Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

CUTTING THROUGH THE NOISE: MACHINE LEARNING PROXIES FOR HIGH DIMENSIONAL NESTED SIMULATION

Xintong Li

Department of Statistics and Actuarial Science
University of Waterloo
200 University Ave W
Waterloo, ON N2L 3G1, CANADA

ABSTRACT

Deep learning models have gained great success in many applications, but their adoption in financial and actuarial applications have been received by regulators with trepidation. The lack of transparency and interpretability of these models raises skepticism about their resilience and reliability, which are important factors for financial stability and insurance benefit fulfillment. In this study, we use stochastic simulation as a data generator to examine deep learning models under controlled settings. Our study shows interesting findings in fundamental questions like “What do deep learning models learn from noisy data?” and “How well do they learn from noisy data?”. Based on our findings, we propose an efficient nested simulation procedure that uses deep learning models as proxies to estimate tail risk measures of hedging errors for variable annuities. The proposed procedure uses deep learning to concentrate simulation budget on tail scenarios while maintaining transparency in estimation.

1 INTRODUCTION

Machine learning models, particularly deep learning models (Hastie et al. 2009; LeCun et al. 2015), have attracted attentions of researchers and practitioners due to the successes in solving real-world tasks such as AlphaGo (Silver et al. 2016) and ChatGPT (OpenAI 2023). Two specialized neural network architectures that are relevant to our study are recurrent neural networks (RNNs) (Williams and Zipser 1989; Sutskever et al. 2014) and long-short-term memory (LSTM) (Hochreiter and Schmidhuber 1997; Chung et al. 2014), as we need to train proxy models that take sequential observations as input.

Besides designing specialized neural network architectures and training algorithms, one branch of neural network research considers the effects of noise in the input data and the error tolerance of neural networks. For example, Luo et al. (2016) showed that adding noise to the input of a convolutional neural network (CNN) can increase the effective receptive field of the network and improve its ability to capture global features. Srivastava et al. (2014) quantified the error tolerance by injecting noise with a custom Boltzmann machine hardware. Yang et al. (2018) studied the error tolerance of a CNN on input image captured under a low-voltage setting. The aforementioned studies use real-world data, as is typically the case for many deep learning studies, where noise is already present in the data. Users of real-world data have little control over the noise level and usually examine the effect of noisy data by *injecting noise*. It is unclear whether the deep learning model trained on noisy data actually learns the real, i.e., noiseless, feature-label relationship. Due to their lack of transparency and interpretability, the adoption of deep learning models in financial and actuarial applications has been received by regulators with some skepticism.

The contributions of our study are two-fold:

1. We propose a two-stage nested simulation procedure that uses deep learning model to improve its efficiency without losing transparency. In essence, a deep learning proxy model is used in the first

stage to identify a set of potential tail scenarios on which computations are performed in the second stage. Our numerical results show that deep learning proxy models can identify the tail scenarios accurately and so the proposed procedure can estimate tail risk measures with similar accuracy while, at the same time, using less simulation budget.

2. We study what deep learning models learn from noisy data by training them using simulated data based on well-designed simulation experiments. This is a novel way to study the effect of noisy data and error tolerance of deep learning models as one can *reduce noise* in the data by increasing the number of replications in a simulation model. This new way of studying deep learning models can provide more direct evidence on their transparency and interpretability.

We are curious about fundamental questions like "What do deep learning models learn from noisy data?" and "How well do deep learning models learn from noisy data?". Data-driven answers to these questions prevail in the existing literature. For example, deep supervised learning models are believed to learn from the given data about the feature-label relationship to predict new labels for unseen features. Cross validation using to assess a subset, i.e., the validation set, of the original data, is a common way to access the quality of learning. Generalization error in the test set is another popular assessment metric. But the test set is also a subset of the original data. Instead of relying solely on real-data (splitting it into multiple subsets), we propose using stochastic simulation outputs as training data for deep learning models. By controlling the simulation design parameters, such as the number of independent replications, we can control the quality (and also the quantity) of the data fed into the deep learning model. In such a controlled environment, we obtain more clear-cut answers to the above fundamental questions.

2 CONCLUSION

The proposed two-stage nested simulation procedure with a machine learning proxy model results in substantial computational savings in estimating CVaR of the hedging loss of a VA contract from accurately identifying the tail scenarios. An LSTM fueled by noisy observation pairs cuts through the noise and learns the true inner simulation model. When new outer scenarios are generated, a trained machine learning proxy model can distinguish between tail and non-tail scenarios without the need to run new inner simulations.

REFERENCES

Chung, J., C. Gulcehre, K. Cho, and Y. Bengio. 2014. "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling". https://arxiv.org/abs/1412.3555. Accessed 7th Sep 2023.

Hastie, T., R. Tibshirani, and J. H. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Volume 2. Springer.

Hochreiter, S., and J. Schmidhuber. 1997. "Long Short-Term Memory". *Neural Computation* 9(8):1735–1780.

LeCun, Y., Y. Bengio, and G. E. Hinton. 2015. "Deep Learning". *Nature* 521(7553):436–444.

Luo, W., Y. Li, R. Urtasun, and R. Zemel. 2016. "Understanding the Effective Receptive Field in Deep Convolutional Neural Networks". *Advances in Neural Information Processing Systems* 29.

OpenAI 2023. "ChatGPT". https://chat.openai.com/chat. Accessed 7th Sep 2023.

Silver, D., A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot et al. 2016. "Mastering the Game of Go with Deep Neural Networks and Tree Search". *Nature* 529(7587):484–489.

Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. 2014. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting". *The Journal of Machine Learning Research* 15(1):1929–1958.

Sutskever, I., O. Vinyals, and Q. V. Le. 2014. "Sequence to Sequence Learning with Neural Networks". In *Advances in Neural Information Processing Systems*, edited by Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, Volume 27: Curran Associates, Inc.

Williams, R. J., and D. Zipser. 1989. "A Learning Algorithm for Continually Running Fully Recurrent Neural Networks". *Neural Computation* 1(2):270–280.

Yang, L., D. Bankman, B. Moons, M. Verhelst, and B. Murmann. 2018. "Bit Error Tolerance of a CIFAR-10 Binarized Convolutional Neural Network Processor". In *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1–5. Institute of Electrical and Electronics Engineers.