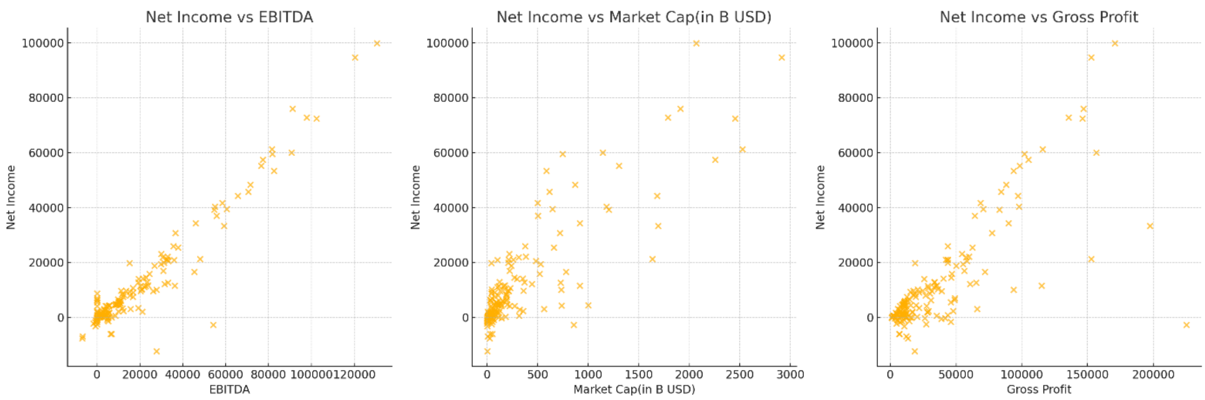

Figure 3. Strong correlation relationship between Net Income and key financial metrics.

Figure 4 shows Net Income with weaker or negative correlation variables—namely Current Ratio and Debt/Equity Ratio—reveal interesting insights. In the case of the Current Ratio, the plot does not show a clear linear relationship with Net Income, reinforcing the weak correlation of −0.34. This suggests that the liquidity position of companies, as reflected by the current ratio, does not significantly impact their profitability. Similarly, the scatter plot for the Debt/Equity Ratio also shows no distinct pattern with Net Income, supporting the very low correlation of −0.28 Companies with varying debt-to-equity ratios appear to exhibit similar levels of profitability, indicating that the capital structure might not play a central role in determining Net Income compared to other financial metrics like EBITDA or Market Cap.

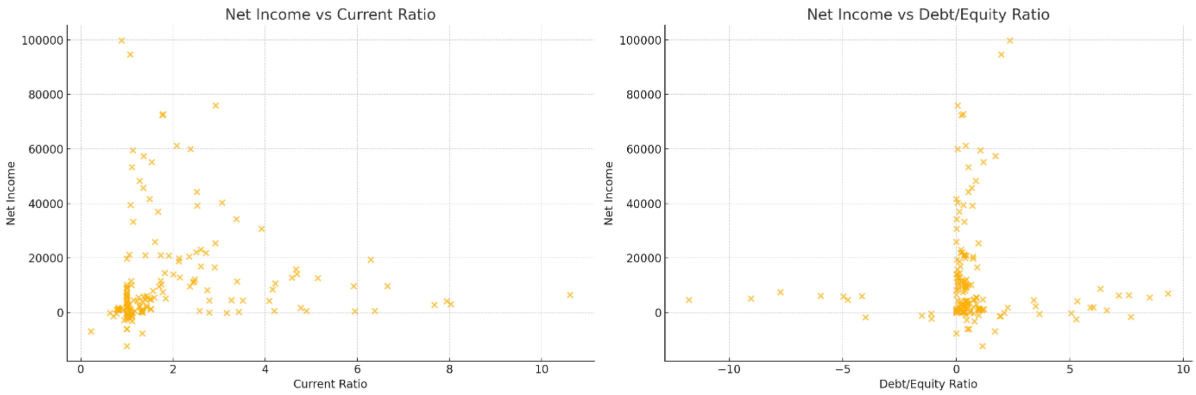

Figure 4. Comparison of negative correlation variables with target variable net Income.

The relationships among predictors in Figure 5—Net Income, EBITDA, Market Cap, and Gross Profit—illustrate the dynamic interplay among these financial metrics, each contributing uniquely to the overall financial narrative of a company. These predictors collectively highlight the operational and market-based dimensions of company performance. For instance, Gross Profit strongly influences Net Income as it represents the initial profitability metric from which all other financial outcomes flow. Companies with higher Gross Profit not only have greater operational efficiency but also tend to achieve higher EBITDA, showcasing their ability to control production costs and optimize margins.

Additionally, the relationship between EBITDA and Market Cap reinforces the idea that operational efficiency is closely tied to investor confidence and market valuation. Larger market caps are often observed in companies with strong EBITDA, as they signify resilience and potential for long-term growth. Similarly, the interconnectedness of Market Cap, Gross Profit, and EBITDA underscores the significance of company size and operational success in shaping profitability metrics. These metrics collectively reflect the strength of a company's financial health, making them indispensable for advanced predictive models.

By examining these interconnected relationships, Figure 5 emphasizes the critical role these variables play in accurately forecasting financial outcomes like Net Income. Their combined impact is vital for understanding trust and private equity performance, as these metrics not only signal operational and financial stability but also influence investor perception and market dynamics. This interconnected framework validates their importance in predictive modeling and decision-making processes, particularly in the context of robust ensemble techniques like stacked generalization.

As shown in Figure 6, the time history curves reveal the trends of Net Income, EBITDA, Market Cap, and Gross Profit of selected companies based on prior analysis from 2009–2023. Net Income exhibits notable fluctuations, showing that the companies in the dataset experience cycles in profitability, likely due to economic conditions or industry-specific trends. The EBITDA trends follow similar fluctuations to those of net income. However, the values are generally higher due to EBITDA's focus on earnings before deducting interest, taxes, and depreciation. Market Cap shows a more stable upward trend, reflecting general company growth and increasing market valuation over time. In contrast, Gross Profit has fluctuations which can also be seen in Net Income and EBITDA, with peaks and declines influenced by changes in sales and costs. These trends further support the earlier positive correlations, underscoring the dynamic relationships between these financial variables.

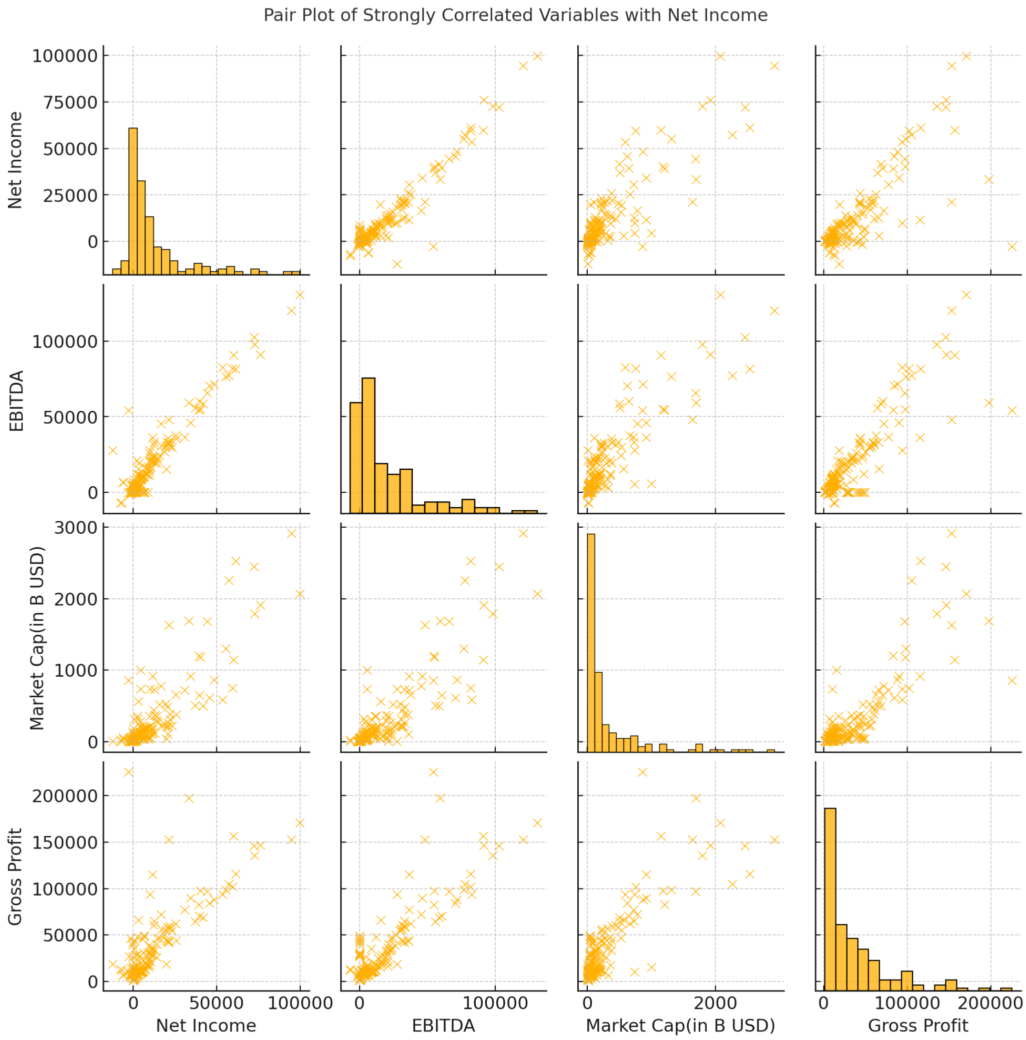

Figure 5. Relationship comparison of correlated variable.

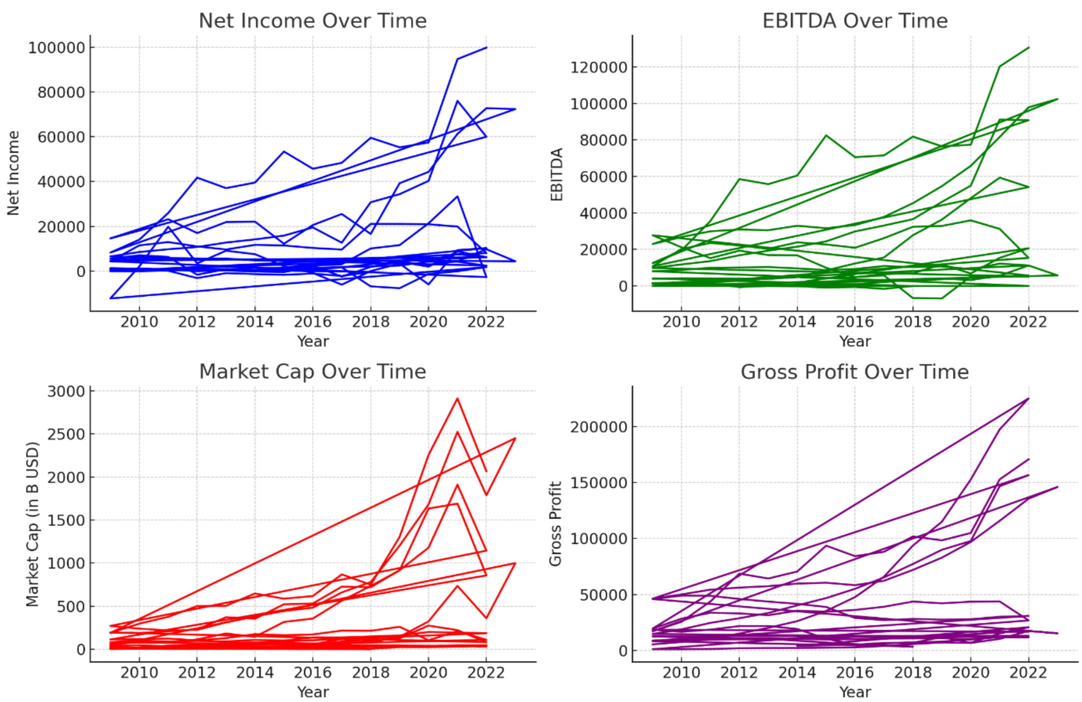

Figure 6. Trends of significant variables.

Stack Generalization for Robust Prediction

The performance of different predictive models was assessed using various metrics, including R-squared, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE), as displayed in Table 1. SVM performed reasonably with an R-squared of 0.47. However, the Random Forest outperformed it with an R-squared of 0.74, indicating a better fit. LR model followed closely with an R-squared of 0.78, reflecting good performance. Notably, the Meta-Model (Neural Network) achieved an impressive R-squared of 0.97, showing exceptional predictive accuracy. It also exhibited the lowest MAE and RMSE, demonstrating its robustness in predicting Net Income accurately.

Table 1. Evaluation metrics for different models.

	R-Square	MAE	RMSE
Support Vector Machine (SVM)	0.47	11172.69	14,043.97
Random Forest (RF)	0.74	7782.94	9910.12
Linear Regression (LM)	0.78	6446.63	8075.48
Meta-Model (NN)	0.97	3287.43	4487.42

Figure 7 compares the residual patterns of the base models (RF, LR, and SVM) and the Meta-Model (Neural Network). The Meta-Model demonstrates a residual distribution characterized by random scatter around the zero line, highlighting its ability to integrate base model predictions effectively without introducing bias. In contrast, SVM exhibits significant bias, as evidenced by the relatively large deviations in residuals. RF and LR show similar performances and achieve relatively good results, though not as good as the Meta-Model. Collectively, Figure 7 and Table 1 reinforce the superiority of the Meta-Model in minimizing residuals and delivering consistent performance, while also highlighting the shortcomings of the base models.

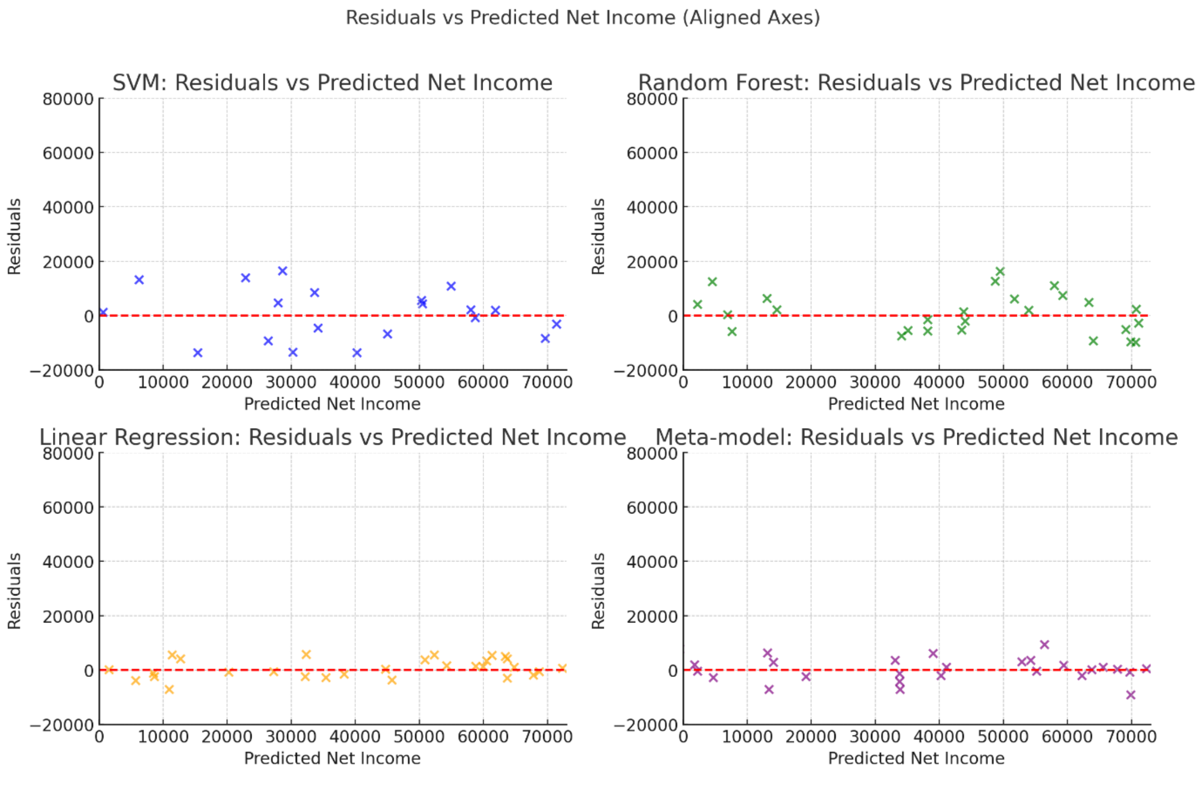

Figure 7. Residual comparison of Based model and Meta-model.

Figure 8 shows how well the predicted values align with the actual values. The red dashed line represents a perfect prediction (where predicted equals actual).

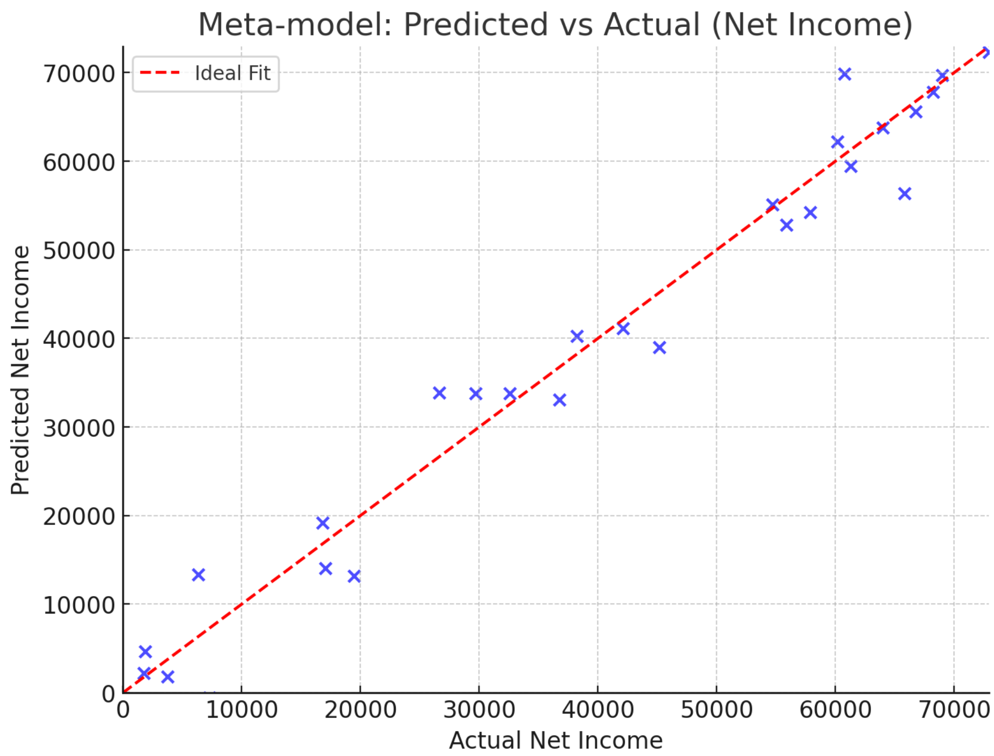

Figure 8. Predicted vs. Actual (Net Income).

Figure 9 compares the predicted and actual values for the net income of different companies randomly selected from the dataset. The closer the predicted values are to the actual values, the better the model's accuracy. For the Meta-Model, the predictions closely align with the actual net income values, as indicated by the near-perfect overlap of the predicted values (represented by the orange dashed line) with the actual values (represented by the blue line). This confirms the high accuracy of the Meta-Model in predicting net income over the test dataset and highlights its superiority over the base models.

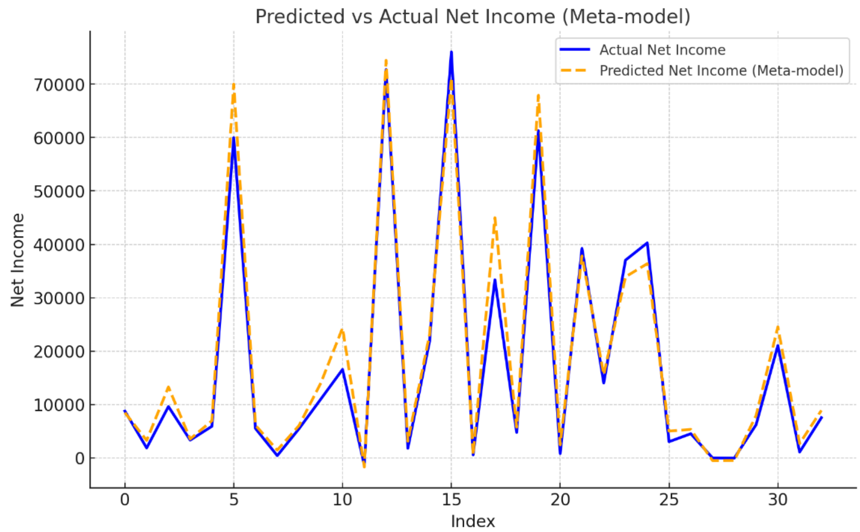

Figure 9. Comparison of actual and predicted value of target variable (Net Income).

4. Conclusions

This study underscores the potential of stacked generalization to enhance the prediction of Trust and Private Equity performance, offering a robust alternative to traditional single-model approaches. By integrating Random Forest (RF), Linear Regression (LR), and Support Vector Machines (SVM) as base models with a Neural Network (NN) meta-model, we achieved significant gains in both accuracy and reliability. The findings highlight the value of combining diverse predictive algorithms to address the complexities of financial datasets, including non-linear relationships and variable interdependencies.

The results confirm that stacked generalization excels at leveraging the strengths of individual models while mitigating their limitations, establishing it as a powerful tool for forecasting financial performance metrics such as Net Income. This capability is essential for informed decision-making in dynamic sectors like Trust and Private Equity, where predictive precision directly influences investment strategies and risk management.

Future research could build on this work by incorporating additional features, testing with sector-specific datasets, or exploring advanced ensemble techniques to further improve predictive accuracy. The methodologies and insights presented in this study contribute to the advancement of financial analytics, fostering more effective data-driven decision-making in the financial sector.

Funding

Not applicable.

Author Contributions

Conceptualization, data collection, analysis, K.X., Y.G. and A.W.; writing—original draft preparation, K.X., Y. G. and A.W.; writing—review and editing, K.X., Y.G. and A.W. All of the authors read and agreed to the published the final manuscript.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The dataset is publicly available on: rish59/financial-statements-of-major-companies2009–2023.

Conflicts of Interest

The authors declare no conflict of interest.

References

1 Holland J. Fund Management, Intellectual Capital, Intangibles and Private Disclosure. *Managerial Finance* 2006; **32**: 277–316.

2 Goodman TH, Neamtiu M, Shroff N, *et al.* Management Forecast Quality and Capital Investment Decisions. *The Accounting Review* 2014; **89**: 331–365.

3 Gan Y, Chen X. The Research on End-to-end Stock Recommendation Algorithm Based on Time-frequency Consistency. *Advances in Computer and Communication* 2024; **5**: 243–259.

4 Tang Y. Investigating the Impact of Digital Transformation on Equity Financing: Empirical Evidence from Chinese A-share Listed Enterprises. *Journal of Humanities, Arts and Social Science* 2024; **8**: 1620–1632.

5 Tang Y. Investigating the Impact of Regional Digital Finance Development on Short-run IPO Performance: Empirical Evidence from China. *Journal of Management Science & Engineering Research* 2024; **7**: 31–43.

6 Liu Y, Wang J. AI-Driven Health Advice: Evaluating the Potential of Large Language Models as Health Assistants. *Journal of Computational Methods in Engineering Applications* 2023; **3**: 1–7. https://doi.org/10.62836/jcmea.v3i1.030106.

7 Luo Z, Yan H, Pan X. Optimizing Transformer Models for Resource-Constrained Environments: A Study on Model Compression Techniques. *Journal of Computational Methods in Engineering Applications* 2023; 1–12. DOI: 10.62836/jcmea.v3i1.030107.

8 Yan H, Shao D. Enhancing Transformer Training Efficiency with Dynamic Dropout. *arXiv*2024, arXiv:2411.03236.

9 Cui W, Zhang J, Li Z, Lopez D, Das' K, Malin B, Kumar S. Evaluating and Improving Generation Consistency of Large Language Models via A Divide-Conquer-Reasoning Approach. 2024. Available online: https://iclr.cc/Conferences/2024/AuthorGuide (accessed on 4 November 2024).

10 Cui W, Zhang J, Li Z, Sun H, Lopez D, Das K, Malin B, Kumar S. Phaseevo: Towards Unified in-Context Prompt Optimization for Large Language Models. *arXiv*2024, arXiv:2402.11347.

11 Zhang J, Cui W, Huang Y, Das K, Kumar S. Synthetic Knowledge Ingestion: Towards Knowledge Refinement and Injection for Enhancing Large Language Models. *arXiv*2024, arXiv:2410.09629.

12 Ma J, Chen X. Fingerprint Image Generation Based on Attention-Based Deep Generative Adversarial Networks and Its Application in Deep Siamese Matching Model Security Validation. *Journal of Computational Methods in Engineering Applications* 2024; **4**: 1–13.

13 Zhang H, Zhu D, Gan Y, *et al*. End-to-End Learning-Based Study on the Mamba-ECANet Model for Data Security Intrusion Detection. *Journal of Information, Technology and Policy* 2024; **2**: 1–17.

14 Gan Y, Zhu D. The Research on Intelligent News Advertisement Recommendation Algorithm Based on Prompt Learning in End-to-End Large Language Model Architecture. *Innovations in Applied Engineering and Technology* 2024; **3**: 1–19.

15 Gan Y, Ma J, Xu K. Enhanced E-Commerce Sales Forecasting Using EEMD-Integrated LSTM Deep Learning Model. *Journal of Computational Methods in Engineering Applications* 2023; **3**: 1–11.

16 Pes B. Learning from High-Dimensional and Class-Imbalanced Datasets Using Random Forests. *Information* 2021; **12**: 286.

17 Erfani SM, Rajasegarar S, Karunasekera S, *et al*. High-Dimensional and Large-Scale Anomaly Detection Using a Linear One-Class SVM with Deep Learning. *Pattern Recognition* 2016; **58**: 121–134.

18 Wagner M, Blom J. The Reciprocal and Non-Linear Relationship of Sustainability and Financial Performance. *Business Ethics: A European Review* 2011; **20**: 418–432.

19 Naimi AI, Balzer LB. Stacked Generalization: an Introduction to Super Learning. *European Journal of Epidemiology* 2018; **33**: 459–464.

20 Tang Y, Li C. Exploring the Factors of Supply Chain Concentration in Chinese A-Share Listed Enterprises. *Journal of Computational Methods in Engineering Applications* 2023; **3**: 1–17.

21 Tang CY, Li C. Examining the Factors of Corporate Frauds in Chinese A-share Listed Enterprises. *OAJRC Social Science* 2023; **4**: 63–77.