Dean&Francis

ISSN 2959-6130

FE005662

Comparative study of the effects of simple and complex neural network models on enterprise performance

Zijie Gu

Shanghai Jiao Tong University Education Group Innovation Education Research Institute, Shanghai, 201201, China
E-mail: 2481440128@qq.com

Abstract:
Overseas studies have shown that complex neural network models perform better in enterprise performance prediction and can accurately capture the complex relationship of data, but there are long training time, overfitting risk, and limited practical application. Research has gradually focused on balancing model complexity and practicality, while the domestic discussion of the difference in the impact of the two is less, mostly focusing on application validation. In specific industries (e.g., traditional manufacturing), simple models have met the demand due to their high efficiency and transparency; however, in high-dimensional and large-scale data scenarios (e.g., science and technology finance), complex models have significant advantages. In the future, it is necessary to combine the industry characteristics and data scale, deepen the model applicability research, and build a differentiated application framework to optimize the enterprise technology selection and resource allocation efficiency.

Keywords: *Simple model, complex model, comparative research, enterprise performance, predictive ability*

1. Introduction

Enterprise performance is a key indicator to measure the survival and development of an enterprise, and plays a vital role in assessing the operation of an enterprise and enhancing its competitiveness. As a powerful predictive and analytical tool, neural network models have been widely used in the field of enterprise performance assessment. By comparing the performance of simple and complex neural network models in terms of their impact on enterprise performance, it can help enterprises to choose the most suitable model, so as to improve the accuracy and practicability of decision-making, optimize the allocation of resources, and promote the sustainable development of enterprises. This study aims to explore the advantages and disadvantages of simple and complex neural network models in enterprise performance prediction, and provide decision-making suggestions with reference value for enterprise practice.

ISSN 2959-6130

2. Advantages of simple neural network models in enterprises

Due to its simplicity and efficiency, the simple model can help some enterprises accelerate the process of performance analysis in specific scenarios, help enterprises better understand the forecast demand and improve competitiveness, and have an important impact on the development and decision-making of enterprises [1]. For example, the linear regression model analyzes the impact of promotional activities on the sales of chain stores, and the results show that the average daily sales will increase by 5.1% with each increase of “full-price activities”. Since the weights of the model are clear and transparent, the management can intuitively adjust the marketing strategy to optimize the operational effect.

2.1 Comparability and effectiveness

Simple models have fewer parameters, such as the input and output layers of a single-layer network, which makes their calculations and adjustments more accurate and avoids “variable anomalies” that may be caused by too many parameters in complex models [2]. Comparing a single-layer perceptron with logistic regression, for example, the specific effect of weights in a linear model can be clearly observed, without the need to consider the interference of hidden layers or complex structures. Simple models usually require only a small number of hyperparameters to be adjusted, such as the learning rate and regularization coefficients, which makes the experimental design more focused and the experimental results more reproducible. When exploring the effect of activation functions on performance, simple models can directly reflect the differences without having to consider complex issues such as gradient vanishing, which is common in deep networks [3]. When dealing with task scenarios with smaller data volumes, lower feature dimensions and linear separability, the simple model shows efficiency, while its deployment cost is lower and the decision-making process is more transparent.

2.2 Generalization ability in limited cases

The reason why the simple model shows stronger generalization ability lies in the good match between its model complexity and data size, as well as its effective overfitting suppression mechanism [4]. When the amount of data is limited, it is easier for the simple model to satisfy the condition of “model capacity ≤ amount of data information”, thus avoiding the loss of generalization ability due to overfitting noise. For example, in a task with only 100 samples, the upper bound of the generalization error of the single-layer perceptron is significantly lower than that of the deep network. This is due to the fact that the simple model has sparser parameters, fewer parameters, and fewer degrees of freedom to be fitted, and is therefore less sensitive to noise in the training data. The simple model is also more effective in the application of strategies such as L2 regularization, due to the lower spatial dimensionality of its parameters, which allows the regularization constraints to work more directly. When the sample size is insufficient (e.g., $N < 100$), the simple model is able to quickly converge to the global optimal solution without falling into a local optimum as in the case of the complex model, thus reducing the impact of stochasticity. For example, in medical image analysis, if there are only tens of cases of labeled data for a certain rare disease, the generalization performance of a shallow MLP may be better than that of a complex model such as ResNet, which reflects the stability of the simple model in small data scenarios. Simple models rely on global weight updates and are less sensitive to local noise (e.g., labeling errors, feature outliers, etc.). For example, in the presence of 10% random noise in industrial sensor data, the prediction error fluctuation range of the single-layer regression model is more than 30% smaller than that of LSTM. These properties allow the simple model to exhibit superior generalization ability in specific situations.

3. Advantages of complex neural network models in enterprises

With sufficient data volume, complex models tend to show superior performance, and their accuracy continues to improve as the data volume increases. For example, Deep Matrix Factorization (DeepMF) improves recommendation click-through rates by more than 30% when dealing with billions of user behavior data. The core advantage of complex models is their ability to solve nonlinear and high-dimensional problems that are difficult for traditional models to cope with, so as to maximize the use of big data resources and achieve accurate prediction and automation. This is of great strategic significance for the development of enterprises.

3.1 Feature extraction and recognition capabilities

Compared to traditional extraction methods, complex models can automatically extract key features directly from raw data, eliminating many tedious intermediate steps. For example, Convolutional Neural Networks (CNNs) automatically capture features such as edges, textures, and local shapes through convolutional kernels, while Transformer utilizes self-attention to discover semantic associations in text (e.g., recognizing the different meanings of “Xiaomi” in the terms of “crop” and “mobile phone brand”) [5]. In addition, the complex model can be dynamically adjusted according to the task goal, and extract task-oriented features from dynamic features. Through layer-by-layer nonlinear transformation, the complex model can gradually abstract features from low-order to high-order to form a multi-granularity representation. Taking image processing as an example, the example layers are as follows: shallow - edges, corner points; medium - local shapes (e.g. wheels, windows); deep - global semantics (e.g. “car”, ‘building’). Complex models are able to approximate arbitrarily complex functions through multi-layer nonlinear transformations, thus solving pattern problems that cannot be handled by linear models. For example, the different-or (XOR) problem, which cannot be solved by a single-layer perceptron, can be easily fitted by a two-layer network. In addition, CNN achieves translation invariance through convolutional kernel parameter sharing, while Transformer handles sequence order through position encoding [6]. These properties make complex models perform particularly well when dealing with high-dimensional, unstructured data and complex pattern recognition tasks.

3.2 Model capacity under big data

Complex models show significant capacity advantages in big data processing, which is mainly attributed to their deep nonlinear structure and large parameter space, enabling them to efficiently fit and generalize high-dimensional and nonlinear data distributions. Model Capacity refers to the ability of a model to learn complex functional relationships. Complex models (e.g., Deep Neural Networks, Transformer, etc.) are able to capture multi-granularity information ranging from microscopic local features to macroscopic global semantics through the multi-layer stacked hidden layer structure and dynamic parameter adjustment mechanism [7]. In the big data environment, this high-capacity feature forms a virtuous circle with the data scale: on the one hand, a large amount of data provides sufficient training samples for complex models, which effectively mitigates the risk of overfitting, and enables the models to explore the parameter space while still maintaining a good generalization ability. On the other hand, the high capacity of complex models enables them to automatically extract and fuse deep features of heterogeneous data from multiple sources. For example, Convolutional Neural Networks (CNNs) are able to learn features such as edges and textures layer by layer from high-level image data, and ultimately realize end-to-end image classification or target detection without relying on manually designed feature engineering [8]. The high-capacity advantage of complex models is mainly reflected in the following three aspects: first, it breaks through the traditional model’s dependence on linearly divisible or low-dimensional data, and solves the problem of intelligent processing of unstructured data such as images, speech, and text [9]; second, through end-to-end learning, it realizes the optimization of the decision output from the original data, which reduces the manual intervention and the accumulation of errors; lastly, it helps to build scalable intelligent systems, e.g., in recommender systems, the Deep Matrix Factorization (DeepMF) model continuously improves the recommendation effect under massive data by capturing the higher-order interactions between users-items.

4. The ison of different degree models in enterprise performance data

Simple models (e.g., linear regression, shallow neural networks) excel at capturing linear relationships, such as the negative correlation between gearing and return on assets (ROA). With smaller data volumes and fewer features, such models perform consistently, but struggle to effectively model nonlinear effects, such as the threshold effect of R&D investment. In contrast, complex models (e.g., Long Short-Term Memory Network LSTM, Transformer, etc.) are able to capture high-dimensional nonlinear relationships, such as industry crossover features. With large data volumes, such models have significantly higher prediction accuracy, with errors (RMSE) that can be reduced by 15-30%. However, complex models are prone to overfitting on small sample data. In terms of computational cost, there are also differences between the two types of models: simple models are fast to train (usually in seconds) and less expensive to deploy, while complex models are more expensive to deploy and require more maintenance.

4.1 Effect of simple models on performance data

In the test prediction, the RMSE of the simple model on the test set was 3.21, which was significantly higher than that of the complex model (2.58), but better than that of the traditional linear regression (3.8). After multiple cross-validations, the standard deviation of prediction error of the simple model is 0.4, which is lower than that of the complex model (0.7), which indicates that the simple model has a lower risk of overfitting. By parsing the model weights, the key drivers of the simple model on firm performance are derived as follows. Meanwhile, the

ISSN 2959-6130

simple model is found to clearly quantify the negative impact of financial leverage (gearing) on ROA, which is consistent with financial management theory. However, the model is weak in capturing non-financial features and non-linear relationships (e.g., the threshold effect of R&D investment). For example, when the R&D investment exceeds hundreds of millions of dollars, the ROA growth rate increases, but the simple model is still calculated by linear weights, which indicates its inadequacy in dealing with nonlinear relationships. In addition, when modeling uniformly across industries, the generalization ability of the model is weak. These are some of the limitations of the simple model.

4.2 Impact of complex models on performance data

In the same test prediction, the RMSE of the complex model on the test set is 2.58, which is 19.6% lower compared to the 3.21 of the simple model, especially in the technology and finance industries. By capturing the nonlinear features, it is found that: when the R&D investment exceeds 500 million RMB, the ROA growth rate increases significantly (showing increasing marginal returns); while when the gearing ratio exceeds 65%, the ROA decreases sharply (the nonlinear inflection point is revealed). In addition, complex models show significant advantages in high-dimensional and nonlinear data processing, but at the same time rely on a large amount of data and computational resources. Analyzing the characteristic contributions of complex models through SHAP (SHapley Additive exPlanations), the following conclusions are drawn: i. In terms of revenue growth rate (average SHAP value +1.8), the positive contribution of high-growth enterprises is significant and exponential; ii. In terms of gearing (average SHAP value -1.5), the negative impact is sharply magnified when it exceeds the critical value (65%) the negative effect is sharply amplified; iii. in R&D investment (average SHAP value +1.2), the marginal benefit of R&D investment in the technology industry is nonlinearly increasing; iv. in the technology industry (average SHAP value +0.9), there is an additional gain to ROA (policy dividend + technological barriers). At the same time, the complex model also reveals interaction effects that cannot be captured by the simple model, such as the synergistic gain of "technology industry × high R&D investment". In addition, complex models are more sensitive to extreme values, e.g., firms with high gearing ratios (>75%) have significantly lower ROA prediction errors. However, the implementation of complex models requires a large number of samples to avoid overfitting and the reliance on tools such as SHAP/LIME leads to higher interpretation costs, which are some of the limitations of complex models.

4.3 Comparison of findings between simple and complex models on performance data

Table 1 compares the two models and extracts the relevant findings

trade	Revenue growth rate of (%)	asset-liability ratio (%)	R & D investment (100 million yuan)	Asset scale (RMB 100 million yuan)	ROA(%)	Using simple models(0/1)	Using complex models(0/1)
science and technology	15.2	42.3	4.5	65.0	9.1	0	1
manufacture	-2.5	68.9	1.2	120.0	-1.8	1	0
finance	22.1	35.0	0.8	80.0	11.5	0	1
retail	5.5	58.0	0.3	15.0	3.2	1	0
science and technology	28.4	30.5	6.8	200.0	14.7	0	1
manufacture	-7.0	72.0	2.1	90.0	-4.5	1	0
finance	3.8	55.0	1.5	45.0	4.0	1	0
science and technology	18.9	25.0	8.2	150.0	12.3	0	1
retail	9.5	50.0	0.6	25.0	5.5	1	0
manufacture	-1.2	65.0	3.0	80.0	-0.5	1	0

The findings show that complex models are prevalent in the Technology and Finance industries, which have higher revenue growth rates, higher R&D investments, and superior return on assets (ROA). In contrast, the Manufacturing and Retail sectors are more likely to use simple models, with weaker or less volatile revenue growth and generally negative ROA. The tabular data shows that R&D investment is significantly higher for companies using complex models, suggesting that technology-intensive companies rely more on complex models to process high-dimensional data. On the other hand, companies using simple models have lower R&D investment, which may be related to their relatively traditional business model.

5. Conclusion

Compared to simple shallow networks, complex neural network models (e.g., LSTM) exhibit higher accuracy (~20% reduction in RMSE) in predicting business performance (e.g., ROA), with particular advantages in high-dimensional nonlinear scenarios, such as technology and finance. However, complex models are significantly more computationally expensive (3x longer training time) and less interpretable, often relying on a posteriori tools (e.g., SHAP) to parse their decision logic. In contrast, simple models perform better in terms of efficiency (training in seconds) and interpretability (can be directly analyzed in terms of weights), and are more suitable for SMEs with limited data or in need of fast business decisions. Therefore, in practical applications, it is recommended that enterprises weigh their choices based on data size, industry characteristics and resource conditions - preferring complex models in large-scale complex scenarios, and simple models in small-scale or explanatory-driven scenarios, in order to achieve the optimal balance between technical utility and cost.

References

[1] Fu, J., & Zhang, L. (2012). Performance evaluation of circular economy green marketing based on AHP and BP neural network model. Science and Technology Management Research, (20), 215–220+242.

[2] Xu, X., Zhou, C., Hu, Z., Lin, S., & Yu, Z. (2024). A survey on lightweight deep neural network model adaptation for edge intelligence. Computer Science, (7), 257–271.

[3] Li, H. (2022). Research and application of image classification based on convolutional neural network [Master's thesis, Hangzhou Dianzi University]. CNKI. https://link.cnki.net/doi/10.27075/d.cnki.ghzdc.2022.000484
DOI: https://doi.org/10.27075/d.cnki.ghzdc.2022.000484

[4] Yi, M. Y. (2022). Machine learning generalization theory [Doctoral dissertation, University of Chinese Academy of Sciences]. Wanfang Data. https://d.wanfangdata.com.cn/thesis/ChhUaGVzaXNOZXdTMjAyNDA5MjAxNTE3MjUSCFk0MDI0OTkxGgh5ejhzaTE2aA%3D%3D

[5] Wang, D., Xu, Y., Li, H., & Hao, Z. (2019). Convolution kernel normalization. Computer Technology and Development, (12), 27–32.

[6] Sun, Y. L. (2024). Computer vision from entry to advanced practical combat. Chemical Industry Press.

[7] Qiu, D. P. (2024). Research and examples of hidden layer structure based on BP neural network. Changjiang Information and Communications, (7), 8–10. https://doi.org/10.20153/j.issn.2096-9759.2024.07.003

[8] Xie, J. Y., & Zhang, J. Y. (2024). A review of graph convolutional neural networks. Journal of Shaanxi Normal University (Natural Science Edition), (2), 89–101. https://doi.org/10.15983/j.cnki.jsnu.2024003

[9] Batexi, Q., Huang, H., & Wang, X. H. (2015). Uyghur speech recognition based on deep neural networks. Computer Engineering and Design, (8), 2239–2244. https://doi.org/10.16208/j.issn1000-7024.2015.08.045