

# Application of Deep Learning for Credit Card Approval: A Comparison with Two Machine Learning Techniques

Md. Golam Kibria and Mehmet Sevkli

**Abstract**—The increased credit card defaulters have forced the companies to think carefully before the approval of credit applications. Credit card companies usually use their judgment to determine whether a credit card should be issued to the customer satisfying certain criteria. Some machine learning algorithms have also been used to support the decision. The main objective of this paper is to build a deep learning model based on the UCI (University of California, Irvine) data sets, which can support the credit card approval decision. Secondly, the performance of the built model is compared with the other two traditional machine learning algorithms: logistic regression (LR) and support vector machine (SVM). Our results show that the overall performance of our deep learning model is slightly better than that of the other two models.

**Index Terms**—Artificial intelligence, machine learning, deep learning, credit risk management.

## I. INTRODUCTION

The growth of the internet has led to a significant rise in credit card usage. It is one of the most used payment methods these days. As the world economy increases, credit card fraud also increasing at an alarming rate [1]. It is also evident that credit card defaulters have also increased significantly. Consequently, the credit card issuing institutions are becoming meticulous in approving credit cards to customers. In addition, the downturn of financial institutions in the USA and Europe during the US subprime mortgage and the European sovereign crisis has raised concerns about risk management properly [2]. Hence, these challenges have attracted significant attention from researchers and practitioners. A wide range of statistical and machine learning techniques have been developed to solve credit card related problems (see [1]-[7]). It is found that machine learning techniques are superior to other traditional statistical techniques in dealing with credit scoring [8]-[11]. In particular, deep learning is a most popular and accurate classification technique that outperforms other machine learning models (e.g. logistic regression (LR), linear discriminant analysis (LDA), multiple discriminant analysis (MDA), k-nearest neighbor (k-NN), decision trees, etc.) [12]. Deep learning is also found to be a state-of-art research area

to solve various practical problems including credit card fraud [6]. Some of the problems for which deep learning technique is found to be the best method to solve are illustrated in Table I.

TABLE I: DEEP LEARNING APPLICATIONS

Area	Problem	Paper
Natural Language Processing (NLP)	Sentiment Analysis	Socher, Perelygin [13], Kim [14], Wehrmann, Becker [15]
	Translation	Bahdanau, Cho [16], Cho, Van Merriënboer [17]
	Question & Answer	Feng, Xiang [18], Dong, Wei [19]
Visual Data Process	Image Classification	Krizhevsky, Sutskever [20], LeCun, Bottou [21]
	Object Detection and Semantic Segmentation	Girshick [22], Girshick, Donahue [23]
	Video Processing	Tsagkatakis, Jaber [24], Karpathy, Toderici [25]
	Speech Emotion Recognition (SER)	Neumann and Vu [26], Han, Yu [27]
Speech and Audio Processing	Speech Enhancement (SE)	Huang, Kim [28], Neumann and Vu [26]
	Social Network Analysis	Zhang, Zhao [29], Huang, Kim [28]
Other Problems	Information Retrieval	Deng, He [30], Shen, He [31]
	Transportation Prediction	Nie, Jiang [32], Ma, Yu [33]
	Autonomous Driving	Geiger, Lenz [34], Hadsell, Erkan [35]
	Biomedicine	Litjens, Sánchez [36], Cireşan, Giusti [37]
	Disaster Management Systems	Tian and Chen [38], Tian and Chen [39]
	Credit card frauds	Niimi [6], Zhang, Han [40], Chaudhary, Yadav [1], Saberi, Mirtalaie [12], Dighe, Patil [41], Pumsirirat and Yan [3], Ong, Huang [10] etc.

In credit card context, most studies used traditional statistical, machine learning, and deep learning techniques to detect credit card fraud and compared the results [1]-[3], [5]-[7], [12], [40], [41]. However, the literature review explores that there is a very little research done to decide whether a customer is to be issued a credit card or not based on their information. Therefore, this study aims to support the decision-makers of whether a customer is to be issued a credit card or not. This study has two objectives. First, it will build a deep learning model based on the best parameters for the credit card dataset. Second, a comparative study between deep learning and traditional machine learning algorithms (Logistic Regression and SVM) will also be conducted.

Manuscript received January 30, 2020; revised November 10, 2020.

Md. Golam Kibria is with the College of Business and Innovation, The University of Toledo, Toledo, OH 43606-3390 USA and on leave from Independent University, Bangladesh (IUB), Bangladesh (e-mail: mkibria@rockets.utoledo.edu).

Mehmet Sevkli is with the College of Business and Innovation, The University of Toledo, Toledo, OH 43606-3390 USA (e-mail: mehmet.sevkli@utoledo.edu).

## II. MODELS

### A. Logistic Regression Model

Logistic Regression (LR) is one of the most commonly applied statistical techniques for credit card analysis [5], [30], [31]. It predicts the likelihood of a result that can just have two states (i.e. a dichotomy). The prediction depends on the use of one or several indicators (numerical and categorical). According to [7], it seeks the best fit parameter to determine the probability of the binary response based on one or more features. Based on independent variables for each credit card application, it provides a probability that is used to classify the application as accepted or rejected [5]. If the probability is larger than the threshold value, it is accepted. Otherwise, it is rejected. LR function takes as input the client characteristics and outputs the probability of default.

$$p = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)} \quad (1)$$

where in the above

- $p$  is the probability of default
- $x_i$  is the explanatory factor  $i$
- $\beta_i$  is the regression coefficient of the explanatory factor  $i$
- $n$  is the number of explanatory variables

For each of the existing data points, it is known whether the client has gone into acceptance or not (i.e.  $p=1$  or  $p=0$ ). The aim in the here is to find the coefficients  $\beta_0, \beta_1, \beta_2, \dots, \beta_n$  such that the model's probability of default equals to the observed probability of default.

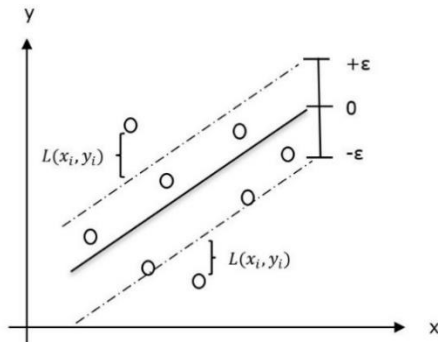


Fig. 1. The graph of support vector regression.

### B. Support Vector Machine (SVM) Model

Support vector machine (SVM) is an algorithm that learns based on instances given and predicts [42]. For instance, an SVM can learn to recognize fraudulent credit card activity by examining hundreds or thousands of fraudulent and non-fraudulent credit card activity reports. SVM was firstly introduced by [43]. It is used as a classification and regression tool to maximize predictive accuracy [2]. SVM is the best fit for supervised learning where data are linearly categorized and examined [7]. Support Vector Regression (SVR) methods aim to approximate the following function

$$f(x) = wx + b \quad (2)$$

by minimizing the following objective function

$$\frac{1}{2} \|w\|^2 + C \frac{1}{n} \sum_{i=1}^n L(x_i, y_i) \quad (3)$$

where  $\|w\|$  is the regularization term,  $L(x_i, y_i)$  is the loss function and  $C$  is the trade-off between model complexity and

error on training dataset. The graphical representation of SVR can be seen in Fig. 1. The advantage of SVR is to present convex solution space resulting in a unique solution.

The data points are not always in a linear classification; the kernel functions enable us to transform the nonlinear dataset into a linear separation format. Fig. 2 shows the transformation of a nonlinear dataset to a linear dataset by using kernel functions.

$$y = \sum_{i=1}^N (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (4)$$

where  $y$  is output,  $\alpha_i$  and  $\alpha_i^*$  are lagrange multipliers,  $x_i$  is input vector,  $K(x_i, x)$  is kernel function, and  $b$  is bias.

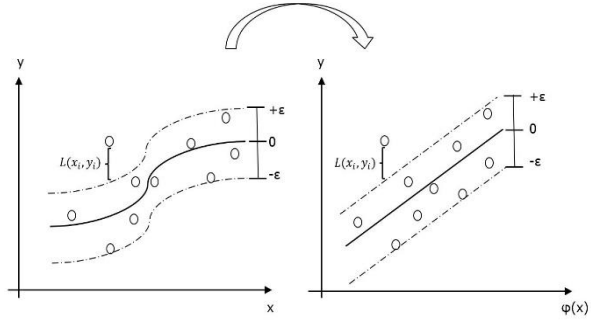


Fig. 2. The transformation of nonlinear dataset to linear dataset by using Kernel functions.

### C. Deep Learning

Deep learning (DL) is a subset of machine learning methods based on artificial neural networks. The core concept of deep learning is automating the extraction of features from the data [43]. According to [44], “deep learning is a class of machine learning algorithms that: (1) use a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. Each successive layer uses the output from the previous layer as input, (2) learn multiple levels of representations that correspond to different levels of abstraction; the levels form a hierarchy of concepts.” Deep learning has recently drawn much attention from researchers in the field of machine learning [6]. It is considered as a robust algorithm for image identification and credit fraud detection [5]. DL is a multi-layer perceptron network that uses a stochastic gradient descent for training [7]. The deep learning principle is similar to an ANN that has many hidden layers. Conversely, non-deep learning feed forward neural networks have only a single hidden layer. The given picture shows the comparison between non-deep learning as in Fig. 3 and deep learning with hidden layers as in Fig. 4.

A sigmoid or a tahn function is applied as an activation function in the deep learning algorithm (see 5, 6).

$$hi(x) = \text{sigmoid}(x^T W \dots i + bi) \quad (5)$$

$$hi(x) = \text{tanh}(x^T W \dots i + bi) \quad (6)$$

## III. DATA

This study used the credit card approval dataset by UCI Machine repository to evaluate the experimental results (see [45]). The UCI Machine Learning Repository is considered to be a good source of data for conducting empirical and methodological research in deep learning. In the dataset, arbitrary names and values were given to the attributes to

maintain the confidentiality of the data. Table II illustrates the details of the dataset.

TABLE II: ATTRIBUTES INFORMATION IN DATASET

Attribute	Type
A1	Nominal
A2	Continuous
A3	Continuous
A4	Nominal
A5	Nominal
A6	Nominal
A7	Nominal
A8	Continuous
A9	Nominal
A10	Nominal
A11	Continuous
A12	Nominal
A13	Nominal
A14	Continuous
A15	Continuous
A16	Dichotomous (Class Attribute)

### A. Data Pre-processing

Some missing values were found in the dataset and taken care of following the appropriate machine learning approach to replace the missing data. All categorical attributes were converted to binary numerical attributes. Then, all data were normalized.

### B. Data Analyzing Platform

Data were analyzed using respective machine learning algorithms (LR, SVM, and DL) with different parameters. The WEKA tool was used for SVM and LR while Python programming language was developed for DL.

## IV. EXPERIMENTAL DESIGN

The main purpose of this study is to build a deep neural network based on parameters that provide the best performance. Different configurations of DL architectures are examined in this study by varying the number of layers and the number of neurons in each layer to see which configuration gives best performance on the data set. A total of 24 different combinations are evaluated for DL in which

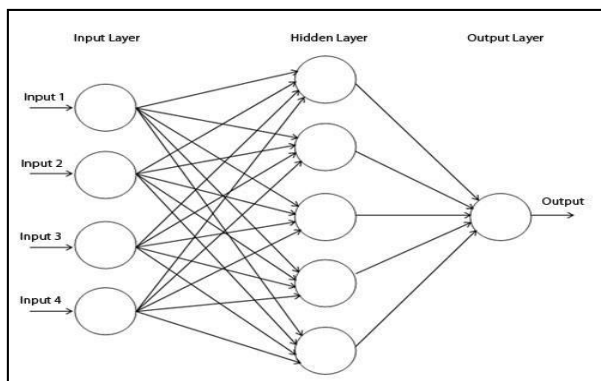


Fig. 3. Single layer hidden neural network.

Based on the confusion matrix, the following classification performance measures are used to evaluate the model performance:

$$\text{Accuracy: } (TP + TN) / (TP + FP + TN + FN) \quad (7)$$

2-, 3-, 5-, and 7-hidden layer networks with 3, 5, 7, 16, 32 and 64 neurons are experimented with. The number of neurons is kept the same in each layer for a single network configuration. For instance, if it is a 5-hidden layer network with 16 neurons, then each of the 5 hidden layers will have 16 neurons. In the first experimentation, the following parameters of the DL were a used-loss function: binary cross-entropy, optimizer: adam, activation function: rectified linear units (ReLU), the batch size for training and prediction: 15 and epochs: 50. A sigmoid function was used in the output layer. The popular 10-fold cross-validation approach is used for model evaluation and model selection to avoid overfitting classifiers [46]. Tuning with a grid search in parameter space is employed for fine-tuning the important parameters to find out the best parameters. After several experiments, Table III shows the best parameters used in the deep learning model:

TABLE III: PARAMETER TUNING

Parameters	Possible values	Best parameter s
batch_size	5, 10, 15, 20, 25, 30, 50	15
epochs	10, 20, 30, 50, 75, 100, 200	100
Optimization	'SGD', 'RMSprop', and 'Adam'	'RMSprop'
Network Weight Initialization	'uniform', 'lecun_uniform', 'normal', 'zero', 'glorot_normal', 'glorot_uniform', 'he_normal', and 'he_uniform'	'uniform'
Activation Function	'softmax', 'relu', 'tanh', 'sigmoid'	'relu'
neurons	1, 3, 5, 7, 10, 15, 20	5

## V. PERFORMANCE EVALUATION

### A. Metrics

The chosen algorithms assume the underlying fraud detection issue as a classification problem. We have considered the confusion matrix given in Table IV for evaluating metrics. However, classical metrics of accuracy and confusion matrix will not be able to capture the actual fraud identification rate due to skewness in instances of each class. Thus, metrics that balance the detection of both classes have been considered.

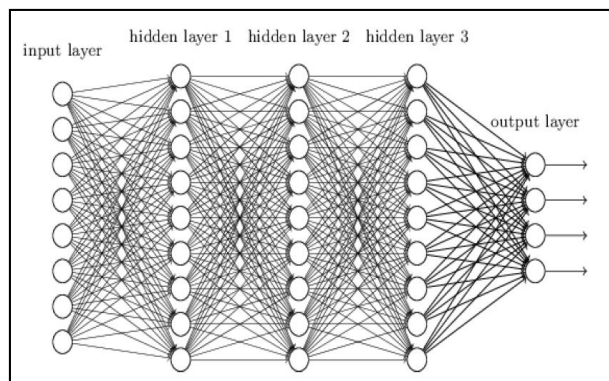


Fig. 4. Deep neural network.

$$\text{Recall (or Sensitive / True positive rate): } TP / (TP + FN) \quad (8)$$

$$\text{Precision: } TP / (TP + FP) \quad (9)$$

$$\text{F1-measure: } 2 \times ((\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})) \quad (10)$$

$$\text{False positive rate: FP} / (\text{FP} + \text{TN}) \quad (11)$$

TABLE IV: CONFUSION MATRIX FOR EVALUATING CLASSIFICATION

		Predicted Class	
		Positive	Negative
Actual Class	Positive	TP	FN
	Negative	FP	TN

### B. Experimental Results

In this paper, three algorithms namely SVM, LR, and DL are compared with each other. The WEKA (Waikato environment for knowledge analysis) tool is used for Support Vector Machine (SVM) and Logistic Regression to calculate the efficiency based on accuracy garnered from the confusion matrix and Python programming language is developed for Deep Learning (DL).

TABLE V: RESULTS

Classifier	F1-Measure	Precision	Recall	False Positive (FP)	Accuracy
<b>SVM</b>	.863	86.80%	86.20%	12.80%	86.23%
<b>LR</b>	.861	86.40%	86.20%	16.10%	86.23%
<b>DL</b>	.886	87.91%	89.26%	16.00%	87.10%

Table V illustrates the experimental results. For each classifier, F1-Measure, Precision, Recall, FP and the accuracy are displayed. As the deep learning results depend on the initial parameters, the algorithm was run for 5 times and the results reported in Table V are the average results of the five experimentations. Accuracy is the percentage of correctly classified instances and provides a measure for the ability to make accurate predictions on previously unseen cases. The F1-measure can reflect the overall performance of the model. The recall metric represents the proportion of the actual rejected applications that have been correctly predicted, while the precision metric denotes the proportion of the correctly predicted rejected applications to the predicted rejected applications. Both recall and precision are important evaluation metrics. In addition, the false positive rate is defined as the proportion of the applications that have been wrongly categorized as positive (false positives).

As shown in Table V, the accuracy rate of DL is the highest at 87.10%. However, the accuracy rates of the other two classifiers are same at 86.23%. Moreover, the precision and recall of DL is higher than that of SVM and LR. Recall value is the same for both SVM and LR while precision slightly differs from each other. The comparative results indicate that deep learning performs better for the credit card dataset.

Specifically, based on the F1-measure, the DL achieves the highest F1-measure score of .886, which indicates the overall performance of the model. F1-measure value for SVM is .863 while .861 for LR. These two algorithms produced almost the same F1-measure scores. In respect to the false positive rate, the SVM outperformed the other two algorithms, 12.80% for SVM, 16.10% for LR and 16% for DL. Based on the all accuracy measures except FP in Table V, we can conclude that the deep learning model performs slightly better than the other two models.

## VI. CONCLUSION

In this paper, we have built a deep learning model based on the best parameters found by the grid search technique. The

built model is then applied for the credit card data set and compared the results with logistic regression and support vector machine models. It is concluded that the deep learning model performed slightly better than the other two models. LR and SVM produced almost the same results. In the future, another experiment can be evaluated using the large dataset to see the comparative accuracy and applicability of these methods.

### CONFLICT OF INTEREST

The authors declare no conflict of interest.

### AUTHOR CONTRIBUTIONS

The first author conducted the research and analyzed the data. The second author collected data and guided the first author throughout the research and data analysis process. Both authors wrote the paper and approved the final version.

### REFERENCES

- [1] K. Chaudhary, J. Yadav, and B. Mallick, "A review of fraud detection techniques: Credit card," *International Journal of Computer Applications*, vol. 45, no. 1, pp. 39-44, 2012.
- [2] C. Luo and D. Wu, "A deep learning approach for credit scoring using credit default swaps," *Engineering Applications of Artificial Intelligence*, vol. 65: p. 465-470, 2017.
- [3] A. Pumsirirat and L. Yan, "Credit card fraud detection using deep learning based on auto-encoder and restricted boltzmann machine," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 1, pp. 18-25, 2018.
- [4] T. B. Trafalis and H. Ince, "Support vector machine for regression and applications to financial forecasting," in *Proc. the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, 2000.
- [5] G. Rushin *et al.*, "Horse race analysis in credit card fraud—deep learning, logistic regression, and Gradient Boosted Tree," in *Proc. 2017 Systems and Information Engineering Design Symposium (SIEDS)*, 2017.
- [6] A. Niimi, "Deep learning for credit card data analysis," in *Proc. 2015 World Congress on Internet Security (WorldCIS)*, 2015.
- [7] S. Mittal and S. Tyagi, "Performance evaluation of machine learning algorithms for credit card fraud detection," in *Proc. 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 2019.
- [8] B. K. Wong and Y. Selvi, "Neural network applications in finance: A review and analysis of literature (1990–1996)," *Information & Management*, vol. 34, no. 3, pp. 129-139, 1998.
- [9] A. Vellido, P. J. Lisboa, and J. Vaughan, "Neural networks in business: A survey of applications (1992–1998)," *Expert Systems with Applications*, 1999, vol. 17, no. 1, pp. 51-70.
- [10] C.-S. Ong, J.-J. Huang, and G.-H. Tzeng, "Building credit scoring models using genetic programming," *Expert Systems with Applications*, vol. 29, no. 1, pp. 41-47, 2005.
- [11] Z. Huang *et al.*, "Credit rating analysis with support vector machines and neural networks: A market comparative study," *Decision Support Systems*, vol. 37, no. 4, pp. 543-558, 2004.
- [12] M. Saberi *et al.*, "A granular computing-based approach to credit scoring modeling," *Neurocomputing*, 2013, vol. 122, pp. 100-115.
- [13] R. Socher *et al.*, "Recursive deep models for semantic compositionality over a sentiment treebank," in *Proc. the 2013 Conference on Empirical Methods in Natural Language Processing*, 2013.
- [14] Y. Kim, "Convolutional neural networks for sentence classification," arXiv preprint arXiv: 1408.5882, 2014.
- [15] J. Wehrmann *et al.*, "A character-based convolutional neural network for language-agnostic Twitter sentiment analysis," in *Proc. 2017 International Joint Conference on Neural Networks (IJCNN)*, 2017.
- [16] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," arXiv preprint arXiv: 1409.0473, 2014.
- [17] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," arXiv preprint arXiv: 1406.1078, 2014.
- [18] M. Feng *et al.*, "Applying deep learning to answer selection: A study and an open task," in *Proc. 2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, 2015.

- [19] L. Dong *et al.*, “Question answering over freebase with multi-column convolutional neural networks,” in *Proc. the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, 2015, vol. 1.
- [20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Advances in Neural Information Processing Systems*, 2012.
- [21] Y. LeCun *et al.*, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, 1998, vol. 86, no. 11, pp. 2278-2324.
- [22] R. Girshick, “Fast r-cnn,” in *Proc. the IEEE International Conference on Computer Vision*, 2015.
- [23] R. Girshick *et al.*, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *Proc. the IEEE Conference on Computer Vision and Pattern Recognition*, 2014.
- [24] G. Tsagkatakis, M. Jaber, and P. Tsakalides, “Goal!! Event detection in sports video,” *Electronic Imaging*, vol. 16, pp. 15-20, 2017.
- [25] A. Karpathy *et al.*, “Large-scale video classification with convolutional neural networks,” in *Proc. the IEEE conference on Computer Vision and Pattern Recognition*, 2014.
- [26] M. Neumann and N. T. Vu, “Attentive convolutional neural network based speech emotion recognition: A study on the impact of input features, signal length, and acted speech,” arXiv preprint arXiv: 1706.00612, 2017.
- [27] K. Han, D. Yu, and I. Tashev, “Speech emotion recognition using deep neural network and extreme learning machine,” in *Proc. Fifteenth Annual Conference of the International Speech Communication Association*, 2014.
- [28] P.-S. Huang *et al.*, “Deep learning for monaural speech separation,” in *Proc. 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014.
- [29] X. Zhang, J. Zhao, and Y. LeCun, “Character-level convolutional networks for text classification,” *Advances in Neural Information Processing Systems*, 2015.
- [30] L. Deng, X. He, and J. Gao, “Deep stacking networks for information retrieval,” in *Proc. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013.
- [31] Y. Shen *et al.*, “Learning semantic representations using convolutional neural networks for web search,” in *Proc. the 23rd International Conference on World Wide Web*, 2014.
- [32] L. Nie *et al.*, “Traffic matrix prediction and estimation based on deep learning for data center networks,” in *Proc. 2016 IEEE Globecom Workshops (GC Wkshps)*, 2016.
- [33] X. Ma *et al.*, “Large-scale transportation network congestion evolution prediction using deep learning theory,” *PloS One*, vol. 10, no. 3, 2015.
- [34] A. Geiger *et al.*, “Vision meets robotics: The kitti dataset,” *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231-1237, 2013.
- [35] R. Hadsell *et al.*, “Deep belief net learning in a long-range vision system for autonomous off-road driving,” in *Proc. 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008.
- [36] G. Litjens *et al.*, “Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis,” *Scientific Reports*, vol. 6, p. 26286, 2016.
- [37] D. C. Cireşan *et al.*, “Mitosis detection in breast cancer histology images with deep neural networks,” in *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2013.
- [38] H. Tian and S.-C. Chen, “A video-aided semantic analytics system for disaster information integration,” in *Proc. 2017 IEEE Third International Conference on Multimedia Big Data (BigMM)*, 2017.
- [39] H. Tian and S.-C. Chen, “MCA-NN: Multiple correspondence analysis based neural network for disaster information detection,” in *Proc. 2017 IEEE Third International Conference on Multimedia Big Data (BigMM)*, 2017.
- [40] X. Zhang *et al.*, “HOPA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture,” *Information Sciences*, 2019.
- [41] D. Dighe, S. Patil, and S. Kokate, “Detection of credit card fraud transactions using machine learning algorithms and neural networks: A comparative study,” in *Proc. 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, 2018.
- [42] W. S. Noble, “What is a support vector machine?” *Nature Biotechnology*, vol. 24, no. 12, pp. 1565-1567, 2006.
- [43] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “A training algorithm for optimal margin classifiers,” in *Proc. the Fifth Annual Workshop on Computational Learning Theory*, 1992.
- [44] L. Deng and D. Yu, “Deep learning: Methods and applications,” *Foundations and Trends R in Signal Processing*, 2014, Now Publishers Inc.: Hanover, MA, USA.
- [45] D. Dua and C. Graff, “UCI machine learning repository,” School of Information and Computer Science, University of California, Irvine, CA, 2019.
- [46] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Information Processing & Management*, vol. 45, no. 4, pp. 427-437, 2009.

Copyright © 2021 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



**Md. Golam Kibria** is a Ph.D. student in the College of Business and Innovation at The University of Toledo, USA. He is also a lecturer (on study leave) in the School of Business at Independent University, Bangladesh (IUB), Bangladesh. He graduated with BBA and MBA in management information systems at the University of Dhaka, Bangladesh in 2011 and 2012, respectively. His current research interest includes machine learning, deep learning, technology adoption, e-government and sustainable development.



**Mehmet Sevkli** got his Ph.D. degree in the field of industrial engineering from Istanbul Technical University in Istanbul, Turkey. His primary research interests include operations research, data analysis/business analytics, predictive and prescriptive analytics, statistics, quantitative management, operations management, and optimization. In particular, he has focused on developing a new metaheuristics approach to combinatorial optimization problems and a fuzzy approach to multi-criteria decision-making problems. He has several papers published in prestigious academic journals and numerous national and international conference papers. He has supervised one PhD dissertation and eight master theses; took part at various academic projects as project coordinator, or project team member. He has teaching experience in diverse fields within the industrial engineering major, namely operations research, operations and supply chain management, facility planning, statistics and management information systems (MIS). Dr. Mehmet Sevkli is currently a visiting professor of Information, Operations and Technology Management, University of Toledo.