Deep Learning Model Implementation Using Convolutional Neural Network Algorithm for Default P2P Lending Prediction

Tiara Lailatul Nikmah [1], Jumanto [1], Budi Prasetiyo [1], Nina Fitriani [2], Much Aziz Muslim [1,3]
[1] Universitas Negeri Semarang, Kampus Sekaran - Gunungpati, 50237Semarang, Indonesia
[2] University of Manchester, Oxford Road Manchester, M13 9PR Manchester, United Kingdom
[3] University Tun Hussein Onn Malaysia, Batu Pahat, 86400 Johor, Malaysia

ARTICLE INFO

Article history:

Received June 06, 2023
Revised August 18, 2023
Published August 22, 2023

Keywords:

P2P Lending;
CNN;
Deep Learning;
Default Risk;
Prediction

ABSTRACT

Peer-to-peer (P2P) lending is one of the innovations in the field of fintech that offers microloan services through online channels without intermediaries. P2P lending facilitates the lending and borrowing process between borrowers and lenders, but on the other hand, there is a threat that can harm lenders, namely default. Defaults on P2P lending platforms result in significant losses for lenders and pose a threat to the overall efficiency of the peer-to-peer lending system. So, it is essential to have an understanding of such risk management methods. However, designing feature extractors with very complicated information about borrowers and loan products takes a lot of work. In this study, we present a deep convolutional neural network (CNN) architecture for predicting default in P2P lending, with the goal of extracting features automatically and improving performance. CNN is a deep learning technique for classifying complex information that automatically extracts discriminative features from input data using convolutional operations. The dataset used is the Lending Club dataset from P2P lending platforms in America containing 9,578 data. The results of the model performance evaluation got an accuracy of 85.43%. This study shows reasonably decent results in predicting p2p lending based on CNN. This research is expected to contribute to the development of new methods of deep learning that are more complex and effective in predicting risks on P2P lending platforms.

Corresponding Author:

Jumanto, Universitas Negeri Semarang, Kampus Sekaran - Gunungpati, Semarang 50229, Indonesia
Email: jumanto@mail.unnes.ac.id

1. INTRODUCTION

The most important and rapidly growing innovation in the financial industry today is Financial Technology (Fintech). Fintech refers to industries that use technology to facilitate financial activities, including electronic payments, online loans, digital investments, and other financial services [1]. Fintech has changed how people access and use financial services, and continues to grow rapidly around the world [2]–[4]. Peer-to-peer (P2P) lending is one of the fintech-based innovations. P2P offers micro-lending services through online channels without intermediaries [5]. P2P lending is growing very rapidly, especially in China. In 2016, more than 13 million people invested in P2P lending platforms, up 134% from 2015, China's largest P2P lending community [6].

This lending system gives lenders greater flexibility to select a suitable risk portfolio [7]. However, on P2P lending platforms there are several risks that must be considered, one of which is the risk of default. Default risk can be caused by several things such as high-interest rates, high credit risk, and lack of supervision and regulation [8], [9]. Defaults often result in significant losses for lenders and pose a threat to the overall efficiency of the P2P lending system. Therefore, it is important to understand default risk P2P lending platforms to be able to find methods that can manage these risks and reduce losses significantly.

Many financial sector risk management efforts have been carried out [10], [11]. One of the efforts to manage risk in the field of P2P lending is to predict the risk of default. Predicting the risk of default on P2P loans can be done in various ways, one of which is with deep learning techniques. A deep learning technique for default risk prediction that is currently very popular is to use classification methods. The classification method will classify loans into default and non-default categories.

Many studies have been conducted to create default risk prediction models using deep learning over the past few years [12]–[16]. Some of these studies such as, Kim and Cho (2019) [17] which use Convolutional Neural Network (CNN) to predict P2P social loan payments to extract features automatically and improve performance. In the research conducted by Kim, CNN obtained the best model with an accuracy of 77.78% using the Lending Club dataset. Then, Kim and Cho (2022) [18] classify an ensemble consisting of diverse convolutional neural networks (CNNs) from GoogLeNet, ResNet, and DenseNet for payment prediction in social loans with labels from unexpired data with appropriate labels into training sets can improve prediction model performance. This study obtained the highest accuracy of 82.8%.

Research conducted by Zhang *et al.* (2020) [19] created a credit scoring model using the online integrated credit scoring model (OICSM). This study compares the proposed method, namely OICSM with GDBT, Wide&Deep, DeepFM, GDBT2NN and OICSM-off methods. Based on the comparison of the results of model performance accuracy, OICSM obtained as the best model with 74.67%.

This study proposes a default risk prediction model using the Convolutional Neural Network (CNN) algorithm. CNNs can process large data quickly and efficiently [20], [21]. This is because CNNs can automatically extract important features from visual data without the need to perform manual feature extraction. In addition, CNNs can process data in parallel by using GPUs (Graphics Processing Units), which speeds up training and inference time. CNN effectively recognizes local patterns in data sequences [22]–[28] by using convolution operations on 1-dimensional data, CNNs can extract important features such as temporal patterns, word order, or other sequence patterns. This allows the model to focus on important information in sequence data and ignore irrelevant context. CNNs also have invariances to positions in the data sequence [29]–[34].This means the CNN model can recognize the same pattern at different positions in the data sequence.

In this study, the data used in this study is Lending Club data from P2P platforms in America. The data went through a preprocessing process with missing value handling. After that, it is divided into 2 parts: data training for the CNN model training process and data testing for model testing. The architecture is built using a 1D convolutional layer and a max-pooling layer. In addition, a dropout layer is applied to prevent overfitting. This study is using the Stochastic Gradient Descent (SGD) optimizer which serves to function to optimize model parameters by iterating through training data in small batches randomly.

2. METHODS

In this study, the prediction of P2P loan default risk uses several stages, namely data preprocessing, data sharing, modeling with CNN and model evaluation.

2.1. Data Description

The dataset in this study taken from the Kaggle repository is the Lending Club dataset at the following link: https://www.kaggle.com/datasets/urstrulyvikas/lending-club-loan-data-analysis. The Lending Club dataset is obtained from customer data from an American P2P lending platform called Lending Club. The Lending Club dataset contains 9,578 records with 8,045 non-default loan data and 1,533 default loan data. This dataset contains 14 numerical features.

2.2. Preprocessing Data

Data preprocessing is the initial stage in data analysis that aims to clean, prepare, and transform raw data into a more suitable form for analysis [35]–[39]. At the data preprocessing stage, the data cleaning process is carried out. The process of cleaning data includes checking missing values. A missing value is a value or data that does not exist or has no known value in a data set [40]–[42]. Missing value checks are performed to check for the presence of NaN or Null values in the data. NaN usually appears caused by various factors such as missing data, invalid data, or data processing errors and Null values are usually used to indicate that a value or object is unavailable or not found. Because they potentially impair model performance, columns with a large number of missing values will be removed. It is crucial to handle missing values appropriately when analyzing data so that the results produced are more accurate and valid [43]. The research flowchart is shown in Fig. 1

Vol. 9, No. 3, September 2023, pp. 802-809

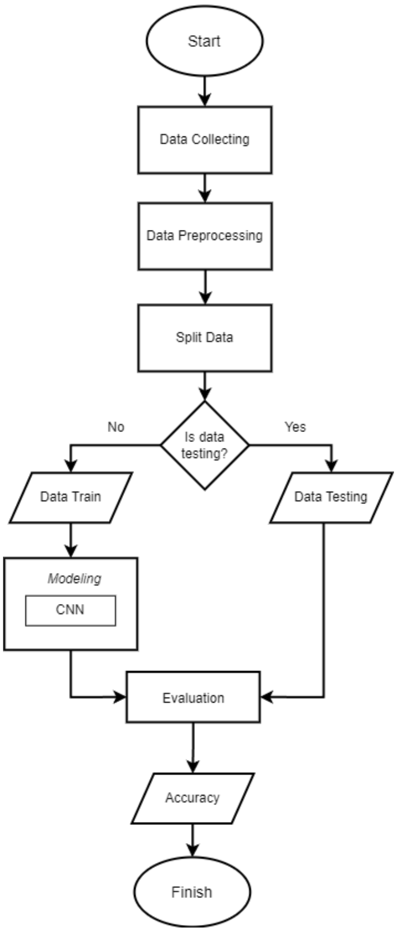

Fig. 1. P2P lending default risk prediction model flowchart

2.3. Split Data

Data that has gone through the data preprocessing stage will then go through the data sharing stage. The data-sharing stage is carried out by the train_test_split method. Data sharing divides the data into 30% testing data and 70% training data with parameters random_state=42.

2.4. Modeling

The CNN algorithm consists of several interconnected layers, such as a convolution layer, an activation layer, a pooling layer, and a weight loss layer. The convolution layer is responsible for extracting features from the image by performing convolution operations against the weight matrix in the input image. The activation layer applies the activation function to the output of the convolution layer to introduce non-linearity.

The following is the architecture of CNN in this study can be seen in Table 1 and the parameters for the training model can be seen in Table 2.

Table 1. CNN Architecture

Layer	Output Shape	Param
Conv1D	(None, 17, 128)	512
MaxPooling1D	(None, 8, 128)	0
Dropout	(None, 8, 128)	0
Conv1D	(None, 6, 128)	49280
MaxPooling1D	(None, 3, 128)	0
Dropout	(None, 3, 128)	0
Flatten	(None, 384)	0
Dense	(None, 128)	49280
Dropout	(None, 128)	0
Dense	(None, 1)	129

Table 2. Parameters Model CNN

Parameters	Value
Optimizer	SGD
Loss	Binary_Crossentropy
Metrics	Accuracy
Batch Size	128, 256
Epoch	10

2.5. 1D Convolutional Layer

This layer is responsible for extracting visual features from the input. Convolution in the convolutional layer involves mathematical operations called shifting and merging [44], [45]. This process involves applying one or more convolution filters to the inputs to generate a new feature map. There are 2 convolution layers, the input will be processed on the first convolution layer which is the 1d convolution layer because the input data used is 1-dimensional data. A 1D convolution layer is a type of convolution layer used in a Convolutional Neural Network (CNN) to process one-dimensional data.

In the 1D convolution layer, the convolution filter moves in only one direction, along the data dimension [46]. Such filters are applied to the inputs by performing convolution operations at each position. This convolution operation involves shifting filters and multiplication of points at each position along the data dimension.

2.6. MaxPooling Layer

MaxPooling Layer performs downsampling operations on the feature map generated by the convolution layer. MaxPooling aims to reduce the number of parameters that must be processed in the network and produce a more concise representation while retaining important information [47]. By reducing dimensions, fewer parameters must be processed in the network, speeding up computations and reducing the need for high computing resources [48].

The MaxPooling layer accepts feature maps as input. This feature map is usually the result of a previous convolution layer that has larger dimensions. MaxPooling layers use fixed-sized windows to scan feature maps. The window is layered on the feature map with specific strides. At each window position, the maximum value within that window is taken which represents the most important value. The MaxPooling operation results in a feature map with smaller dimensions compared to the input feature map.

2.7. Dropout Layer

Deep neural networks generally overfit parameters and tend to overfit. Regularization techniques have been introduced to address this problem. One of the regularization techniques dropouts. The dropout layer receives input from the previous layer in the neural network, which is usually a feature map or activation output [49]. The dropout layer has a parameter called dropout probability or dropout rate [50]. Dropout probability controls the proportion of neuron units to be deactivated during training. The dropout layer will randomly select the neuron units to be deactivated [51]. This is done independently for each node of the training data. It prevents the coadaptation of neurons during training. We set the dropout at probability 0.2 after all convolutional layers.

2.8. SGD Optimizer

SGD serves to optimize model parameters by iterating through training data in small batches at random. SGD uses a stochastic approach where model parameters are updated after each sample or small batch and the concept of Gradient Descent where model parameters are updated by moving a step towards a gradient of derivative loss functions against those parameters [52]. The purpose of SGD is to find the value of a parameter that yields a minimal value of the loss function [53]. The weights and biases in the neural network model are initialized with random values. For each training iteration, SGD randomly selects from a small subset of training data (batches) from the entire training dataset. The training data in the batch is passed through the neural network in advance. After forward propagation, the loss value is calculated based on the difference between the output produced by the network and the expected target. After calculating the loss, the loss gradient to the weight and bias is calculated using the backpropagation method. Weight and bias are updated using a previously calculated gradient. SGD updates by reducing the gradient with the learning rate, the learning rate used is 0.01. This process continues over several iterations until it reaches the specified maximum number of iterations.

.

3. RESULTS AND DISCUSSION

The study was tested with the Lending Club dataset which has 14 numerical features. In the Lending Club dataset, missing value data was not found after checking the missing value at the data preprocessing stage. Next, the data goes into the data sharing stage dividing 30% of testing data and 70% of training data. Training data is used to train the model. The model is built using the CNN algorithm. The construction of this model uses a model imported from the Keras library. The results of the CNN model training can be seen in Table 3.

Table 3. CNN Model Training Results

Epoch	Accuracy	Loss
1	**0.8543**	**0.4160**
2	0.8332	0.4523
3	0.8339	0.4505
4	0.8453	0.4301
5	0.8383	0.4450
6	0.8331	0.4505
7	0.8383	0.4437
8	0.8500	0.4227
9	0.8422	0.4369
10	0.8328	0.4522

The evaluation results in Table 3 show the accuracy of the CNN model that has been carried out as many as 10 epochs. The highest accuracy was obtained at epoch 1 with an accuracy of 85.43% and the lowest loss value was obtained from epoch 1 with a total of 0.4160%. This high accuracy result is obtained because convolution allows the extraction of local patterns from data sequences. Convolution layers can recognize the same pattern at different positions in the data sequence. For example, a convolution layer can recognize the same sequence pattern even though the pattern occurs at different positions in the data sequence. This helps in custom pattern recognition and improves model performance on specific tasks. In addition, successive convolutions can extract increasingly abstract features from data sequences, aiding in obtaining richer and more complex feature representations.

Research using the same dataset has been carried out in several related works. A comparison of the proposed method with previous research can be seen in Table 4.

Table 4. Comparison of the performance of the proposed method with previous studies

Author (Year)	Dataset	Algorithm	Accuracy (%)
Kim & Cho (2019) [17]	Lending Club	CNN	77.78%
Kim & Cho (2022) [18]	Lending Club	CNN +DeseNet + ResNet + GoogleNet	82.8%
Zhang (2020) [19]	Lending Club	OICSM	74.67%
Kim & Cho (2019b) [12]	Lending Club	DenseNet-BC	79.6%
Proposed Method	**Lending Club**	**CNN**	**85.43%**

Based on Table 4, this study has an accuracy performance advantage by applying the CNN algorithm to the Lending Club dataset. Compared to previous studies, this study also applied datasets that had previously been used. However, our proposed method could improve performance values on loan default predictions using the Lending Club dataset. Successive convolution layers can recognize increasingly abstract and high-level features. In addition, the use of MaxPooling layers can perform spatial down sampling to reduce data dimensions, eliminate redundancy, increase computational efficiency, and enable more effective data processing.

This study was successful in using CNN to improve prediction accuracy on p2p lending platforms. However, in the future, the accuracy of this research can still be improved by handling pre-processed data with more effective methods or handling unbalanced data. In addition, it can be tried with a modified CNN architectural approach or the CNN method combined with certain methods to obtain even higher accuracy results

4. CONCLUSION

In this study, it can be concluded that, CNN obtains high accuracy due to its computationally optimized design and using parallel data processing technology. In addition, CNN also uses the convolution layer to extract essential information from data sequences more efficiently and obtain rich and abstract feature representations. By implementing the Lending Club Dataset from the P2P lending platform in America which contains 9,578 data with 14 features, the CNN Architecture is designed using 2 1D convolutional layers, 2 Max pooling layers, dropout layers and SGD optimizer. Conducted 10 epoch training with the highest accuracy of 85.43%. This research can still be developed by adding feature selection methods, feature extraction and data imbalance handling.

REFERENCES

[1] F. I. Lessambo, "Banking Regulation and Fintech Challenges," in *Fintech Regulation and Supervision Challenges within the Banking Industry: A Comparative Study within the G-20*, pp. 1-26, 2023, https://doi.org/10.1007/978-3-031-25428-4_1.

[2] T. Renduchintala, H. Alfauri, Z. Yang, R. Di Pietro, and R. Jain, "A Survey of Blockchain Applications in the FinTech Sector," *J. Open Innov. Technol. Mark. Complex.*, vol. 8, no. 4, p. 185, 2022, https://doi.org/10.3390/joitmc8040185.

[3] M. A. Muslim *et al.*, "New model combination meta-learner to improve accuracy prediction P2P lending with stacking ensemble learning," *Intell. Syst. with Appl.*, vol. 18, p. 200204, 2023, https://doi.org/10.1016/j.iswa.2023.200204.

[4] Y. Dasril, M. A. Muslim, M. F. Al Hakim, J. Jumanto, and B. Prasetiyo, "Credit Risk Assessment in P2P Lending Using LightGBM and Particle Swarm Optimization," *Regist. J. Ilm. Teknol. Sist. Inf.*, vol. 9, no. 1, pp. 18–28, 2023, https://doi.org/10.26594/register.v9i1.3060.

[5] X. Lu, T. Lu, C. Wang, and R. Wu, "Can social notifications help to mitigate payment delinquency in online peer-to-peer lending?," *Prod. Oper. Manag.*, vol. 30, no. 8, pp. 2564–2585, 2021, https://doi.org/10.2139/ssrn.3758144.

[6] R. Ge, J. Feng, B. Gu, and P. Zhang, "Predicting and Deterring Default with Social Media Information in Peer-to-Peer Lending," *J. Manag. Inf. Syst.*, vol. 34, no. 2, pp. 401–424, 2017, https://doi.org/10.1080/07421222.2017.1334472.

[7] Z. U. Rehman, N. Muhammad, B. Sarwar, and M. A. Raz, "Impact of Risk Management Strategies on the Credit Risk Faced by Commercial Banks of Balochistan," *Financ. Innov.*, vol. 5, no. 1, pp. 1–13, 2019, https://doi.org/10.1186/s40854-019-0159-8.

[8] O. Raiter, "Macro-Economic and Bank-Specific Determinants of Credit Risk in Commercial Banks," *Empir. Quests Manag. Essences*, vol. 1, no. 1, pp. 36–50, 2021, https://researchberg.com/index.php/eqme/article/view/28.

[9] M. A. Muslim *et al.*, "An Ensemble Stacking Algorithm to Improve Model Accuracy in Bankruptcy Prediction," *J. Data Sci. Intell. Syst.*, 2023, https://doi.org/10.47852/bonviewJDSIS3202655.

[10] S. Aisyah and Sulastri, "Analisa Data Mining Default Kartu Kredit Menggunakan Algoritma Neural Network," *Din. Inform.,* vol. 10, no. 1, pp. 47–52, 2018, https://doi.org/10.35315/informatika.v10i1.8132.

[11] R. Wijayanti and Sulastri, "Analisa Klasifikasi Kartu Kredit Menggunakan Algoritma Naive Bayes," *Din. Inform.*, vol. 10, no. 2, pp. 80–86, 2018, https://doi.org/10.35315/informatika.v10i2.8137.

[12] J. Y. Kim and S. B. Cho, "Predicting repayment of borrows in peer-to-peer social lending with deep dense convolutional network," *Expert Syst.*, vol. 36, no. 4, 2019, https://doi.org/10.1111/exsy.12403.

[13] J. Duan, "Financial system modeling using deep neural networks (DNNs) for effective risk assessment and prediction," *J. Franklin Inst.*, vol. 356, no. 8, pp. 4716–4731, 2019, https://doi.org/10.1016/j.jfranklin.2019.01.046.

[14] Z. Ma, W. Hou, and D. Zhang, "A credit risk assessment model of borrowers in P2P lending based on BP neural network," *PLoS One*, vol. 16, no. 8, p. e0255216, 2021, https://doi.org/10.1371/journal.pone.0255216.

[15] C. Wang, D. Han, Q. Liu, and S. Luo, "A deep learning approach for credit scoring of peer-to-peer lending using attention mechanism LSTM," *IEEE Access*, vol. 7, pp. 2161–2168, 2018, https://doi.org/10.1109/ACCESS.2018.2887138.

[16] K. Bastani, E. Asgari, and H. Namavari, "Wide and deep learning for peer-to-peer lending," *Expert Syst. Appl.*, vol. 134, pp. 209–224, 2019, https://doi.org/10.1016/j.eswa.2019.05.042.

[17] Kim and Cho, "Towards Repayment Prediction in Peer-to-Peer Social Lending Using Deep Learning," *Mathematics*, vol. 7, no. 11, p. 1041, 2019, https://doi.org/10.3390/math7111041.

[18] J.-Y. Kim and S.-B. Cho, "Ensemble of diverse deep neural networks with pseudo-labels for repayment prediction in social lending," *Sci. Prog.*, vol. 105, no. 3, p. 00368504221124004, 2022, https://doi.org/10.1177/00368504221124004.

[19] Z. Zhang, K. Niu, and Y. Liu, "A Deep Learning Based Online Credit Scoring Model for P2P Lending," *IEEE Access*, vol. 8, pp. 177307–177317, 2020, https://doi.org/10.1109/access.2020.3027337.

[20] D. Jin, Y. Lu, J. Qin, Z. Cheng, and Z. Mao, "SwiftIDS: Real-time intrusion detection system based on LightGBM and parallel intrusion detection mechanism," *Comput. \& Secur.*, vol. 97, p. 101984, 2020, https://doi.org/10.1016/j.cose.2020.101984.

[21] B. Prasetiyo, Alamsyah, M. F. Al Hakim, J. Jumanto, and M. H. Adi, "Differential augmentation data for vehicle classification using convolutional neural network," in *AIP Conference Proceedings*, vol. 2614, no. 1, 2023, https://doi.org/10.1063/5.0126720.

[22] X. Zhang, L. Wang, Y. Zhao, and Y. Su, "Graph-based place recognition in image sequences with CNN features," *J. Intell. \& Robot. Syst.*, vol. 95, pp. 389–403, 2019, https://doi.org/10.1007/s10846-018-0917-2.

[23] A. Mohammed and R. Kora, "A comprehensive review on ensemble deep learning: Opportunities and challenges," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 2, pp. 757–774, 2023, https://doi.org/10.1016/j.jksuci.2023.01.014.

[24] J. Weng *et al.*, "Deep learning for classification of thyroid nodules on ultrasound: validation on an independent dataset," *Clin. Imaging*, vol. 99, pp. 60–66, 2023, https://doi.org/10.1016/j.clinimag.2023.04.010.

[25] A. Parashar, A. Parashar, W. Ding, M. Shabaz, and I. Rida, "Data Preprocessing and Feature Selection Techniques in Gait Recognition: A Comparative Study of Machine Learning and Deep Learning Approaches," *Pattern Recognit. Lett.*, 2023, https://doi.org/10.1016/j.patrec.2023.05.021.

[26] M. Landauer, S. Onder, F. Skopik, and M. Wurzenberger, "Deep learning for anomaly detection in log data: A survey," *Mach. Learn. with Appl.*, vol. 12, p. 100470, 2023, https://doi.org/10.1016/j.mlwa.2023.100470.

[27] Y. Sun, J. Li, Y. Xu, T. Zhang, and X. Wang, "Deep learning versus conventional methods for missing data imputation: A review and comparative study," *Expert Syst. Appl.*, vol. 227, p. 120201, 2023, https://doi.org/10.1016/j.eswa.2023.120201.

[28] S. Coulibaly, B. Kamsu-Foguem, D. Kamissoko, and D. Traore, "Deep learning for precision agriculture: A bibliometric analysis," *Intell. Syst. with Appl.*, vol. 16, p. 200102, 2022, https://doi.org/10.1016/j.iswa.2022.200102.

[29] M. P. A. van den Ende and J.-P. Ampuero, "Automated seismic source characterization using deep graph neural networks," *Geophys. Res. Lett.*, vol. 47, no. 17, p. e2020GL088690, 2020, https://doi.org/10.1029/2020GL088690.

[30] V. Chandrasekhar, J. Lin, O. Morère, H. Goh, and A. Veillard, "A practical guide to CNNs and Fisher Vectors for image instance retrieval," *Signal Processing*, vol. 128, pp. 426–439, 2016, https://doi.org/10.1016/j.sigpro.2016.05.021.

[31] Y. Zhang *et al.*, "A rotation meanout network with invariance for dermoscopy image classification and retrieval," *Comput. Biol. Med.*, vol. 151, p. 106272, 2022, https://doi.org/10.1016/j.compbiomed.2022.106272.

[32] A. Jamali, S. K. Roy, and P. Ghamisi, "WetMapFormer: A unified deep CNN and vision transformer for complex wetland mapping," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 120, p. 103333, 2023, https://doi.org/10.1016/j.jag.2023.103333.

[33] L. Abdelrahman, M. Al Ghamdi, F. Collado-Mesa, and M. Abdel-Mottaleb, "Convolutional neural networks for breast cancer detection in mammography: A survey," *Comput. Biol. Med.*, vol. 131, p. 104248, 2021, https://doi.org/10.1016/j.compbiomed.2021.104248.

[34] A. Anton, N. F. Nissa, A. Janiati, N. Cahya, and P. Astuti, "Application of Deep Learning Using Convolutional Neural Network (CNN) Method For Women's Skin Classification," *Sci. J. Informatics*, vol. 8, no. 1, pp. 144–153, 2021, https://doi.org/10.15294/sji.v8i1.26888.

[35] W. F. Abror, A. Alamsyah, and M. Aziz, "Bankruptcy Prediction Using Genetic Algorithm-Support Vector Machine (GA-SVM) Feature Selection and Stacking," *J. Inf. Syst. Explor. Res.*, vol. 1, no. 2, pp. 103–108, 2023, https://doi.org/10.52465/joiser.v1i2.180.

[36] V. A. Sindagi and V. M. Patel, "A survey of recent advances in CNN-based single image crowd counting and density estimation," *Pattern Recognit. Lett.*, vol. 107, pp. 3–16, 2018, https://doi.org/10.1016/j.patrec.2017.07.007.

[37] A. F. Mulyana, W. Puspita, and J. Jumanto, "Increased accuracy in predicting student academic performance using random forest classifier," *J. Student Res. Explor.*, vol. 1, no. 2, pp. 94–103, 2023, https://doi.org/10.52465/josre.v1i2.169.

[38] F. Sultana, A. Sufian, and P. Dutta, "Evolution of Image Segmentation using Deep Convolutional Neural Network: A Survey," *Knowledge-Based Syst.*, vol. 201–202, p. 106062, 2020, https://doi.org/10.1016/j.knosys.2020.106062.

[39] A. D. Lestari, N. A. Syarifudin, and Y. J. Nurriski, "Application of pest detection on vegetable crops using the cnn algorithm as a smart farm innovation to realize food security in the 4.0 era," *J. Soft Comput. Explor.*, vol. 3, no. 2, 2022, https://doi.org/10.52465/joscex.v3i2.72.

[40] K. Hastuti, "Analisis komparasi algoritma klasifikasi data mining untuk prediksi mahasiswa non aktif," *Semantik*, vol. 2, no. 1, 2012, http://publikasi.dinus.ac.id/index.php/semantik/article/view/132.

[41] S. Raj and S. Masood, "Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques," *Procedia Comput. Sci.*, vol. 167, pp. 994–1004, 2020, https://doi.org/10.1016/j.procs.2020.03.399.

[42] Q. Lin, X. Bao, and C. Li, "Deep learning based missing data recovery of non-stationary wind velocity," *J. Wind Eng. Ind. Aerodyn.*, vol. 224, p. 104962, 2022, https://doi.org/10.1016/j.jweia.2022.104962.

[43] R. A. Hughes, J. Heron, J. A. C. Sterne, and K. Tilling, "Accounting for missing data in statistical analyses: multiple imputation is not always the answer," *Int. J. Epidemiol.*, vol. 48, no. 4, pp. 1294–1304, 2019, https://doi.org/10.1093/ije/dyz032.

[44] B. Wu *et al.*, "Shift: A zero flop, zero parameter alternative to spatial convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 9127–9135, 2018, https://arxiv.org/abs/1711.08141.

[45] H. Isyanto, A. S. Arifin, and M. Suryanegara, "Fast and Accurate Voice Biometrics with Deep Learning Algorithm of CNN Depthwise Separable Convolution Model and Fusion of DWT-MFCC Methods," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 8, no. 3, p. 431, 2022, https://doi.org/10.26555/jiteki.v8i3.24515.

[46] M. Qiao, S. Yan, X. Tang, and C. Xu, "Deep convolutional and LSTM recurrent neural networks for rolling bearing fault diagnosis under strong noises and variable loads," *Ieee Access*, vol. 8, pp. 66257–66269, 2020, https://doi.org/10.1109/ACCESS.2020.2985617.

[47] G. Tolias, R. Sicre, and H. Jégou, "Particular object retrieval with integral max-pooling of CNN activations," *arXiv Prepr. arXiv1511.05879*, 2015, https://arxiv.org/abs/1511.05879.

[48] Y. Umuroglu *et al.*, "Finn: A framework for fast, scalable binarized neural network inference," in *Proceedings of the 2017 ACM/SIGDA international symposium on field-programmable gate arrays*, pp. 65–74, 2017, https://arxiv.org/abs/1612.07119.

[49] S. Park and N. Kwak, "Analysis on the dropout effect in convolutional neural networks," in *Computer Vision--ACCV 2016: 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers,*

Part II 13, pp. 189–204, 2017, https://doi.org/10.1007/978-3-319-54184-6_12.
[50] H.-M. Huang *et al.*, "Implementation of dropout neuronal units based on stochastic memristive devices in neural networks with high classification accuracy," *Adv. Sci.*, vol. 7, no. 18, p. 2001842, 2020, https://doi.org/10.1002/advs.202001842.
[51] Y. Wei, Q. Duan, R. Yuan, X. Yan, and Y. Yang, "Dropout neuronal unit with tunable probability based on NbOx stochastic memristor for efficient suppression of overfitting," *Microelectron. Eng.*, vol. 259, p. 111778, 2022, https://doi.org/10.1016/j.mee.2022.111778.
[52] K. Kawaguchi and H. Lu, "Ordered sgd: A new stochastic optimization framework for empirical risk minimization," in *International Conference on Artificial Intelligence and Statistics*, pp. 669–679, 2020, https://lis.csail.mit.edu/pubs/kawaguchi-aistats20a.pdf.
[53] E. M. Dogo, O. J. Afolabi, N. I. Nwulu, B. Twala, and C. O. Aigbavboa, "A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks," in *2018 international conference on computational techniques, electronics and mechanical systems (CTEMS)*, pp. 92-99, 2018, https://doi.org/10.1109/CTEMS.2018.8769211.

BIOGRAPHY OF AUTHOR

Tiara Lailatul Nikmah received her bachelor degree in computer science from Universitas Negeri Semarang, Indonesia. She has a research interest in Data Mining, Machine Learning, and Artificial Intelligence Applications. She can be contacted at the email address: tiaralaila21@students.unnes.ac.id.

Jumanto received his Master Degree in computer science from Universitas Gadjah Mada, Indonesia. Currently, He is a lecturer in the Department of Computer Science Universitas Negeri Semarang, Indonesia. He has a research interest in Pattern Recognition, Image Processing, Machine Learning, and Artificial Intelligence Applications. He can be contacted at the email address: jumanto@mail.unnes.ac.id.

Budi Prasetiyo is currently a lecturer and head of laboratory in Department of Computer Science Universitas Negeri Semarang (UNNES), Indonesia. He received his bachelor's degree in Mathematics from UNNES. Then, He obtained his Master's Degree in Computer Science from Universitas Diponegoro, Indonesia. His research interest is in Data Mining, Computational Mathematics, Machine Learning, and Artificial Intelligence. He can be contacted at email: bprasetiyo@mail.unnes.ac.id.

Nina Fitriani is currently pursuing her Ph.D. in English education University of Manchester, UK. She received her bachelor's degree in English education from Universitas Sumatera Utara (USU). Her research interest is education, language, digital literacy. She can be contacted at email: fitrianinina@gmail.com.

Much Aziz Muslim is currently pursuing his Ph.D. in the faculty of management technology at Universiti Tun Hussein Onn Malaysia (UTHM) and working as a lecturer in the department of computer science Universitas Negeri Semarang (UNNES), Indonesia. He received his master's degree in computer science from Universitas Gadjah Mada, Indonesia. The scope of research he is currently working on is in the fields of Computer Science, Data Mining, and Networking. He can be contacted at email: a212muslim@mail.unnes.ac.id.