# The BP Neural Network with Adam Optimizer for Predicting Audit Opinions of Listed Companies

Hua-Ping Wu, *Member, IAENG,* Lin Li

***Abstract*—The risk of material misstatement is closely related to the types of audit opinions on the financial statements for listed companies. Consequently, investors or Certified Public Accountants need to predict the audit opinion types of financial statements to make better investment decisions or to conduct audit work. In this paper, an audit opinions prediction model is developed derived from the BP neural network with Adam optimizer by using 14 raw financial data sets of Chinese listed companies which are from the retail and wholesale industry. The results relying on data indicating that the prediction accuracy of the model is 94.05%. It is considerably better than the normal BP neural network at 83.78% in the entire sample. It is useful for predicting audit opinions on financial statements of listed companies.**

***Index Terms*—BP Neural Network, Audit Opinion, Adam Optimizer, Raw Financial Data**

## I. Introduction

The users of accounting information are usually familiar with listed companies through regular reports, such as interim reports, annual reports, prospectuses, and so on. These reports offer valuable advice to investors for making appropriate investment decisions. However, the financial reports of many listed companies could not completely and truly reflect the business performance and financial status since providing false financial statements or some irregularities which including illegal guarantees and pledges. Evidently, it is disadvantageous to the users of accounting information who make financial decisions by using public data. The aim of an audit is to give guarantee that financial statements do not contain material misstatements to protect the interests for the users of financial statements. The types of audit reports can be categorized into standard unqualified opinion and non-standard unqualified suggestion which includes four opinion, namely unqualified one with paragraph, qualified, disclaimer and adverse. According to the accounting supervision report of the 2018 annual report for listed companies issued by the CSRC(China Securities Regulatory Commission), 219 corporations have been issued with non-standard unqualified audit opinions, of which 38 have been issued with disclaimer of opinions, 82 with qualified opinions, and 99 with unqualified opinions with paragraph.

The audit opinion type is closely related to the financial status of a company. Zhang et al. studied the relationship between the audit opinion type of a companys financial statements and the debt characteristics of a Chinese company, which showed that companies with non-unqualified audit opinions have more interest rates and short-term debt[1]. Moalla and Baili studied the correlation between a company's credit rating and audit opinion type, in which it is more likely that a company with a higher credit rating or an optimistic outlook receives an unqualified audit opinion[2]. Therefore, the users of financial statements could judge a companys financial status according to the audit report opinion type. Certified Public Accountants (CPAs) can also establish models to assess whether a material misstatement poses a risk to the financial statements of an audited corporation. This improves the efficiency and effectiveness of audit work.

A few researchers have used statistical algorithms or machine learning algorithms to predict material misstatement risks using the financial statements. In the opinion of Sharma et al., machine learning has developed and paid much attention to audit analysis in recent years. They used a variety of machine learning classification algorithms to detect financial statement fraud, then compared classification techniques and obtained the result based on accuracy level[3]. Stanisic et al. developed a model that predicts audit opinion type using a variety of statistical and machine learning models, respectively, and concluded that it should be used in different conditions according to the advantages presented by each condition[4].

With the emergence of big data and artificial intelligence, deep learning neural networks have been popularly applied in various fields. Many scholars have applied the deep learning neural network model to predict the types of audit opinions or to identify fraud in financial statements. Temponeras used a deep neural network model to predict financial statement fraud in Greek corporations, and the accuracy of the model reached 93.7%[5]. Ma et al. proposed a model derived from PCA(principal component analysis) and the BP(Back Propagation) neural network for corporate financial fraud identification; the experimental results showed that the financial statement fraud identification rate has improved significantly, compared to the discriminant analysis model and logistic regression model. In addition, the BP neural network model can accurately identify financial fraud and ensure a company's financial security[6]. Bahrami et al. believed that auditors' opinions would be influenced by the views of financial statement users. They used an integrated ANN-PSO(Artificial Neural Network-Particle Swarm Optimization) model to predict audit opinions, and the results showed that an ANN-PSO model was more accurate than an ANN model[7].

Most models established by scholars use financial ratios as input values. However, Bao et al. used 28 original financial data sets obtained mainly from financial statements as input values, and created a prediction model using ensemble

Manuscript received Apr. 24th, 2020; revised Mar. 19th, 2021. This work was supported by the National Social Foundation of China under Grant No. 16CGL070.

H.P. Wu is with Accounting R&D Center, Chongqing University of Technology, Chongqing, 400054, China, e-mail: hpwu@cqut.edu.cn.

L. Li is with Accounting R&D Center, Chongqing University of Technology, Chongqing, 400054, China.

learning to detect whether financial fraud exist in financial statements. They considered that the fraud prediction model using the original financial data is more accurate, and the results showed that the accuracy of prediction using financial ratios is inferior compared to using raw financial data[8].

Based on it, in this paper, it puts forward a BP neural network model with Adaptive Moment Estimation (Adam) optimizer. The proposed model is developed to predict the audit opinion type and to identify whether there may be a material misstatement risk in financial statements using raw financial data from Chinese listed companies which are the retail and wholesale industry. As a result, the prediction accuracy of the model is significantly better than that of the traditional BP neural network. It could assist investors or CPAs in making better investment decisions or to conduct audit work more effectively.

## II. Back Propagation Neural Network

The BP neural network was proposed by Rumelhart and McClelland in 1986, who optimized the model by calculating the gradient by back propagation and continuously updated the weights and bias of the input parameters of iteration[9]. The BP neural network is typically made up of an input, hidden and output layers. The hidden one should be composed of one or more layers. The neural network structure with a single hidden one is shown in Figure 1.

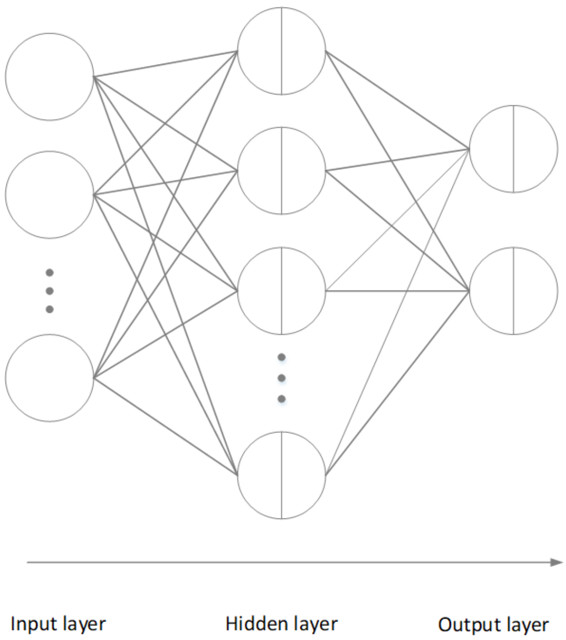


Fig. 1: Three-layer neural network structure

The BP neural network model owns a strong generalization capability and could be well-adapted to the data of any nonlinear structure. However, it also has some disadvantages, such as slow convergence speed, numerous iterations, and the ability to fall into the local optimum easily. If the sample size is too small and the characteristic value of the sample data is too large in the BP neural network model, the trained model can easily produce the phenomenon of overfitting, which leads to the model's good performance in the training set, but not the ideal performance in the test set. In 2012, Hinton proposed the Dropout algorithm to solve the overfitting problem of the BP neural network[10]. The Dropout algorithm reduces the relationship of complex co-adaptation between neurons by randomly deleting some neurons in the neural network, which strengthens the generalization ability of the model. The neural network structure after adding the Dropout algorithm is shown in Figure 2.

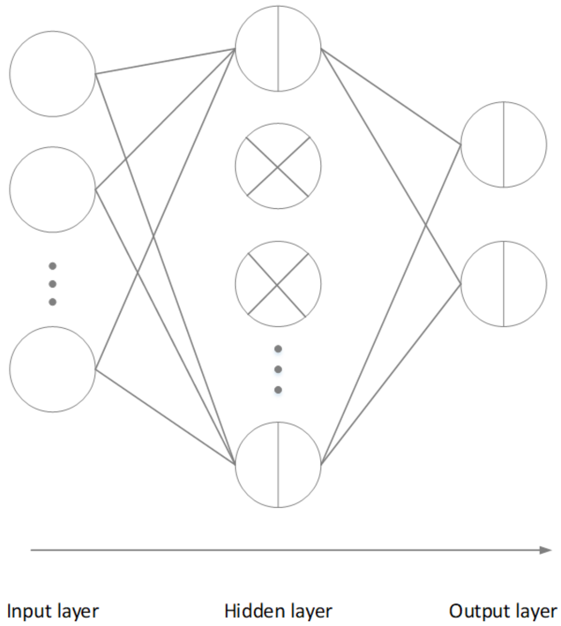


Fig. 2: Three-layer neural network structure with Dropout

The BP neural network uses the gradient descent algorithm(GDA) to optimize the loss function in the process of back propagation.The main procedures of the GDA is shown in Table I.

TABLE I: Main procedures of the GDA

| Gradient descent algorithm: |
|---|
| Input: |
| $w$: Weight of input variables |
| $b$: Bias of input variables |
| $L$: Loss function |
| $\alpha$: Learning rate |
| Steps: |
| 1: Calculate the gradient of the weight and bias of the target loss function for the current variable,$g = \nabla L(w, b)$ |
| 2: Update $w$ and $b$, which $w = w - \alpha \cdot g$, $b = b - \alpha \cdot g$ |
| 3: Repeat steps 1 - 2, until the conditions settled |

The traditional gradient descent algorithm is slow in gradient descent and easily falls into the local optimum, which leads to the unsatisfactory effect of the model. Therefore, in 2014, Kingma and Ba proposed the Adam optimization algorithm[11]. It is based on the adaptive estimation of lower-order moments and is easy to implement,with high computational efficiency and low memory consumption, is suitable for the problem of large data and parameters, is an intuitive interpretation of the hyper-parameters of the algorithm, and usually requires limited adjustment. Based on the traditional gradient descent algorithm, the Adam algorithm increases the first-order and second-order momentum. The first-order momentum is the moving average of the exponential in the gradient direction at each moment, and the second-order momentum is the sum of the squares of all gradient values. The main calculation formula of the iterative process for the Adam algorithm is shown in Equations (1),(2),(3), and (4).

$$m = \beta_1 \cdot m + (1 - \beta_1) \cdot g \tag{1}$$

$$v = \beta_2 \cdot v + (1 - \beta_2) \cdot g^2 \tag{2}$$

$$w = w - \alpha \cdot \frac{m}{\sqrt{v}} \tag{3}$$

$$b = b - \alpha \cdot \frac{m}{\sqrt{v}} \tag{4}$$

where $m$ and $v$ represent the first-order and second-order momentum, respectively, $w$ and $b$ represent the weight and the bias of the input variables, respectively, $g$ represents the gradient of loss function, $\alpha$ represents the learning rate, and$\beta_1$ and $\beta_2$ are hyper-parameters.

## III. Experiment Analysis

### A. Experimental Environment

Experimental platform: Windows 10 operating system, Intel(R) Core(TM) i7-7700@3.20GHz CPU, 16GB memory.

Experimental tools: Python, PyTorch, Scikit-Learn, Pandas, Matplotlib.

The basic process of the experiment is shown in Figure 3.

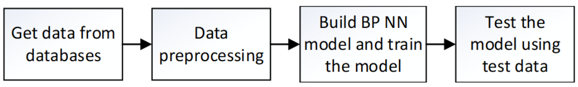


Fig. 3: Basic process of the experiment

### B. Data Selection and Preprocessing

In this paper, the raw financial data and audit report opinions are from all Chinese A-stock listed companies in the retail and wholesale industry (according to the industry classification of the 2012 edition of the CSRC) from year 2013 to 2018 in the CSMAR database, and is a total of 923 items.

The 17 financial indices selected for the variables in the input layer refer to the fraud detection model established by Cecchini et al. in 2010 with 40 raw financial indices[12]. The output variable is the predicted audit opinion, in which the type of unqualified opinion (including standard unqualified one and unqualified one with paragraph) is 0, and the type of non-unqualified opinion which includes qualified, qualified one with paragraph, disclaimer, and adverse is 1.

For improved accuracy of the modeled data, the original data set should be firstly preprocessed. The steps in data preprocessing are as follows:

Step 1: Browse the data set using Pandas and find empty values in the data. Use the Simple Imputer class from Scikit-Learn to substitute the empty values with 0. Replace the value of the type of audit opinion in the data set with 0 for unqualified opinions and 1 for non-unqualified opinions.

Step 2: Due to the significant difference of order of magnitude among each input variable, z-score standardization will be performed for the data to present a standard normal distribution for improved accuracy of the model. The standardized formula is illustrated in Equation 5:

$$x^* = \frac{x - \mu}{\sigma} \tag{5}$$

where $x$ represents the original data, $\sigma$ represents the standard deviation of the data, and$\mu$ represents the mean of the data.

Step 3: Use the train_test_split() method in Scikit-Learn to divide the data set into training and test, in which the training set accounts for 80% of the total data set.

Step 4: Over-sampling of the training set samples. It was found that the number of unqualified opinions and non-unqualified opinions in the total data set totaled 898 and 25, respectively, accounting for 97.29% and 2.71%. It is a serious imbalance. If the number of samples in a certain class is too small, the model will not obtain enough information for training, which may affect the accuracy of the model. Under-sampling or over-sampling techniques are generally used to solve the problem of sample disequilibrium in statistics. Under-sampling is used to achieve sample equilibrium by reducing the number of samples in more categories, while the over-sampling technique is used to achieve balance by increasing the number of samples in fewer categories. Therefore, in the paper, the Synthetic Minority Over-sampling Technique is used to balance the samples. The sample of the training set is balanced by using the Imblearn package in Python. Finally, the number of samples of both unqualified and non-unqualified opinions totaled 716.

Step 5: To select indicators with a strong explanatory ability, the SelectKBest class in Scikit-Learn is used to conduct the chi-square test on variables and output the $P$ value, and 14 variables with $P < 0.05$ are selected as the final input variables. As per Table II,the 17 original financial indices and their $P$ values are selected.

TABLE II: Variable code

| Variable name | P values |
|---|---|
| Monetary capital | 0.00005 |
| Accounts receivable | 0.00088 |
| Inventory | 0.00015 |
| Current assets | 0.00003 |
| Fixed assets | 0.000006 |
| Total assets | 0.00002 |
| Accounts payable | 0.0012 |
| Current liabilities | 0.00003 |
| Long-term loan | 0.0006 |
| Total liabilities | 0.0001 |
| Total owners' equity | 0.00005 |
| Revenue | 0.0005 |
| Cost of sales | 0.011 |
| Profit before tax | 0.12(deleted) |
| Net profit | 0.16(deleted) |
| Cash flow from operating activities | 0.57(deleted) |
| Basic earnings per share | 0.013 |
| Audit opinion | |

### C. Parameter Setting and Model Training of BP Neural Network

In this paper, it uses Pytorch to build the BP neural network model, which is structured as a three-layer neural network. Here, the activation function of the hidden layer is the RELU, the activation one of the output layer is the SoftMax, and the loss one is the CrossEntropy.

The parameters of the model are as follows.

Input layer: the number of neurons is 14, which refers to the number of input original financial indices variables.

Hidden layer: The number of neurons is determined according to the empirical formula, namely, $h = \sqrt{i+o}+a$, where $i$ represents the amount of neurons of the input layer, $o$ represents the amount of neurons of the output layer, and$a$ is an integer from 1 to 10. The BP neural network of the Adam optimizer is used to carry out the loop by calculating the accuracy of the model under each value. The result is shown in Figure 4. It is found that the accuracy of the model is greatest when the number of neurons in the hidden layer is 14. Therefore, the number of neurons is set to 14.

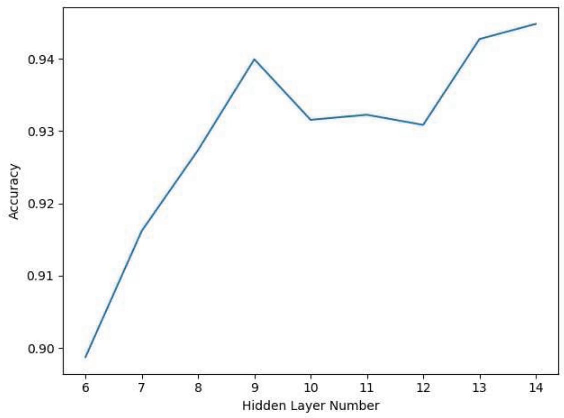


Fig. 4: The number of hidden layer neurons and the accuracy of the model

Output layer: since the output result is a binary classification, the number of neurons is set to 2.

Other parameters: the learning rate is set to 0.1, $\beta_1$ and $\beta_2$ are set to 0.9 and 0.999, the number of model iterations is set to 1000, the dropout probability is set to 0.5, and the global random seed is set to 0.

Then, using Pytorch to build and train the neural network model with the traditional gradient descent and Adam optimization algorithm, respectively. The loss function curve of the model and the accuracy curve of model prediction are obtained, as illustrated in Figures 5 and 6, respectively. It is found that the loss function convergence rate of the Adam optimization algorithm is faster, the model accuracy is higher. Obviously, the prediction effect is better than the traditional gradient descent algorithm.

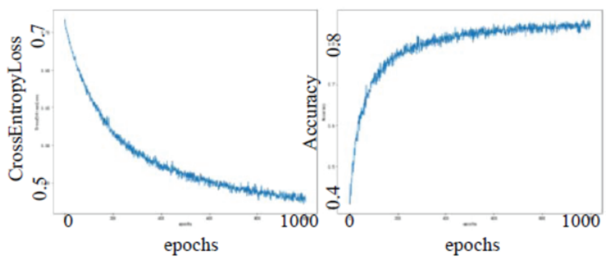


Fig. 5: Loss function and accuracy curve of gradient descent training for 1000 times

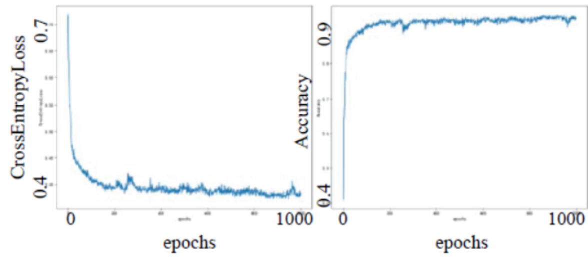


Fig. 6: Loss function and accuracy curve of Adam optimizer training for 1000 times

Finally, the results are presented in Table 3. The prediction results of the BP neural network model trained by the traditional gradient descent algorithm are as follows: the prediction accuracy of the unqualified opinion and non-unqualified opinion is 82.12% and 88.55%, respectively. And the overall sample accuracy is 85.34%. The prediction results of the BP neural network model trained by the Adam optimization algorithm are as follows: the prediction accuracy of the unqualified opinion and non-unqualified opinion is 93.44% and 92.88%, respectively. And the overall accuracy of the sample is 93.16%. Therefore, the training results of the BP neural network model with Adam optimizer are better than one.

TABLE III: Accuracy of model training

| | BP | Adam-BP |
|---|---|---|
| Unqualified opinion | 82.12% | 93.44% |
| Non-unqualified opinion | 88.55% | 92.88% |
| Total | 85.34% | 93.16% |

### D. Test Model with Test Data Set

The test data set is input into the model for prediction to examine the generalization ability of the model. The test data set contains 182 unqualified opinions and 3 non-unqualified ones. The goal of the model is to find the 3 non-unqualified opinions and discover them as few times as possible. Here, the confusion matrix is used to evaluate the model. The confusion matrix is a multi-dimensional evaluation one of two classification problems, which can play a very good effect when evaluating imbalanced samples. The confusion matrix can be divided into the following results according to the true category of the data and the predicted category, as shown in Table 4.

TABLE IV: Confusion matrix

| | | Predicted | |
|---|---|---|---|
| | | Positive | Negative |
| Actual | Positive | TP | FN |
| | Negative | FP | TN |

The confusion matrix is used to obtain three used evaluation indexes, namely precision rate, recall rate and accuracy rate. The precision rate measures the proportion of samples that the model identifies as positive that are actually positive. The recall rate is the proportion of the positive class samples correctly classified in the prediction process to the real positive class samples. The formulas are as follows.

$$precision = \frac{TP}{TP + FP} \tag{6}$$

$$recall = \frac{TP}{TP + FN} \tag{7}$$

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{8}$$

The purpose of the paper is to examine the risk of material misstatement in financial statements. Therefore, the non-unqualified opinions are classified as positive. The confusion matrix of test set is based on the traditional neural network and the Adam optimization neural network. It is as shown in Table 5.

TABLE V: Confusion matrix of test data

| | | | Predicted | |
|---|---|---|---|---|
| | | | Non-Unqualified | Unqualified |
| Actual | Traditional BP | Non-Unqualified | 3 | 0 |
| | | Unqualified | 30 | 152 |
| | Adam-BP | Non-Unqualified | 3 | 0 |
| | | Unqualified | 11 | 171 |

The evaluation indexes of the two models can be calculated from Table 5. They are as shown in Table 6 and Fig 7.

TABLE VI: Comparison of two models

| | Traditional BP | Adam-BP |
|---|---|---|
| precision | 9.10% | 21.43% |
| recall | 100.00% | 100.00% |
| accuracy | 83.78% | 94.05% |

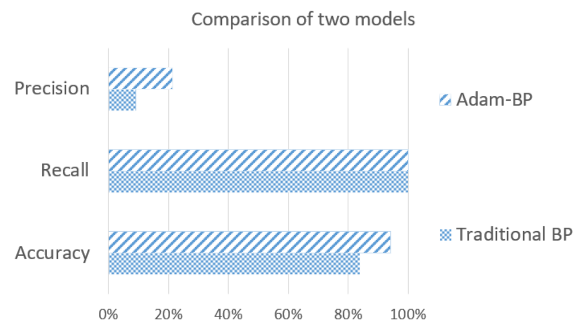


Fig. 7: Comparison of two models

It can be seen that although all the non-unqualified opinions are accurately found (recall rate) in the two models , but in precision rate and total accuracy rate, BP neural network with Adam optimizer performs better than traditional neural network, which means that the BP neural network with Adam optimizer has a better effect.

## IV. Conclusion

According to the above mentioned experiment, there is a stronger correlation between 14 indices and the audit opinion type of financial statements from Chinese wholesale and retail industry listed companies. The prediction accuracy and precision of the BP neural network model with Adam optimizer is higher than the traditional BP neural network model, while the Adam optimizer models convergence speed is faster and has a stronger generalization ability. Therefore, financial statement audit opinion types of Chinese wholesale and retail industry listed companies can be predicted. It provides investors with assistance in decisions, reduces investment risk, aids CPAs to identify risks of material misstatement in the financial statements of listed companies more efficiently, and improves the audit work of CPAs.

## References

[1] J. Zhang, "Modified audit opinions and debt contracting: evidence from china," *Asia Pacific Journal of Accounting & Economics*, vol. 27, no. 2, pp. 218–241, 2020.

[2] R. Baili, "Credit ratings and audit opinion: evidence from tunisia," *Journal of Accounting in Emerging Economies*, vol. 9, no. 1, pp. 103–125, 2019.

[3] A. Sharma, M. Patel, and M. Tiwari, "A comparative study to detect fraud financial statement using data mining and machine learning algorithms," *International Research Journal of Engineering and Technology*, vol. 6, no. 8, pp. 1492–1495, 2019.

[4] N. Stanišić, T. Radojević, and N. Stanić, "Predicting the type of auditor opinion: Statistics, machine learning, or a combination of the two?" *The European Journal of Applied Economics*, vol. 16, no. 2, pp. 1–58, 2019.

[5] G. S. Temponeras, S.-A. N. Alexandropoulos, S. B. Kotsiantis, and M. N. Vrahatis, "Financial fraudulent statements detection through a deep dense artificial neural network," in *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*. IEEE, 2019, pp. 1–5.

[6] X. Ma, X. Li, Y. Song, X. Zheng, Z. Zhang, and R. He, "A bp neural network for identifying corporate financial fraud," in *2019 IEEE International Conference on Intelligence and Security Informatics (ISI)*. IEEE, 2019, pp. 191–193.

[7] F. Bahrami, J. Rezazadeh, and F. Sarraf, "Forecasting audit opinion based on multilevel perceptron neural network model using one-goal particle swarm optimisation," *International Journal of Management Practice*, vol. 13, no. 1, pp. 86–102, 2020.

[8] Y. Bao, B. Ke, B. Li, Y. J. Yu, and J. Zhang, "Detecting accounting fraud in publicly traded us firms using a machine learning approach," *Journal of Accounting Research*, vol. 58, no. 1, pp. 199–235, 2020.

[9] D. E. Rumelhart and J. L. McClelland, "The pdp research group: Parallel distributed processing: Explorations in the microstructure of cognition," *Foundations*, vol. 1, pp. 3–44, 1986.

[10] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv preprint arXiv:1207.0580*, 2012.

[11] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[12] M. Cecchini, H. Aytug, G. J. Koehler, and P. Pathak, "Detecting management fraud in public companies," *Management Science*, vol. 56, no. 7, pp. 1146–1160, 2010.