JURNAL ONLINE INFORMATIKA
JOIN (Jurnal Online Informatika)
Volume 7 No. 2 | December 2022: 168-176
DOI: 10.15575/join.v7i2.840
p-ISSN: 2528-1682
e-ISSN: 2527-9165

Anti-Corruption Disclosure Prediction Using Deep Learning

Victor Gayuh Utomo[1], Tirta Yurista Kumkamdhani[2], Galih Setiarso[3]
[1,2]Department of Information System, Universitas Semarang, Indonesia
[3]Department of Information Technology, Universitas Semarang, Indonesia

Article Info
Article history:
Received December 03, 2021
Revised April 19, 2022
Accepted July 07, 2022
Published December 26, 2022

Keywords:
Anti-corruption Disclosure
Deep learning
Logistic regression
Machine learning

ABSTRACT
Corruption gives major problem to many countries. It gives negative impact to a nation economy. People also realized that corruption comes from two sides, demand from the authority and supply from corporate. On that regard, corporates may have their part in fight against corruption in the form of anti-corruption disclosure (ACD). This study proposes new method of ACD prediction in corporate using deep learning. The data in this study are taken from every companies listed in Indonesia Stock Exchange (IDX) from the year 2017 to 2019. The companies can be categorized in 9 categories and the data set has 8 features. The overall data has 1826 items in which 1032 items are ACD and the other 794 items are non-ACD. In this study, the deep neural network or deep learning is composed from input layer, output layer and 3 hidden layers. The deep neural network uses Adam optimizer with learning rate 0.0010, batch size 16 and epochs 500. The drop out is set to 0.05. The accuracy result from deep learning in predicting ACD is considered good with the average training accuracy is 74.76% and average testing accuracy is 76.37%. However, the loss result isn't good with average training loss and testing loss are respectively 51.76% and 50.96%. Since the aim of the study to find the possibility of deep learning as alternative of logistic regression in ACD prediction, accuracy comparison from deep learning and logistic regression is held. Deep learning has average prediction accuracy of 76.37% is better than logistic regression with average accuracy of 67.15%. Deep learning also has higher minimum accuracy and maximum accuracy compared to logistic regression. This study concludes that deep learning may give alternatives in ACD prediction compared the more common method of logistic regression.

Corresponding Author:
Victor Gayuh Utomo
Department of Information System,
Universitas Semarang,
Jl. Soekarno Hatta, Tlogosari Semarang 50196, Indonesia
Email: victor@usm.ac.id

1. INTRODUCTION

Corruption gives major problem to many countries. It gives negative impact to a nation economy. Study shows that areas with higher corruption index, that shows worse corruption, have worse economic growth [1]. Corruption also gives negative impact on stock market capitalization and number of transactions [2]. The negative impact of corruption not limited in economy only, it even damages the environment [3]. Since corruption is tangled with politic, it is hard to be eliminated [4]. Other than the obvious law enforcement [5], the work against corruption includes implementing better accounting standard [6] and implementing information technology in transaction [7].

However, people also realized that corruption comes from two sides, demand from the authority and supply from corporate [8]. On that regard, corporates may have their part in fight against corruption in the form of anti-corruption disclosure (ACD). ACD includes accounting to combat bribery, board and senior management responsibilities, human resource to combat bribery, responsible business relationship, external verification and assurance, codes of conduct and whistle blowing [9].

Studies show that the corporate adoption of ACD is low [9], [10] despite the fact that ACD may bring positive impact to the corporate financial performance both in short term and long term [11]. Based on this fact, many studies are held to find the determinant factor for a corporate to adopt ACD. These studies hold experiments by creating model with regression method [12]–[14]. While nothing wrong by using regression model, this study propose a new method with machine learning, specifically deep learning.

In machine learning, computer may learn and found useful information from a set of data [15]. In particular, machine learning algorithm uses a set of data called training data set and extracts the information accordingly. After the learning process, machine learning creates a model. This model can be used to process another data. For instance, in supervised learning, machine learning creates a model that able to classify data into a certain label. In order to evaluate the model, machine learning use another data, called testing data set, produces result that compared with ground truth [16]. Interestingly, while machine learning has certain algorithm, it is easy enough to incorporate another method in it, such as fuzzy neural network, a combination of fuzzy algorithm and neural network (algorithm in machine learning) [17]. This study propose deep learning as machine learning method. Deep learning is a subset of machine learning that imitates the works of human brain [18]. There are many studies involving deep learning, for example in pneumonia classification [19], air temperature prediction [20] and text classification [21].

There are several studies in fighting corruption with machine learning. A study in Brazil analyzes budget to predict corruption. The study use machine learning with gradient boost tree algorithm [22]. A study of corruption prediction in Italy also use machine learning. Classification tree algorithm is used to analyze the Italian National Institute of Statistics data [23]. Another study use deep learning to predict corruption by analyze 3 major news media coverage [24]. However, those studies use machine learning in predicting the corruption itself. Study that incorporate ACD with machine learning [25] uses feature selection to find the determinant factor of ACD adoption.

This study proposes new method of ACD prediction in corporate instead of finding the determinant factor. The proposed method is deep learning. Prediction result from deep learning method will be compared with prediction result from regression method. The aim of the comparison is to determine whether deep learning may give alternative to the regression method that is commonly used.

2. MATERIAL & METHOD

In this study of ACD there are two sources for the data. First data source comes from the companies' annual report. Since all the companies are listed in Indonesia Stock Exchange (IDX), the information are available at the IDX official website. The second data source comes from the United Nations Global Compact (UNGC) in the form of the company membership.

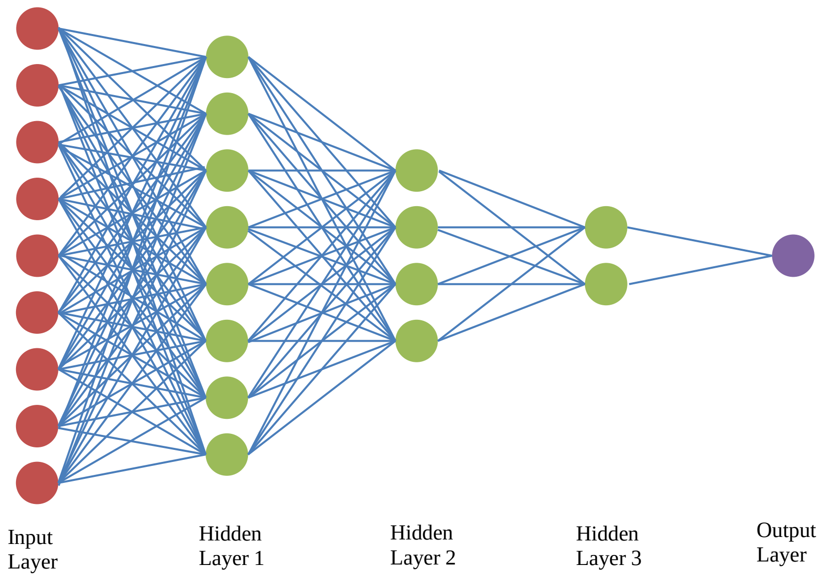

Figure 1. Structure of Deep Learning in This Study

The data are taken from every companies from the year 2017 to 2019. The companies can be categorized in 9 categories, i.e. 1) agriculture, 2) chemical, 3) consumer, 4) finance, 5) infrastructure, 6) mining, 7) property, 8) trading and 9) miscellaneous. The data set has 8 features, i.e. 1) foreign ownership, 2) government ownership, 3) government tender, 4) international operation, 5) independent commissioner, 6) governance committee, 7) big 4 auditor and 8) UNGC membership. The overall data has 1826 items in which 1032 items are ACD and the other 794 items are non-ACD.

Deep neural network or deep learning is composed from input layer, output layer and hidden layer. Multiple hidden layer is the characteristic of deep learning [26]. In this study, there are 3 hidden layer as illustrated in Figure 1. The input layer has 9 neurons. The number of neurons at this layer equals to the number of features in the data. The output layer has 1 neurons that is equals with the number of output. While hidden layer 1 until hidden layer 3 has 8, 4 and 2 neurons respectively.

Deep learning in this study is set to have 500 epoch or iteration. Deep learning also need another configuration called hyperparameters. These hyperparameters include learning rate and batch size. The number of learning rate and batch size are decided with a process called hyperparameters tuning. The tuning is a required process to avoid unwanted condition such as underfit or overfit.

As deep learning is a part of neural network, the research method in this study follows the common method in neural network research. The method in this study is taken from [27] with some adjustment considering the context of the study and can be seen in Figure 2.

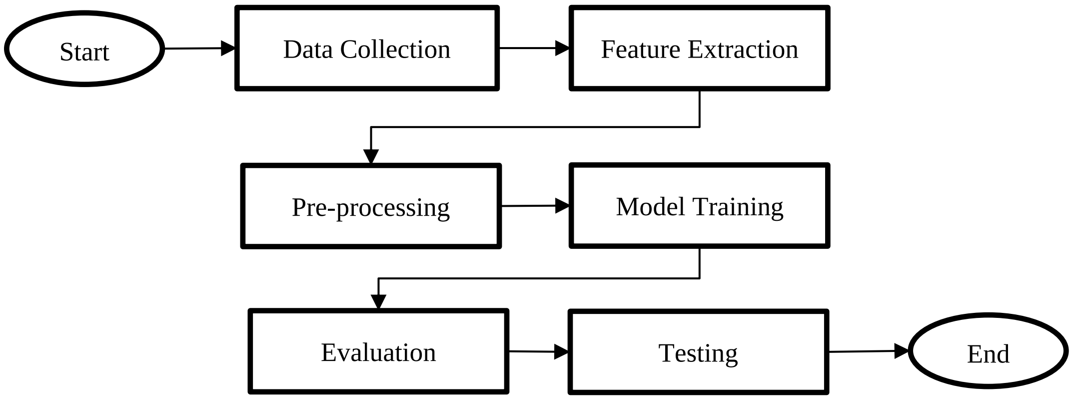

Figure 2. Research Method

Data collection in this study is held by manually compile the annual report from the IDX website and analyze UNGC website. The study rather forced to collect the data manually since the IDX website provides inconsistent download page for the annual report.

Another step that requires manual work is feature extraction. The manual nature of the feature extraction of this study can’t be avoided since most of the required features need manual interpretation of the companies’ annual report.

In data pre-processing step, the study prepares the data so it is ready for the actual machine learning process. There are 2 processes conducted in this step, data cleansing and data normalization. In data cleansing, incomplete data are removed. Data cleansing is applied for two types of data, companies with unpublished annual report and companies with incomplete features. While it is compulsory for listed companies to publish their annual reports, a few companies didn’t publish their annual reports. Hence, the study choose to remove these data. The other type of incomplete data includes companies with certain annual reports that the study unable to extract the necessary features. For example, the companies didn’t state their auditor in their annual reports. In such situation, the study unable to determine whether the companies used big 4 auditor or not. The data normalization process is required since the features is a mixed of feature with numeric value and features with Boolean value.

Model training is held by feed the training data to the configured deep learning algorithm. The exact deep learning configuration is discussed at latter part of this chapter. The model is created after the algorithm trained all the training data. The training data takes 80% of the data. In deep learning algorithm, training process produces metrics such as accuracy and loss.

Evaluation is held right after model training. In evaluation, another part of data that are different from training data, feed to the model created from previous step. This process produces prediction result. The prediction result compared with the actual value or ground truth to create another metrics. The metrics act as the basis to tune the hyperparameters of the deep learning model.

The last step in the method is testing. Testing works just like evaluation by feed data into the created model. The metrics produced in this step are also accuracy and loss. However there are differences. First, that the testing data takes another data that is different with evaluation data. Second, the testing method aims to actually measure the performance of the created model.

This study repeat the whole method for each category data and collects the necessary metrics to measure the performance of the model. Since the aim of the study is to provide alternative in anti-corruption disclosure (ACD) prediction, the model performance compared with performance of regression model. Regression model chosen since the previous studies mostly use this method. In this study, the actual regression used is logistic regression. Logistic regression considered appropriate in this study since the model produces binary value. Another reason is that logistic regression used widely in social sciences, which suitable with ACD, for prediction model [28].

3. RESULTS AND DISCUSSION

After data collection, feature extraction and pre-processing step, the result is machine learning ready data. The summary of the data is presented at Table 1. For each category, the number of data of company that has anti-corruption disclosure compared to the company that has no anti-corruption disclosure is considered balance, at least in machine learning. There are no data that only a small percentage of its counterpart.

Table 1. Data Summary

Categories	Has ACD	No ACD
agriculture	330	200
chemical	1300	800
consumer	760	710
finance	1890	720
infrastructure	980	990
mining	890	440
property	1290	840
trade	2100	2050
miscellaneous	780	550

Since the deep learning in this study use Adam optimizer, hyperparameters tuning should include initial learning rate, β_1, β_2 and weight decay [29]. However, this study limits the tuning process to learning rate and batch size. The dropout parameter is set to a fix number at 0.05 along with the number of epoch set to 500.

Table 2. Evaluation Result

Learning Rate	Batch Size	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
0.0005	16	0.7368	0.527	0.7692	0.5048
0.0005	32	0.7368	0.4864	0.7692	0.487
0.0005	64	0.7171	0.5813	0.7436	0.5527
0.001	16	0.7697	0.4534	0.7692	0.4619
0.001	32	0.7632	0.506	0.7949	0.4983
0.001	64	0.7303	0.5144	0.7692	0.5066
0.002	16	0.7829	0.448	0.7436	0.4984
0.002	32	0.7434	0.4924	0.7692	0.4461
0.002	64	0.625	0.6633	0.5987	0.6786

Hyperparameters tuning is conducted based on the result of model training and evaluation. The tuning process use data from Property category with combination of learning rate 0.0005, 0.010, 0.020 and batch size 16, 32, 64. The result shown at table 2.

Overall, a good deep learning model should have high accuracy and low loss. However, the tuning process also need to address the overfitting or underfitting symptom. If the training has much better performance than the validation then the model shows overfitting tendency and if the training has much worse performance than the validation then the model shows underfitting tendency. The tuning process need to find model with the least difference between training and validation both in accuracy and loss value. To obtain the delta between the training accuracy and validation accuracy this study use the following formula

$$\delta_a = |a_t - a_v|$$

With δ_a is delta accuracy, a_t is training accuracy and a_v is validation accuracy. The formula eliminates the tendency of overfitting or underfitting and focus on similarity between training and validation value. Obtaining the delta between the training loss and validation loss is conducted with the following formula

$$\delta_l = |l_t - l_v|$$

With δ_l is delta loss, l_t is training loss and l_v is validation loss. After calculate delta accuracy and delta loss from table, the result is presented as line chart in Figure 3.

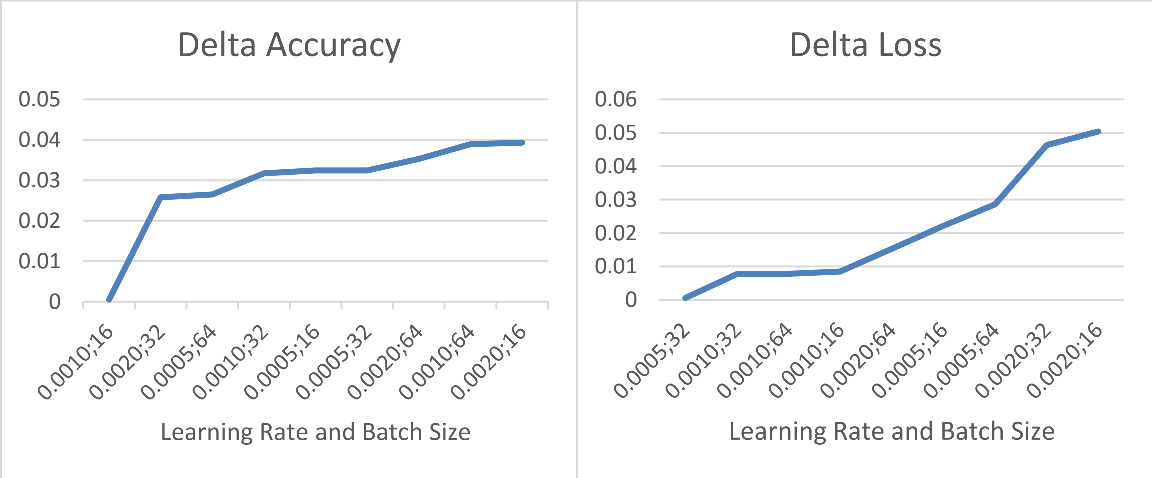

Figure 3. Model Tuning; (a) Delta Accuracy, and (b) Delta Loss

Figure 3a shows delta accuracy with respect to combination of learning rate and batch size while Figure 3b shows delta loss with respect to combination of learning rate and batch size. From Figure 3, the smallest delta accuracy is achieved from learning rate 0.0010 and batch size 16 (0.0010;16) while the smallest delta loss is achieved from (0.0005;32). Since the smallest value from delta accuracy and delta loss comes from different combination the analysis is taken to the next smallest value. After analyzing the four smallest value, there are two combinations to be found, i.e. (0.0010;32) and (0.0010;16). Further analysis shows that both combinations have high testing accuracy, high validation accuracy, low testing loss and low validation loss. However, while both have similar delta loss values, the delta accuracy of (0.0010;16) is smaller than (0.0010;32). After the tuning, the complete hyperparameters values is presented at Table 3.

Table 3. Hyperparameters Value

Hyperparameters	Value
Optimizer	Adam
Learning Rate	0.0010
Batch Size	16
Epochs	500
Dropout	0.05

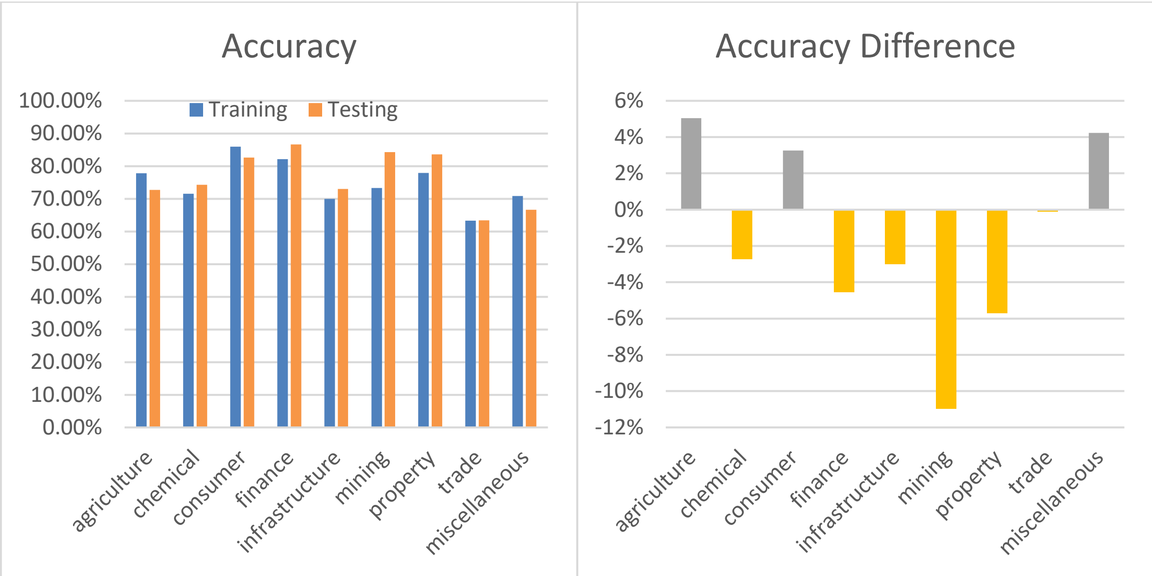

Figure 4. Accuracy after Testing; (a) Accuracy for Each Categories, and (b) Accuracy Difference for Each Categories

The next step in this study is to use the data from all categories for data training and testing. The data are split between training and testing with 80:20 ratio. With the aforementioned hyperparameters value, the accuracy of ACD prediction from each categories are taken and presented at Figure 4 along with accuracy difference. The accuracy difference is calculated with $accuracy_{training} - accuracy_{testing}$.

Figure 4a shows the accuracy value for each categories. For each categories there are two bars, first bar indicates the accuracy from training step and the second bar indicates the accuracy from testing step. Figure 4b shows the difference between training and testing accuracy. Positive value indicates overfitting while negative value indicates underfitting.

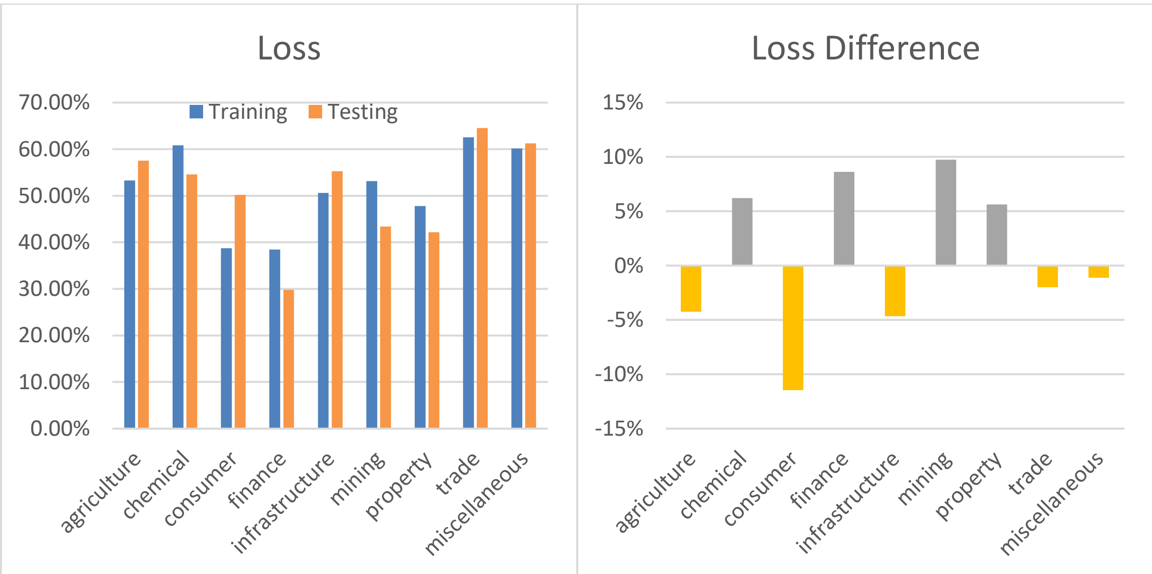

Figure 5. Loss after Testing; (a) Loss for Each Categories, and (b) Loss Difference for Each Categories

From Figure 4a, 8 out of 9 categories managed to achieve around 70% or above accuracy both in training and testing result. The result is considered good. Moreover, the average training accuracy is 74.76% and average testing accuracy is 76.37%. Comparing the average accuracies show that the model is well tuned with only a little sign of undefitting. The accuracy difference, shown at Figure 4b, also supports the result. From 9 categories, 3 categories show little sign of overfitting (with difference between 0 to 5%), 4 show little sign of underfiting (with difference between 0 to -5%), 1 category with almost no difference and 1 category

p-ISSN: 2528-1682
e-ISSN: 2527-9165

with heavy underfitting (around -11% difference). Despite the rather low number of accuracy in Trade category and underfitting in Mining category, this study consider the proposed method of deep learning may predicts ACD with good accuracy.

Like in the tuning analysis, this study also measures the loss metric from both training and testing result. The loss value and the loss differences from each categories are shown at Figure 5. The accuracy difference is calculated with $loss_{training} - loss_{testing}$.

Figure 5a shows the loss value for each categories. For each categories there are two bars, first bar indicates the loss from training step and the second bar indicates the loss from testing step. Figure 5b shows the difference between training and testing loss. Opposed with accuracy difference, in loss difference positive value indicates underfitting while negative value indicates overfitting.

From Figure 5a, only 5 out of 9 categories managed to achieve around 50% or below loss both in training and testing result. The result is considered not good despite the average training loss and testing loss are respectively 51.76% and 50.96%. Comparing the average losses show that the model is well tuned. However, the loss difference, shown at Figure 5b, do not support the result. From 9 categories, 1 category show little sign of underfitting (with difference between 0 to 5%), 3 categories show sign of underfiting (with difference between 5 to 10%), 4 categories show little sign of overfitting (with difference between 0 to -5%) and 1 category shows heavy overfitting (around -11% difference). The average loss shows good number due to the cancelling effect among the categories. This study consider the proposed method of deep learning in ACD prediction has below average loss performance.

This study also compares the performance between deep learning with the logistic regression. The logistic regression is applied to data in every categories. The ratio between training data and testing data is also 80:20. The performance prediction is measured with accuracy metric. The comparison of the accuracy of deep learning and logistic regression shown at Figure 6.

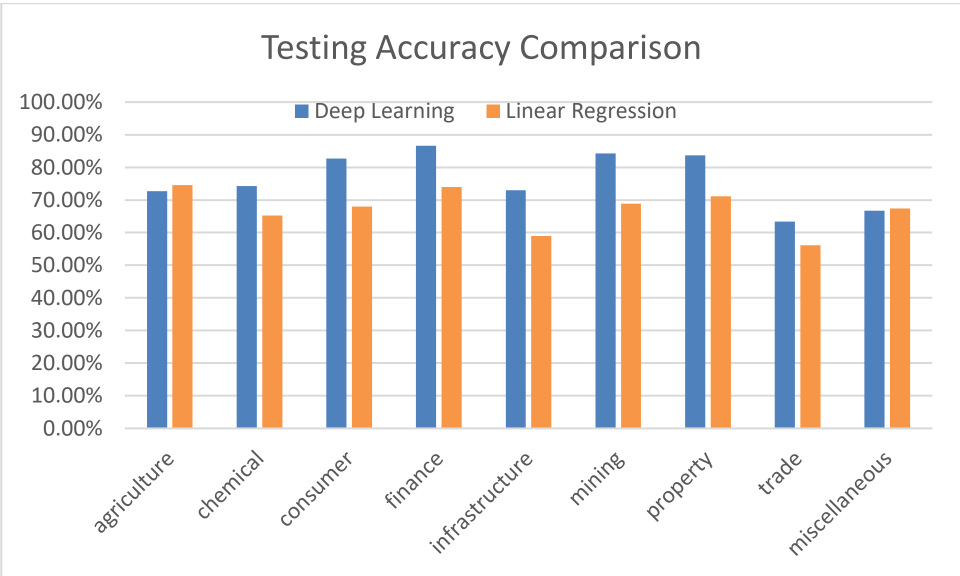

Figure 6. Testing Accuracy Comparison for Each Categories

Figure 6 shows the accuracy value for each categories. For each categories there are two bars, first bar indicates the accuracy from deep learning testing process and the second bar indicates the accuracy from logistic regression testing process. Only accuracy is measured for this comparison since logistic regression has no loss metric.

From Figure 6, deep learning accuracy is better than logistic regression in 7 out of 9 categories. By average, deep learning has accuracy of 76.37% while logistic regression has accuracy of 67.15%. Deep learning has 9.27% advantage in accuracy over logistic regression. The minimum accuracy in deep learning is reached in Trade category at 63.35% and still higher than the minimum accuracy of logistic regression, also from Trade category, at 56.14%. The same pattern also applied at maximum accuracy. The deep learning in Finance category at 86.67% and logistic regression in Agriculture category at 74.55%.

This study finds that in all prediction accuracy comparison, deep learning performs better than logistic regression in ACD prediction. As suggested in introduction, the aim of the study is to find whether deep learning may acts as alternative to logistic regression. In that regard, deep learning may acts as alternative to logistic regression. Not only as an alternative, deep learning is even a better alternative to logistic regression in ACD prediction.

4. CONCLUSION

ACD (Anti-Corruption Disclosure) is important in fighting corruption. Most ACD prediction is held with logistic regression. This study proposes new method of ACD prediction in corporate using deep learning. The data in this study are taken from every companies listed in Indonesia Stock Exchange (IDX) from the year 2017 to 2019. The companies can be categorized in 9 categories, i.e. 1) agriculture, 2) chemical, 3) consumer, 4) finance, 5) infrastructure, 6) mining, 7) property, 8) trading and 9) miscellaneous. The data set has 8 features, i.e. 1) foreign ownership, 2) government ownership, 3) government tender, 4) international operation, 5) independent commissioner, 6) governance committee, 7) big 4 auditor and 8) UNGC membership. The overall data has 1826 items in which 1032 items are ACD and the other 794 items are non-ACD.

In this study, the deep neural network or deep learning is composed from input layer, output layer and 3 hidden layers. The deep neural network uses Adam optimizer with learning rate 0.0010, batch size 16 and epochs 500. The drop out is set to 0.05.

The performance of deep learning in predicting ACD is analyzed with accuracy and loss. In accuracy, deep learning is considered good with the average training accuracy is 74.76% and average testing accuracy is 76.37%. However, the loss result isn't good with average training loss and testing loss are respectively 51.76% and 50.96%.

Since the study aim is to find deep learning as alternative to logistic regression, predictions from both methods are compared. Comparison is held by comparing the ACD prediction accuracy. Average accuracy of deep learning (76.37%) is higher than logistic regression (67.15%). Not only in average, the minimum and maximum accuracy reached by deep learning also higher than logistic regression. In that regard, this study concludes that deep learning may acts as alternative, if not better, to the commonly used logistic regression as ACD prediction method.

5. REFERENCES

[1] A. Alfada, "The destructive effect of corruption on economic growth in Indonesia: A threshold model," *Heliyon*, vol. 5, no. 10, p. e02649, 2019.

[2] I. Missaoui, M. Brahmi, and J. BenRajeb, "Quantitative relationship between corruption and development of the Tunisian stock market," *Public Munic. Financ.*, vol. 7, no. 2, pp. 39–47, 2018.

[3] Y. Liu and F. Dong, "Corruption, economic development and haze pollution: Evidence from 139 global countries," *Sustain.*, vol. 12, no. 9, 2020.

[4] S. Xu, M. Qiao, B. Che, and P. Tong, "Regional anti-corruption and CSR disclosure in a transition economy: The contingent effects of ownership and political connection," *Sustain.*, vol. 11, no. 9, pp. 1–14, 2019.

[5] E. Sundari and A. Retnowati, "THE WEAKNESS OF THE CONTROL SYSTEM FOR FIGHTING CORRUPTION IN THE JUDICIAL PROCESS : THE CASE OF INDONESIA INTERNATIONAL JOURNAL OF SOCIAL POLICY AND LAW (IJOSPL)," *Int. J. Soc. Policy Law*, vol. 02, no. 01, pp. 93–102, 2021.

[6] V. Tawiah, "The impact of IPSAS adoption on corruption in developing countries," *Financ. Account. Manag.*, no. June 2020, pp. 1–22, 2021.

[7] A. Addo and C. Avgerou, "Information Technology and Government Corruption in Developing Countries : Evidence from Ghana Customs INFORMATION TECHNOLOGY AND GOVERNMENT CORRUPTION IN DEVELOPING COUNTRIES : EVIDENCE FROM GHANA CUSTOMS Atta Addo Surrey Business School Chrisanthi Avger," *MIS Q.*, no. January, 2021.

[8] B. Gauthier, J. Goyette, and W. A. K. Kouamé, "Why do firms pay bribes? Evidence on the demand and supply sides of corruption in developing countries," *J. Econ. Behav. Organ.*, vol. 190, no. October, pp. 463–479, 2021.

[9] A. Issa and A. Alleyne, "Corporate disclosure on anti-corruption practice: A study of social responsible companies in the Gulf Cooperation Council," *J. Financ. Crime*, vol. 25, no. 4, pp. 1077–1093, 2018.

[10] H. Nobanee, O. F. Atayah, and C. Mertzanis, "Does anti-corruption disclosure affect banking performance?," *J. Financ. Crime*, vol. 27, no. 4, pp. 1161–1172, 2020.

[11] N. K. Karim, A. Animah, and E. E. Sasanti, "Pengungkapan Anti Korupsi Dan Kinerja Keuangan Perusahaan: Studi Kasus Perusahaan Terdaftar Di Indeks Sri Kehati," *J. Ris. Akunt. Aksioma*, vol. 15, no. 2, p. 28, 2017.

[12] D. Zulvina and D. Adhariani, "Anti-corruption disclosure and firm value: Can female CEOs and CFOs have moderating roles?," *Int. J. Innov. Creat. Chang.*, vol. 10, no. 11, pp. 771–794, 2020.

[13] M. A. K. Masud, S. M. Bae, J. Manzanares, and J. D. Kim, "Board directors' expertise and corporate corruption disclosure: The moderating role of political connections," *Sustain.*, vol. 11, no. 16, 2019.

[14] M. A. Odriozola and I. Á. Etxeberria, "Determinants of corporate anti-corruption disclosure: The case of the emerging economics," *Sustain.*, vol. 13, no. 6, 2021.

[15] G. Suresh, D. A. S. Kumar, D. S. Lekashri, D. R. Manikandan, and C.-O. Head, "Efficient Crop Yield

Recommendation System Using Machine Learning For Digital Farming," *Int. J. Mod. Agric.*, vol. 10, no. 1, p. 2021, 2021.

[16] R. Fiebrink, "Machine learning education for artists, musicians, and other creative practitioners," *ACM Trans. Comput. Educ.*, vol. 19, no. 4, 2019.

[17] F. Abdali-Mohammadi, M. N. Meqdad, and S. Kadry, "Development of an IoT-based and cloud-based disease prediction and diagnosis system for healthcare using machine learning algorithms," *IAES Int. J. Artif. Intell.*, vol. 9, no. 4, pp. 766–771, 2020.

[18] H. Qasim El-Mashharawi, I. A. Alshawwa, and M. Elkahlout, "Grape Type Classification Using Deep Learning," *Int. J. Acad. Eng. Res.*, vol. 3, no. 12, pp. 41–45, 2019.

[19] O. Stephen, M. Sain, U. J. Maduh, and D. U. Jeong, "An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare," *J. Healthc. Eng.*, vol. 2019, 2019.

[20] P. H. Gunawan, D. Munandar, and A. Z. Farabiba, "Long Short-Term Memory Approach for Predicting Air Temperature In Indonesia," *J. Online Inform.*, vol. 5, no. 2, p. 161, 2020.

[21] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep Learning Based Text Classification: A Comprehensive Review," vol. 1, no. 1, pp. 1–43, 2020.

[22] E. Ash, S. Galletta, and T. Giommoni, "A Machine Learning Approach to Analyze and Support Anti-Corruption Policy," *SSRN Electron. J.*, vol. 2020, no. April 2020, pp. 1–34, 2020.

[23] G. De Blasio, A. D. Ignazio, and M. Letta, "Predicting Corruption Crimes with Machine Learning . A Study for the Italian Municipalities," 2020.

[24] A. Weichselbraun, S. Hörler, C. Hauser, and A. Havelka, "Classifying News Media Coverage for Corruption Risks Management with Deep Learning and Web Intelligence," *ACM Int. Conf. Proceeding Ser.*, vol. Part F1625, pp. 54–62, 2020.

[25] V. G. Utomo, R. A. Dewi, D. Marutho, and A. Hidayat, "Analyze Corporate Anti-Corruption Disclosure with Feature Selection," in *2021 International Seminar on Application for Technology of Information and Communication (iSemantic)*, 2021, pp. 234–237.

[26] S. Nosratabadi *et al.*, "Data science in economics: Comprehensive review of advanced machine learning and deep learning methods," *Mathematics*, vol. 8, no. 10, pp. 1–25, 2020.

[27] F. A. Nugraha, N. H. Harani, R. Habibi, and R. N. S. Fatonah, "Sentiment Analysis on Social Distancing and Physical Distancing on Twitter Social Media using Recurrent Neural Network (RNN) Algorithm," *J. Online Inform.*, vol. 5, no. 2, p. 195, 2020.

[28] P. Sur and E. J. Candès, "A modern maximum-likelihood theory for high-dimensional logistic regression," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 116, no. 29, pp. 14516–14525, 2019.

[29] L. N. Smith, "A disciplined approach to neural network hyper-parameters: Part 1 -- learning rate, batch size, momentum, and weight decay," pp. 1–21, 2018.