Journal of Computer Technology and Applied Mathematics
Vol. 1, No. 1, 2024

Ensemble Fusion: Optimizing Market Prediction with Neural Networks, Residual Networks and Xgboost

ZHU, Mengran [1*] ZHANG, Ye [2] ZHANG, Xinyu [3]

[1] Miami University, USA
[2] University of Pittsburgh, USA
[3] North Carolina State University, USA

**** ZHU, Mengran is the corresponding author, E-mail: mengran.zhu0504@gmail.com***

Abstract: In the realm of financial markets, accurate prediction of market trends plays a pivotal role in guiding investment decisions and maximizing returns. This paper presents an innovative ensemble model that combines neural networks (NN), residual networks (Resnet), and Xgboost, offering a comprehensive approach to market prediction. Through extensive experimentation and evaluation, our ensemble model demonstrates remarkable performance enhancements over individual models and other ensemble configurations. By integrating the predictive strengths of NN, Resnet, and Xgboost, our ensemble achieves significant improvements in predictive accuracy, underscoring the potential of ensemble learning in refining market prediction strategies and empowering traders and investors with enhanced decision-making capabilities. This research contributes to advancing the field of quantitative trading by providing a robust and effective framework for market prediction, offering insights and opportunities for practitioners to navigate the complexities of financial markets with greater confidence and success.

Keywords: Quantitative Trading, Market Prediction, Ensemble Learning, Neural Networks, Resnet, Xgboost.

DOI: https://doi.org/10.5281/zenodo.11060009

1 Introduction

The financial domain is an arena of relentless dynamism, characterized by incessant fluctuations that present both challenges and opportunities for investors and traders. In this intricate milieu, the proficiency to forecast market directions with accuracy is not merely advantageous but pivotal for the formulation of strategic decisions and the realization of trading acumen. The integration of machine learning and sophisticated data analytics into this field marks a significant stride towards refining predictive models, aiming to furnish stakeholders with a nuanced understanding and a strategic edge in this competitive sector.

This scholarly endeavor delves into the potential of ensemble learning within the financial market prediction sphere, a domain where the amalgamation of diverse predictive models can offer a holistic and robust forecasting mechanism. Ensemble learning, by virtue of its methodology, amalgamates the predictive prowess of varied models to curtail the limitations inherent in singular approaches and amplify overall predictive accuracy. The research at hand ventures into an innovative ensemble model that synergistically merges neural networks, residual networks, and gradient boosting frameworks, each acclaimed for its distinct capabilities in capturing market intricacies.

Contrary to isolated investigations of neural networks, residual networks, or gradient boosting methodologies in financial forecasts, this study pioneers an integrative approach, postulating that the confluence of these technologies can unlock superior predictive performance. This ensemble model is envisioned to leverage neural networks for their aptitude in identifying intricate data patterns, utilize residual networks for their robustness in deep learning architectures, and harness gradient boosting for its adeptness in deciphering nonlinear dependencies, thereby presenting a multifaceted approach to market prediction.

2 Related Work

The field of market prediction has witnessed extensive research and development, with scholars and practitioners exploring various methodologies to enhance predictive accuracy and robustness. In this section, we provide a brief overview of related work in the domain of quantitative trading and market prediction.

ZC Lipton et al. [1] provides an overview of recurrent neural networks (RNNs) and their applications in sequence learning tasks, including time series prediction in financial markets.Y Bengio et al.[2]discuss deep learning techniques for unsupervised and transfer learning, highlighting their relevance in financial data analysis.K He et.al;. [3]

introduces residual networks (Resnets) and demonstrates their effectiveness in training deep neural networks, inspiring applications in various domains, including finance.

T Chen et al.[4] present Xgboost, a scalable implementation of gradient boosting machines, and discuss its applications in predictive modeling, including financial forecasting.ZH Zhou [5] provides an in-depth exploration of ensemble learning methods, including bagging, boosting, and stacking, with applications in diverse domains, including finance.This survey paper reviews the application of deep learning techniques, such as convolutional neural networks (CNNs) and RNNs, in time series forecasting tasks, including stock price prediction [6].Yao et al.[7] introduce NDC-Scene, improving monocular 3D Semantic Scene Completion.

A Géron [8] provides practical insights into machine learning techniques using Scikit-Learn and TensorFlow, with examples applicable to financial data analysis and prediction.This classic textbook covers a wide range of machine learning algorithms, including ensemble methods, and their theoretical underpinnings, with applications in various fields, including finance [9].R Caruana et al.[10] present an empirical comparison of various supervised learning algorithms, including decision trees, random forests, and gradient boosting, highlighting their performance in predictive modeling tasks.Yao et al.[11] propose AerialLaneNet, a deep learning method for mapping lanes from aerial images, with superior performance.

S Hochreiter et al. [12] introduces long short-term memory (LSTM) networks, a type of RNN architecture capable of capturing long-range dependencies, with applications in sequential data analysis, including financial time series prediction.L Breiman et al.[13] introduces the random forest algorithm, a popular ensemble learning method based on decision trees, and discusses its applications in classification and regression tasks, including financial prediction. S Ruderet al. [14] provides an overview of gradient descent optimization algorithms, including stochastic gradient descent (SGD) and its variants, and discusses their applications in training neural networks for financial forecasting.Wu et al. [15] show that visually pleasing outdoor environments improve thermal comfort, aiding sustainability in building design.

RE Schapire et al.[16] introduces AdaBoost, a popular boosting algorithm, and discusses its theoretical properties and applications in ensemble learning, including financial prediction tasks.EF Fama et al.[17] proposes the Fama-French three-factor model, which considers market risk, size, and value factors in explaining stock returns, informing quantitative trading strategies and risk management techniques.Q Peng [18] proposes a dual-augmentor framework that significantly improves domain generalization in 3D human pose estimation through two distinct augmentors and meta-optimization techniques. This comprehensive textbook provides a detailed introduction to deep learning techniques, including CNNs, RNNs, and deep reinforcement learning, with applications in various domains, including finance [19].

Yao et al. [20] compare Transformers and CNNs for depth estimation, introducing the DGR module and optimal transport theory for enhancement. Hiu et al. [21]introduces a novel Deep Reinforcement Learning (DRL) approach for efficient mobile robot path planning.Ru et al.[22] propose RA*, a trajectory planning algorithm for AUVs in near-bottom cruising, ensuring safety in complex underwater environments with constraints.Yao, J., et al. [23] propose NDC-Scene for monocular 3D Semantic Scene Completion, showing superior performance.Yao, J., et al.[24] propose AerialLaneNet for lane-level map extraction from aerial images, showing superior performance.

Previous research in quantitative trading has laid the foundation for the development of sophisticated predictive models, leveraging techniques ranging from neural networks to ensemble learning. While individual models have demonstrated promising results, our study seeks to explore the synergies of combining diverse methodologies to further enhance predictive performance in market prediction tasks.

3 Quantitative Analysis and Feature Engineering

3.1 Data Preprocessing and Exploration

The initial phase of our analysis focused on the preprocessing of the dataset, wherein missing data points, particularly within the `resp` attribute indicating future price fluctuations, were imputed with mean values, as shown in Figure 1. This approach facilitated a consistent dataset, crucial for the integrity of subsequent analytical processes.

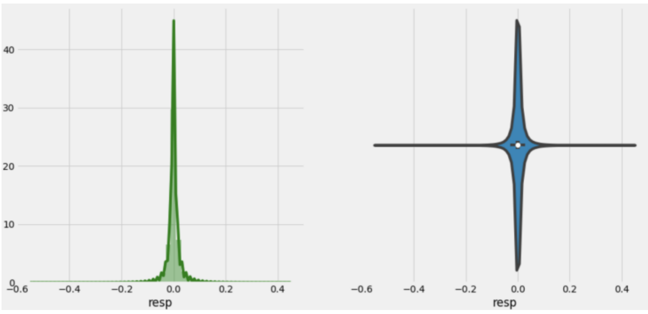

Figure 1 Distribution of resp

3.2 Feature Analysis

An in-depth examination of the dataset's features was conducted to discern underlying patterns and potential anomalies. A notable observation was the prevalence of instances with a zero value in the weight attribute, rendering them non-contributory to the model's calculations, as shown in Figure 2. These instances were retained during training to maintain dataset authenticity yet excluded during testing phases to enhance model accuracy. Additionally, the variability in feature means and their distributions necessitated a comprehensive feature set for model training, underscoring the complexity of the financial data under investigation, as shown in Figure 3.

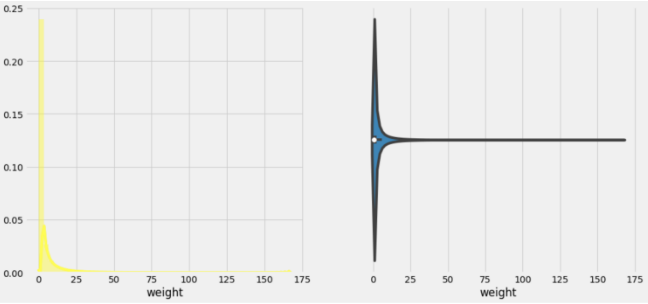

Figure 2 Distribution of weight

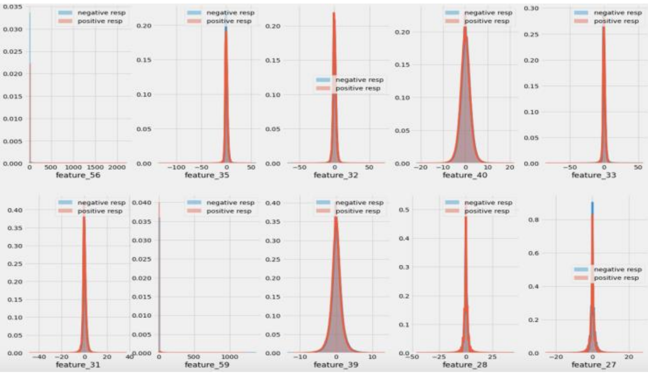

Figure 3 Distribution of other features

3.3 Temporal Analysis and Noise Reduction

The dataset's temporal dimension was scrutinized to identify significant patterns and correlations over time, essential for predictive modeling in financial contexts. A focused analysis on white noise within the time series data aimed to segregate random fluctuations from meaningful trends, thereby refining the model's forecasting capability, as shown in Figure 4.

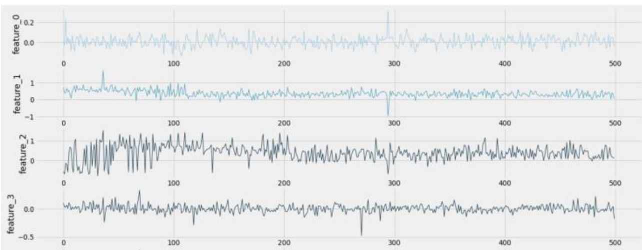

Figure 4 The noise of features

3.4 Dimensionality Reduction through PCA

Principal Component Analysis (PCA) was employed as a strategic measure to reduce the dataset's dimensionality while retaining critical informational content. The PCA results indicated a high degree of correlation among certain features, suggesting the potential for data representation with fewer principal components, as shown in Figure 5. This dimensionality reduction not only simplifies the model but also aids in mitigating the risk of overfitting, thereby enhancing model generalizability.

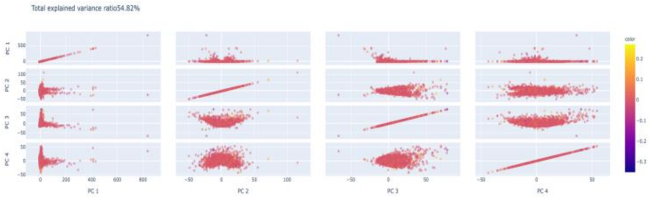

Figure 5 Total explained variance ratio

4 Methodology

The ensemble model we propose is a sophisticated integration of three advanced predictive models: a Neural Network (NN), a Residual Network (ResNet), and eXtreme Gradient Boosting (XGBoost). This amalgamation harnesses the distinctive strengths of each model, yielding a composite model with superior predictive accuracy and robustness. The ensemble's architecture is schematically depicted in Figure 6, illustrating the synergistic combination of these models to formulate a consolidated predictive output.

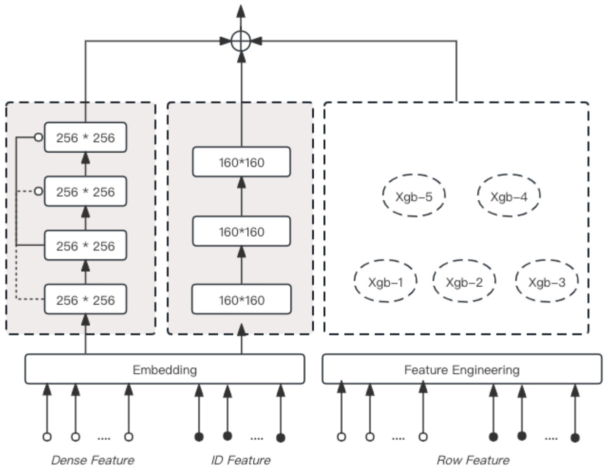

Figure 6. Schematic diagram of the ensemble model architecture

4.1 Neural Network Ensemble

The Neural Network (NN) Ensemble is a sophisticated aggregation framework that combines the outputs of multiple neural network models to enhance predictive accuracy and robustness. This approach leverages the diversity in neural network architectures, where each model, NN_i, is uniquely designed to unravel different patterns and features from the data, thus providing a multifaceted perspective on the problem at hand.

Let M denote the total number of neural networks within the ensemble. Each neural network, NN_i , is trained independently on the dataset, allowing it to develop its own hypothesis of the underlying data structure. Upon receiving an input x, each model within the ensemble generates a prediction $\widehat{y_i}(x)$ reflective of its learned patterns and associations.

The ensemble's prediction, $\hat{y}_{\mathrm{NN\ ensemble}}(\mathrm{x})$, is not a mere average but a weighted sum of all individual model predictions. The formulation is as follows:

$$\hat{\boldsymbol{y}}_{NN\ ensemble}(\boldsymbol{x}) = \sum_{i=1}^{M} \boldsymbol{w_i} * \hat{\boldsymbol{y}}_{\boldsymbol{i}}(\boldsymbol{x}) \#(\mathbf{1})$$

In this equation, w_i represents the weight assigned to the i-th neural network's prediction. These weights are not arbitrary; they are meticulously optimized through techniques such as cross-validation to ensure that each model's contribution is proportional to its predictive reliability and accuracy. The optimization process aims to maximize the ensemble's overall performance, mitigating the impact of any individual model's biases or overfitting tendencies.

4.2 NN + ResNet Ensemble

The NN + ResNet Ensemble represents a synergistic integration of two powerful architectures, each designed to exploit different levels of data abstraction. The ensemble merges the capabilities of Neural Networks (NN) and Residual Networks (ResNet), offering a composite model that benefits from the unique strengths of each.

The Neural Network component, $f_{NN}(x)$, is adept at capturing shallow or surface-level patterns in the dataset. These patterns, often linear or near-linear, provide a baseline understanding of the data's structure. NN models are typically quick to train and can offer rapid insights into the dataset's primary characteristics.

Conversely, the ResNet component, $f_{ResNet}(x)$, is designed to unearth deeper, more complex patterns. ResNet achieves this through its innovative use of skip connections, which mitigate the vanishing gradient problem, allowing the network to learn richer, more nuanced representations of the data. This capability makes ResNet particularly valuable for tasks where the data encompasses intricate hierarchical or non-linear relationships.

The ensemble prediction, $\hat{y}_{NN\ +\ ResNet}(x)$, is not a simple combination but a strategically weighted sum of the outputs from both the NN and ResNet modules:

$$\hat{y}_{NN\ +\ ResNet}(x) = \alpha * f_{NN}(x) + (1 - \boldsymbol{\alpha}) * f_{ResNet}(x) \#(2)$$

Here, α is a critical hyperparameter that determines the relative influence of the NN and ResNet components within the ensemble. Its value is optimized through an exhaustive search, such as grid search or randomized search, where various α values are tested, and the one yielding the best performance on a validation set is selected.

This optimization process is crucial as it ensures that the ensemble leverages the predictive strengths of both NN and ResNet in a balanced manner. The objective is to combine the rapid, efficient pattern recognition of the NN with the deep, comprehensive feature extraction of the ResNet, resulting in a model that is both accurate and robust, capable of handling a wide variety of data scenarios.

4.3 NN + ResNet + XGBoost Ensemble

Building on the foundation laid by the Neural Network and ResNet models, the NN + ResNet + XGBoost Ensemble introduces an additional layer of sophistication by incorporating the XGBoost algorithm, renowned for its effectiveness in sequential decision-making and handling of tabular data. This ensemble is the culmination of our integrative approach, designed to exploit the distinct advantages of each constituent model: the NN's adeptness at capturing surface-level patterns, the ResNet's proficiency in extracting deeper, more abstract features, and XGBoost's strategic sequential decision-making that enhances the model's interpretability and effectiveness in diverse scenarios.

The predictive output of this comprehensive ensemble, $\hat{y}_{NN\ +\ ResNet\ +\ XGB}(x)$, synthesizes the individual predictions from the NN, ResNet, and XGBoost models. The integration is mathematically represented as a weighted sum of these predictions:

$$\hat{\boldsymbol{y}}(\boldsymbol{x}) = \boldsymbol{\beta} * \boldsymbol{f}_{\boldsymbol{NN}}(\boldsymbol{x}) + \boldsymbol{\gamma} * \boldsymbol{f}_{\boldsymbol{ResNet}}(\boldsymbol{x}) + (\mathbf{1} - \boldsymbol{\beta} - \boldsymbol{\gamma}) * \boldsymbol{f}_{\boldsymbol{XGBoost}}(\boldsymbol{x}) \#(\mathbf{3})$$

In this formulation, β and $\boldsymbol{\gamma}$ are critical hyperparameters that modulate the influence of the NN and ResNet outputs, respectively, within the ensemble. The optimization of these parameters is meticulously conducted through cross-validation techniques to ascertain the optimal mix that maximizes the ensemble's predictive performance while ensuring robustness and reliability across varying market conditions.

This ensemble represents an ambitious endeavor to amalgamate diverse modeling techniques, each with its proven strengths, into a unified, comprehensive predictive tool. The strategic integration of NN, ResNet, and XGBoost models aims to provide a holistic view of the data, capturing patterns and relationships at multiple levels of abstraction. By harmonizing these varied approaches, the NN + ResNet + XGBoost Ensemble aspires to set a new benchmark in predictive accuracy, offering nuanced insights and robust forecasts that are crucial for informed decision-making in the realm of quantitative trading.

5 Metrics

The evaluation metric employed to assess the trading model's performance encapsulates both profitability and stability, pivotal for gauging effectiveness in the test dataset's simulated trading scenario.

The daily profit, $\boldsymbol{D_i}$, is given by:

$$\boldsymbol{D_i} = \sum \boldsymbol{weight_i} * \boldsymbol{resp_i} * \boldsymbol{action_i} \#(\mathbf{4})$$

Where $weight_i$ denotes the weight of the stock, $resp_i$ represents the return, and $action_i$ is a binary variable indicating the execution of a trade.

To quantify the model's stability, we calculate the utility score, U, as follows:

$$\mathbf{U} = \mathbf{min}(\mathbf{max}(\mathbf{t}, \mathbf{0}), \mathbf{6}) * \sum \mathbf{p_i} \#(\mathbf{5})$$

Where p_i is the daily profit. The term t is the Sharpe ratio, formulated to assess the risk-adjusted return:

$$\mathbf{t} = \frac{\sum \mathbf{p_i}}{\sqrt{\sum \mathbf{p_i^2}}} * \sqrt{\frac{\mathbf{250}}{|\mathbf{i}|}} \#(\mathbf{6})$$

where p_i denotes the daily profit and |i| represents the number of unique trading days in the test dataset. The factor 250 adjusts for the annualized rate, assuming 250 trading days in a year.

This utility score, U, merges the total profit with a penalty for volatility, emphasizing the model's consistent performance over time. A higher utility score signifies not only profitability but also stability, indicating the model's robustness across diverse market conditions.

6 Experimental Results

The effectiveness of different models in refining market prediction strategies was rigorously assessed, with results summarized in Table 1. Each model's performance was evaluated based on both public and private scores, providing comprehensive insights into their predictive capabilities in real-world trading scenarios.

TABLE 1. Experimental Results

Model	Public	Private
NN Ensemble	11428.04	11012.23
NN + Resnet Ensemble	11531.85	11113.98
NN +Resent+Xgboost Ensemble	11602.23	11498.92

The experimental results highlight the effectiveness of ensemble learning techniques in refining market prediction strategies. While individual models such as neural networks and residual networks demonstrate promising performance, the ensemble of multiple models outperforms them, underscoring the importance of diversity in model architectures. The NN + Resnet + Xgboost Ensemble model, in particular, showcases the potential of integrating multiple algorithms to achieve superior predictive accuracy, paving the way for more robust and reliable models in quantitative trading.

7 Conclusion

In conclusion, our study presents an innovative ensemble model that combines neural networks, residual networks, and Xgboost to enhance market prediction accuracy. Through rigorous experimentation, we've demonstrated significant performance improvements over individual models and other ensemble configurations. By integrating diverse methodologies, our ensemble approach outperforms individual models, emphasizing the importance of model diversity. Particularly, the NN + Resnet + Xgboost Ensemble model showcases the potential of integrating multiple algorithms for superior predictive accuracy, offering a robust framework for quantitative trading. Our findings contribute to advancing the field by providing traders and investors with reliable predictive models, facilitating confident decision-making in navigating financial markets.

Acknowledgments

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

Funding

Not applicable.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author Contributions

Not applicable.

SUAS Press

About the Authors

ZHU, Mengran

Miami University, USA.

ZHANG, Ye

University of Pittsburgh, USA.

ZHANG, Xinyu

North Carolina State University, USA.

References

[1] Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A critical review of recurrent neural networks for sequence learning. arXiv preprint arXiv:1506.00019.

[2] Bengio, Y. (2012, June). Deep learning of representations for unsupervised and transfer learning. In Proceedings of ICML workshop on unsupervised and transfer learning (pp. 17-36). JMLR Workshop and Conference Proceedings.

[3] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[4] Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).

[5] Zhou, Z. H. (2012). Ensemble methods: foundations and algorithms. CRC press.

[6] Torres, J. F., Hadjout, D., Sebaa, A., Martínez-Álvarez, F., & Troncoso, A. (2021). Deep learning for time series forecasting: a survey. Big Data, 9(1), 3-21.

[7] Yao, J., Li, C., Sun, K., Cai, Y., Li, H., Ouyang, W., & Li, H. (2023, October). Ndc-scene: Boost monocular 3d semantic scene completion in normalized device coordinates space. In 2023 IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 9421-9431). IEEE Computer Society.

[8] Géron, A. (2022). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. " O'Reilly Media, Inc.".

[9] Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: springer.

[10] Caruana, R., & Niculescu-Mizil, A. (2006, June). An empirical comparison of supervised learning algorithms. In Proceedings of the 23rd international conference on Machine learning (pp. 161-168).

[11] Yao, J., Pan, X., Wu, T., & Zhang, X. (2024, April). Building lane-level maps from aerial images. In ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 3890-3894). IEEE.

[12] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[13] Breiman, L. (2001). Random forests. Machine learning, 45, 5-32.

[14] Ruder, S. (2016). An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747.

[15] Wu, C., Cui, J., Xu, X., & Song, D. (2023). The influence of virtual environment on thermal perception: physical reaction and subjective thermal perception on outdoor scenarios in virtual reality. International Journal of Biometeorology, 67(8), 1291-1301.

[16] Schapire, R. E. (1990). The strength of weak learnability. Machine learning, 5, 197-227.

[17] Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of financial economics, 33(1), 3-56.

[18] Peng, Q., Zheng, C., & Chen, C. (2024). A Dual-Augmentor Framework for Domain Generalization in 3D Human Pose Estimation. arXiv preprint arXiv:2403.11310.

[19] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

[20] Yao, J., Wu, T., & Zhang, X. (2023). Improving depth gradient continuity in transformers: A comparative study on monocular depth estimation with cnn. arXiv preprint arXiv:2308.08333.[17] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

[21] Liu, H., Shen, Y., Yu, S., Gao, Z., & Wu, T. (2024). Deep Reinforcement Learning for Mobile Robot Path Planning. arXiv preprint arXiv:2404.06974.

[22] Ru, J., Yu, H., Liu, H., Liu, J., Zhang, X., & Xu, H. (2022). A Bounded Near-Bottom Cruise Trajectory Planning Algorithm for Underwater Vehicles. Journal of Marine Science and Engineering, 11(1), 7.

[23] Yao, J., Li, C., Sun, K., Cai, Y., Li, H., Ouyang, W., & Li, H. (2023, October). Ndc-scene: Boost monocular 3d semantic scene completion in normalized device coordinates space. In 2023 IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 9421-9431). IEEE Computer Society.

[24] Yao, J., Pan, X., Wu, T., & Zhang, X. (2024, April). Building lane-level maps from aerial images. In ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 3890-3894). IEEE.