Received: 10 September 2018 | Revised: 31 January 2019 | Accepted: 1 March 2019

DOI: 10.1111/exsy.12403

SPECIAL ISSUE PAPER

WILEY Expert Systems

Predicting repayment of borrows in peer-to-peer social lending with deep dense convolutional network

Ji-Yoon Kim | Sung-Bae Cho

Department of Computer Science, Yonsei University, Seoul, Korea

Correspondence
Sung-Bae Cho, Department of Computer Science, Yonsei University, 50 Yonsei-ro, Sinchon-dong, Seodaemun-gu, Seoul, Korea.
Email: sbcho@yonsei.ac.kr

Funding information
Korea Electric Power Corporation, Grant/Award Number: R18XA05

Abstract

In peer-to-peer lending, it is important to predict the repayment of the borrower to reduce the lender's financial loss. However, it is difficult to design a powerful feature extractor for predicting the repayment as user and transaction data continue to increase. Convolutional neural networks automatically extract useful features from big data, but they use only high-level features; hence, it is difficult to capture a variety of representations. In this study, we propose a deep dense convolutional network for repayment prediction in social lending, which maintains the borrower's semantic information and obtains a good representation by automatically extracting important low- and high-level features simultaneously. We predict the repayment of the borrower by learning discriminative features depending on the loan status. Experimental results on the Lending Club dataset show that our model is more effective than other methods. A fivefold cross-validation is performed to run the experiments.

KEYWORDS

deep learning, dense convolutional networks, peer-to-peer lending, repayment prediction

1 | INTRODUCTION

As online services continue to grow rapidly, a large amount of data is being generated (Chen, Mao, & Liu, 2014). Big data have been used to address critical issues that occur in the social, economic, and technical spheres (Kim & Cho, 2018; Ronao & Cho, 2016a). Peer-to-peer (P2P) lending is one of the FinTech services that directly match lenders with borrowers through online platforms. It generates huge transactional data and attracts many users (Zhao et al., 2017). Transactional data on P2P lending is used to address issues such as credit scorings and default prediction.

The lenders suffer from financial losses when their borrowers do not pay or partially pay them in the repayment period. The lenders may suffer due to the default of the borrowers (Xu, Chen, & Chau, 2016). To reduce the financial risk of the lenders, it is important to predict defaults and assess the creditworthiness of the borrowers (Serrano-Cinca, Gutiérrez-Nieto, & López-Palacios, 2015). If we effectively leverage a large amount of transactional data processed online, we can support the prediction of borrowers' repayment (Yan, Yu, & Zhao, 2015). Related studies generally select features such as the borrower characteristics and credit history from the lending data or extract new features for predicting the repayment (Malekipirbazari & Aksakalli, 2015). They attempted to model repayment prediction and credit assessment by employing a machine learning algorithm (Namvar, Siami, Rabhi, & Naderpour, 2018).

It is difficult to define a discriminative feature extractor for repayment prediction. There is a limit to extract good low- and high-level data representation by capturing the relationship between large and complex data using statistical and manual methods (Philip Chen & Zhang, 2014). A poor data representation can decrease the performance even for an advanced and sophisticated machine learner (Najafabadi et al., 2015).

The Lending Club in the United States provides complete loan data containing the current loan status and latest payment information (https://www.lendingclub.com). From the data of 2007 to 2017, it consists of approximately 1.7 million instances with 111 attributes. The structure of the data is huge and complex, and it is difficult to extract useful high-level features for repayment prediction. Table 1 summarizes the data attributes.

TABLE 1 The summary of Lending Club data

Type	Description	No. of variable
Predictor	Current status of the loan (charged off or fully paid)	1
Borrower information	Demographic or personal information of the borrowers (such as ID, grade, employment length, home ownership, annual income, and zip code)	18
Loan characteristics	Information for loan type (such as term, interest rate, instalment, purpose, issue day, total payment, and loan amount)	24
Credit history	Information of the lenders for credit card, check card, and delinquent record	68

Convolutional neural network (CNN), one of the typical deep learning methods, automatically extracts high-level and complex features from big data (Chen & Lin, 2014; LeCun, Bottou, Bengio, & Haffner, 1998). It is possible to capture the statistical pattern of input hierarchically by extracting general features from the lower layers and high-level features from the higher layers with several convolution operations (Ronao & Cho, 2016b). It often outperforms the state-of-the-art performance achieved by handcrafted features (Qiu, Wu, Ding, Xu, & Feng, 2016).

Because the CNN classifier uses only high-level features, it is difficult to capture a good representation from large and complex P2P lending data. As an alternative, advanced CNNs have been proposed to maximize semantic information flow from low to high levels (Huang, Sun, Liu, Sedra, & Weinberger, 2016; Xie, Girshick, Dollar, Tu, & He, 2017; Zagoruyko & Komodakis, 2016). One of them, a dense convolutional network (DenseNet; Huang, Liu, Matten, & Weingerger, 2017), is a feed-forward network in which all layers are directly connected. This achieves a good representation by using general and high-level features.

In this study, we propose a model employing DenseNet (Huang et al., 2017) for directly learning the representation in social lending data to not only improve the classification performance but also obtain a good feature representation for repayment prediction. The proposed network maintains the borrower's semantic information and obtains a powerful representation by learning low- and high-level features simultaneously through dense connectivity. A dense block extracts diversified features of borrowers, and a transition layer merges representative features. Our model classifies the loan status of a borrower by learning the pattern while maintaining and extracting discriminative features according to the loan status and improves the overall performance with a good representation.

The main contributions of this study are as follows: (a) We present a powerful architecture based on a deep dense connected convolutional network, which can obtain diversified features including low- and high-level features to address the repayment prediction of borrowers. (b) Instead of using a traditional feature extraction method, the proposed method automatically extracts a good representation in the feature space from a subset of borrowers. That is, it distinguishes the characteristics of borrowers and obtains an excellent representation without dimension reduction or feature extraction. Our model shows how the systems that do not require handcrafted features outperform other machine learning algorithms. (c) Extensive experiments conducted on the Lending Club dataset demonstrate that the proposed method achieves the highest performance among the conventional machine learning and deep learning methods for predicting the repayment.

This paper is organized as follows. Section 2 discusses the related works on P2P social lending. Section 3 describes the proposed DenseNet architecture for predicting the repayment. In Section 4, we present the experimental results of our proposed architecture compared with other methods and analyse the results. Section 5 concludes this study.

2 | RELATED WORK

As P2P lending transactions increase, it becomes very important to predict the repayment (Milne & Parboteeah, 2016). Many researchers for repayment prediction in P2P social lending employ machine learning methods to improve the classification accuracy for defaults, for example, logistic regression (Guo, Zhou, Luo, Liu, & Xiong, 2016; Lin, Li, & Zheng, 2017; Zhang, Li, Hai, Li, & Li, 2017) and trees such as decision tree (DT) and random forest (RF; Malekipirbazari & Aksakalli, 2015; Serrano-Cinca & Gutiérrez-Nieto, 2016). They also try to extract efficient features to improve the performance of models that predict the repayment in P2P social lending. In this section, a literature review is presented focusing on feature engineering.

Some researchers selected features using statistical methods (Polena & Regner, 2016) or information gains (Chen, 2017) before training machine learning models. The selected features are as follows: loan amount, debt-to-income ratio, home ownership, delinquencies, etc. Other researchers extracted new features based on existing ones. They extracted features related to P2P lending as handcrafted or extracted new features using feature extraction techniques such as latent Dirichlet allocation. In particular, Jiang, Wang, Wang, and Ding (2017) presented a debt default prediction method combined with soft information related to text description. They extracted topic features such as asset, income, work, family, business, and agriculture from texts to provide qualitative information. These studies have drawbacks in that it is difficult to compare their performances and generalize these models because they derive unique features (Ronao & Cho, 2016b).

Studies that use neither feature selection nor feature extraction selected only preprocessed features. They occasionally labelled new classes to utilize a large amount of data. Fu (2017) labelled "good" and "bad" for 1,320 K data to predict the default. Vinod Kumar, Natarajan, Keerthana, Chinmayi, and Lakshmi (2016) similarly labelled "good" and "bad" including instances such as "current" and "late." However, assigning a new class makes it difficult to compare the performance.

As data are large and complex, it is difficult to extract the discriminative features of borrowers by employing existing data mining and machine learning techniques (Sohangir, Wang, Pomeranets, & Khoshgoftaar, 2018). Other methods that can extract meaningful patterns are required in big data. Deep learning provides an opportunity to address the issues for learning the patterns that exist in big data (Najafabadi et al., 2015) and is more effective when learning patterns in complex data. In this study, we show that a variety of features can be automatically extracted from big data for predicting repayment.

3 | THE PROPOSED METHOD

Figure 1 shows the proposed architecture for repayment prediction in social lending using DenseNet. We assume that it would be easier to predict the loan status in the feature space if the network captures the inherent properties of the lending data of a borrower and generalizes them to other borrowers. The feature space is learned to obtain a good representation from the social lending data of many borrowers. The lending data are projected onto the learned feature space using the trained network and are predicted using the softmax classifier.

3.1 | Dense convolutional network

3.1.1 | Dense connectivity

DenseNet has direct connections from every layer to all subsequent layers, known as dense connectivity. The low- and high-level features extracted from the social lending data are reused using dense connectivity. Assume that the output passed through the l^{th} layer from the normalized lending data be x_l. The output of the l^{th} layer is

$$x_l = H_l([x_0, x_1, ..., x_{l-1}]), \tag{1}$$

where $H_l(\cdot)$ is a composite function consisting of a rectified linear unit (Nair & Hinton, 2010) as its activation function and a 3×3 convolution operation, and $[x_0, x_1, \ldots, x_{l-1}]$ represent the concatenation of feature maps generated in convolutional layers 1,2, … ,$l-1$. The concatenated general and high-level features in social lending data are used for training. Rectified linear unit, which outputs the input x when $x \geq 0$ and zero when $x < 0$, is an activation function that can effectively propagate the gradient.

3.1.2 | Dense block

DenseNet consists of several dense blocks and transition layers. A dense block contains a directly connected convolution layer, and the size of the feature maps in each block is maintained. We introduced a 1×1 convolution before each 3×3 convolution to increase the computational efficiency

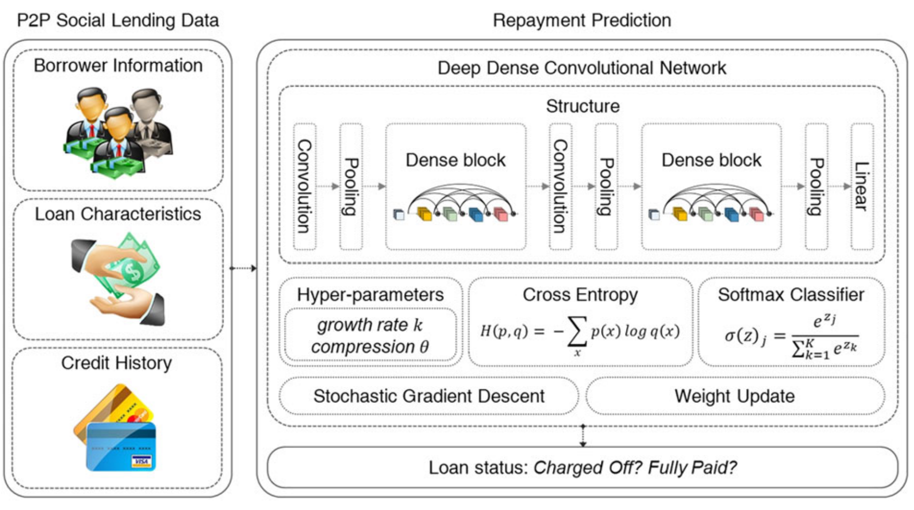

FIGURE 1 The proposed architecture for repayment prediction. P2P: peer-to-peer

and to reduce the number of feature maps in the dense block that acts as a bottleneck layer. Various features are extracted through the dense block from the inherent properties such as the borrower information or loan product characteristics from the Lending Club dataset.

DenseNet uses hyperparameter k as the growth rate to control the excessive increase in the number of feature maps. It enables to control the amount of newly generated information at a fixed growth rate. When H_l generates k feature maps, the number of input feature maps in the l^{th} layer is given by

$$l^{th} layer = k_0 + k \times (l - 1), \tag{2}$$

where k_0 is the number of feature maps in the input layer of the first dense block.

3.1.3 | Transition layer

The transition layer for the down-sampling layer is referred between dense blocks. This layer consists of convolution and pooling layers, which reduce the size of feature maps. We merge representative features from social lending data through the transition layer. The hyperparameter θ, as a compressor, is designed to achieve model compactness in the transition layer. θ has a value between 0 and 1. When the number of feature maps passed a dense block is n, the number of feature maps passed the transition layer is θn.

3.1.4 | Classification layer

The combination of global average pooling layer and softmax classifier is used as the classification layer for the repayment prediction. The feature maps generated by repeating several dense blocks and transition layers from the social lending data are arranged in the feature vector $f^l = [f_1, f_2, \ldots, f_l]$ in one dimension through the global average pooling layer. The final layer, the softmax classifier, represents the loan status d of the borrower (charged off and fully paid).

$$P(d|f) = \underset{d \in D}{\operatorname{argmax}} \frac{\exp\left(f^{L-1} w^L + b^L\right)}{\sum_{k=1}^{N_D} \exp\left(f^{L-1} w_k\right)}, \tag{3}$$

where L represents the index of the final layer, b represents the bias, w represents the connected weight, and N_D is the number of classes of loan status.

3.2 | Architecture

This section describes the details of the architecture and training of DenseNet. The goal of the overall architecture is to minimize the complexity of the model and reduce the size of the parameters while maintaining the depth of an optimal network. However, DenseNet includes many structures based on the combination of hyperparameters, which affect the process of capturing the features of the borrower (He & Sun, 2015). To design an optimal structure, it is necessary to understand the domain; hence, in this study, we aim to predict the repayment involved in social lending from the borrower's information.

3.2.1 | Structure

We investigate various architectures for learning good representations using validation sets to use the architecture suitable for repayment prediction. We use the best-performed architecture as described in Table 2. Our structure consists of two dense blocks and one transition layer. The convolutional and pooling layers in DenseNet use learnable parameters, and these layers are optimized during training.

The lending data from 64 feature maps pass an initial convolution. The dense blocks use 1×1 and 3×3 convolution as a bottleneck layer. The transition layer consists of 1×1 convolution and average pooling. After each convolutional layer in the dense block and transition layer, a dropout with a 0.25 probability (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014) is applied. In the classification layer, global average pooling is performed followed by the softmax classifier. Figure 2 shows the proposed dense block and transition layer.

3.2.2 | Training

After modelling the network, it is necessary to define the loss function to evaluate and improve the results. DenseNet is trained using optimization algorithms to minimize cross-entropy. The errors are propagated forward. The weights of DenseNet are updated using stochastic gradient descent (Bottou, 2010), which minimizes categorical cross-entropy in minibatches of the lending data.

TABLE 2 The proposed DenseNet structure

Layers	Output size	No. of Parameter	Configuration
Convolutional layer	(72, 64)	256	1 × 3 conv, stride 1, zero padding
Pooling layer	(73, 64)	-	1 × 2 max pool, stride 1, zero padding
Dense block	(73, 448)	196,608	$\begin{pmatrix} 1\times 1\ conv \\ 1\times 3\ conv \end{pmatrix} \times 3$
Transition layer	(72, 224)	100,352	1 × 1 conv, stride 1 1 × 2 avg pool, stride 1
Dense block	(72, 608)	196,608	$\begin{pmatrix} 1\times 1\ conv \\ 1\times 3\ conv \end{pmatrix} \times 3$
Classification layer	(608) (2)	- 1,218	Global avg pool Fully connected, softmax

$$\theta_{t+1} = \theta_t - \eta \sum_{i=1}^{n} \nabla_\theta J(\theta_t)/n, \tag{4}$$

where θ represents the parameters of DenseNet (not compression), η is the learning rate, n is the batch size, and $\nabla_\theta J(\theta)$ represents the gradient of the objective function $J(\theta)$.

We train our model using stochastic gradient descent with a 10^{-6} weight decay and 0.9 momentum. It is trained for 500 epochs with a batch size of 512. The learning rate is set to 10^{-2} initially and is divided by 5 at every 100th epoch.

3.3 | Dropout

Deep neural networks are generally excessively parameterized and are likely to overfit. Regularization techniques have been introduced to address these problems. One of the regularization techniques, dropout, randomly sets activation to zero with some probability during training (Srivastava et al., 2014). It is performed independently for each node of the training data. This prevents coadaptation of neurons during training. If the dropout probability is very large or very small, overfitting or underfitting may occur, respectively. We set the dropout at the probability of 0.25 after all convolutional layers.

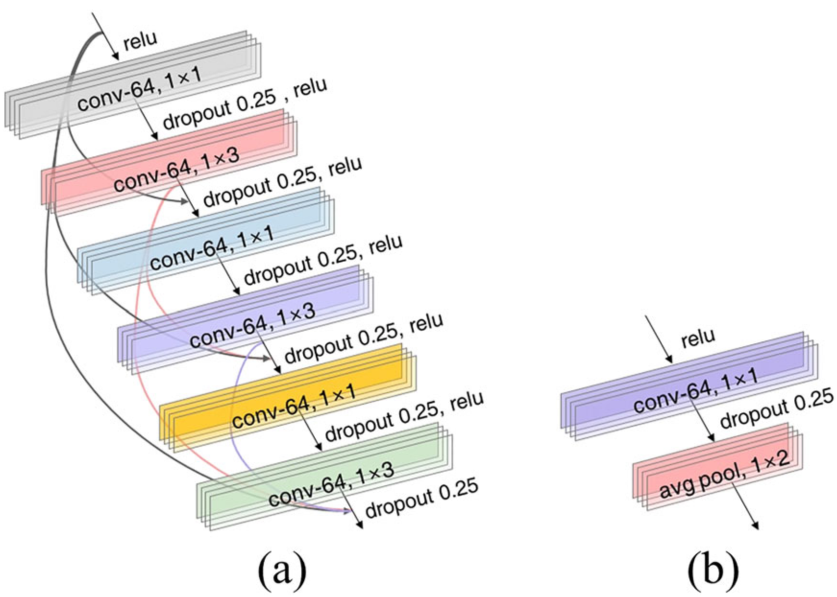

FIGURE 2 (a) Dense block and (b) transition layer

4 | EXPERIMENTS

4.1 | Dataset

In this study, we used the dataset of the Lending Club, which is the loan platform with the highest number of loans in the United States (https://www.lendingclub.com). Since 2007, Lending Club has been providing borrowers' information such as loan purpose, annual income, and home ownership. In this dataset, the loan status consists of "fully paid," "charged off," "current," "In grace period," "Late (16–30 days and 31–120 days)," and "default." In this study, only the loan records belonging to fully paid and charged off were selected because we predicted the repayment of the borrower whose repayment period was expired. From the data of 2015–2016, 152,161 loan entries were used in the experiment (March 2017) from which 33,745 entries (22.2%) were charged off and 118,416 (77.8%) were fully paid.

Many related works have used the Lending Club dataset to study repayment prediction, and some preprocessing methods were employed by Vinod Kumar et al. (2016). The dataset consists of 110 attributes including borrower information, borrower's credit history, loan characteristics, and loan status, which are predictive variables as shown in Table 1. Attributes that cannot be used for prediction, such as borrower's ID, URL, and loan explanation, attributes with more than 50% of missing values, redundant attributes, attributes with the same value, and attributes that are filled after the borrower starts to repay, were not considered. After the preprocessing, 63 attributes and 143,823 instances were used. Table 3 lists the attributes that were not considered, and Table 4 lists the features used in the experiment.

All variables are scaled in the [0, 1] range to be used as inputs to DenseNet. The categorical variables are created as dummy variables, and the continuous variables are min–max normalized using Equation (5) (Kantardzic, 2011) by removing 1% of the outliers. The dataset is split in the ratio 80:20 for the training and test sets, respectively.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}. \tag{5}$$

4.2 | Experimental set-up

The experiment consists of a comparison with other methods and an analysis of misclassification cases to evaluate the usefulness of DenseNet. The first experiment compares the performance of deep learning models including CNN, Inception v3 (Szegedy et al., 2015), residual network (ResNet; He, Zhang, Ren, & Sun, 2016), and Inception-ResNet (Szegedy, Ioffe, Vanhoucke, & Alemi, 2017) with various widths and depths in the test set. We demonstrate the performance of DenseNet by comparing the number of parameters used and accuracy. We also compare the cross-entropy and accuracy of the validation set for CNN, Inception, ResNet, and DenseNet to verify if overfitting occurs. The second experiment presents the performance comparisons between conventional machine learning methods and above-mentioned deep learning models in the validation set. A fivefold cross-validation is employed to evaluate the overall performance and generalization ability of the classifier (Brown & Mues, 2012). A statistical analysis is performed for quantitative evaluation. The third experiment presents the analysis misclassification cases using a confusion matrix. We can verify the features that DenseNet did not extract by performing this analysis. The data of the well-classified borrowers and those of the misclassified borrowers are filtered, and the difference depending on the loan status is examined.

The hyperparameters are adjusted while maintaining the best configuration in the validation set. We test the values of the growth rate k from 32 to 128, compression θ from 0 to 1 added by 0.25, and dropout from 0 to 1 added by 0.25. We saved the model that achieved the highest performance in the validation set. Table A1 shows the comparison of the performance by the hyperparameters in the validation set.

TABLE 3 List of eliminated attributes

Category	Eliminated variables
Attributes that cannot be used for prediction	id, member_id, emp_title, url, desc, title
Attributes with more than 50% missing values	mths_since_last_delinq, mths_since_last_major_derog, mths_since_recent_bc_dlq, mths_since_recent_revol_delinq, open_il_6m, open_il_12m, open_il_24m, mths_since_rcnt_il, total_bal_il, il_util, mths_since_last_record, open_rv_12m, open_rv_24m, max_bal_bc, open_acc_6m, all_util, inq_fi, total_cu_tl, inq_last_12m
Redundant attributes	annual_inc_joint, dti_joint, verified_status_joint
Attributes with the same value	policy_code, pymnt_plan, out_prncp, out_prncp_inv
Attributes that are filled after the borrower starts to repay	grade, sub_grade, funded_amnt, funded_amnt_inv, issue_d, total_pymnt, total_pymnt_inv, total_rec_prncp, total_rec_int, total_rec_late_fee, recoveries, collection_recovery_fee, last_pymnt_d, last_pymnt_amnt, next_pymnt_d, tot_coll_amt

TABLE 4 List of the used feature

Category	Feature	Description	Type
Predictor	Loan status	Charged off, fully paid	Binary
Borrower info	Annual income	The self-reported annual income	Numeric
	Employment length	Employment length in years (<1–10>)	Nominal
	Home ownership	Rent, own, mortgage, other	Nominal
	Address state/zip code	The state and zip code provided by the borrower in the loan application	Nominal/ Numeric
Loan product info	Loan amount	The listed amount of the loan applied for by the borrower	Numeric
	Term	36 or 60 months	Nominal
	Interest rate	Interest rate on the loan	Numeric
	Instalment	The monthly payment owed by the borrower	Numeric
	Purpose	Purpose for the loan request	Nominal
	Verification status	If income was verified, not verified, or if the income source was verified by LC	Nominal
	Initial list status	The initial listing status of the loan	Nominal
	Debt to income	A ratio calculated using the borrower's total monthly debt payments on the total debt obligations	Numeric
	Application type	Individual application or a joint application with two co-borrowers	Nominal
	Last credit pull day	The most recent month LC pulled credit for this loan	Date
Credit info	Total current balance	Total current balance of all accounts	Numeric
	Total bankcard limit	Total bankcard high credit/credit limit	Numeric
	Account now delinquent	The number of accounts on which the borrower is now delinquent	Numeric
	and so on (44 variables)		

LC: lending club

The hyperparameters of the deep learning methods used in the first experiment were set as follows: The CNN structure had two sets of convolutional and pooling layers and was trained for 100 epochs with a batch size of 512. Inception, ResNet, and Inception-ResNet were trained for 200 epochs with a batch size of 512, and the models with the highest performance were saved. Inception consists of three inception modules, and ResNet consists of one convolution and three identity blocks. The filter size was set to 1×3, and a 1×1 convolution for the bottleneck layer was used before the 1×3 convolution. A dropout with the probability of 0.25 was used before the last fully connected layer. The optimizer of the CNN model was RMSProp (Tieleman & Hinton, 2012), and Adadelta (Zeiler, 2012) was used for the remaining models.

The parameters of the machine learning method in the second experiment were set as follows: k was set to 3 in k-nearest neighbours, multi-layer perceptron had 15 hidden layers, support vector machine was set as the radial basis function kernel, DT was set with a maximum depth of 25, and RF was set with a maximum depth of 30. The maximum depths of DT and RF were chosen from 1 to 40.

4.3 | Results and analysis

4.3.1 | Performance comparison with deep learning methods

Table 5 shows the main results of the performance comparisons with other CNN methods in the test set. DenseNet-BC with k = 128 and θ = 0.5 achieved the highest performance. DenseNet-BC (k = 64) shows a remarkable performance with a fewer number of parameters than other models. Our model demonstrates the highest performance of 79.6%, which is about 3 percentage points higher than that of ResNet, which is one of the popular CNN models.

TABLE 5 Comparison with other CNN models

Model	Parameter	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CNN	2,242,370	75.9	80.7	90.7	85.4
Inception v3	2,430,818	76.5	80.4	92.5	86.0
ResNet	669,026	76.4	80.6	91.9	85.9
Inception-ResNet v4	3,300,610	77.8	79.3	**96.8**	87.2
DenseNet-BC (k = 64)	508,674	79.3	**81.8**	94.5	87.4
DenseNet-BC (k = 128)	2,117,058	**79.6**	81.2	96.2	**88.0**

We compared the cross-entropy and accuracy of the validation set for CNN, Inception, ResNet, and DenseNet to verify if overfitting occurs. In other CNN models except for DenseNet, the loss increased during training, but the loss of DenseNet decreased and accuracy steadily increased. Figure 3 shows the cross-entropy and accuracy with respect to the number of epochs for the four models.

4.3.2 | Performance comparison with other methods

We performed a fivefold cross-validation using the model that showed the highest performance. The deep learning methods show significantly higher performances than the conventional machine learning ones. Among the machine learning methods, tree-based methods such as RF and DT achieved high performance. Figure 4 shows the performance comparison for the fivefold cross-validation.

We performed a paired *t* test to demonstrate the usefulness of the proposed method. Table 6 shows the paired *t*-test results for the accuracy of DenseNet and the different classifiers in the fivefold cross-validation. In all the tests, DenseNet showed statistically significant differences at the significance level of 0.01.

4.3.3 | Analysis of misclassification cases

Table 7 shows the confusion matrix of DenseNet. Our model well classified the repaid borrowers, but not the nonrepaid borrowers. This may have occurred because there were fewer cases in which the repayment is not completed in the social lending dataset used.

Emekter, Tu, Jirasakuldech, and Lu (2015) presented the critical features in the repayment prediction model using the Lending Club dataset: interest rate, home ownership, revolving line utilized, total funded, etc. We compared the well-classified and misclassified data in the confusion matrix with respect to these features. Figure 5 shows the distribution of sample data as follows: The top row of Figure 5 shows the well-classified data (TP and TN), whereas the bottom row shows the misclassified data (FP and FN). The distributions of the misclassified and well-classified data are totally reverse.

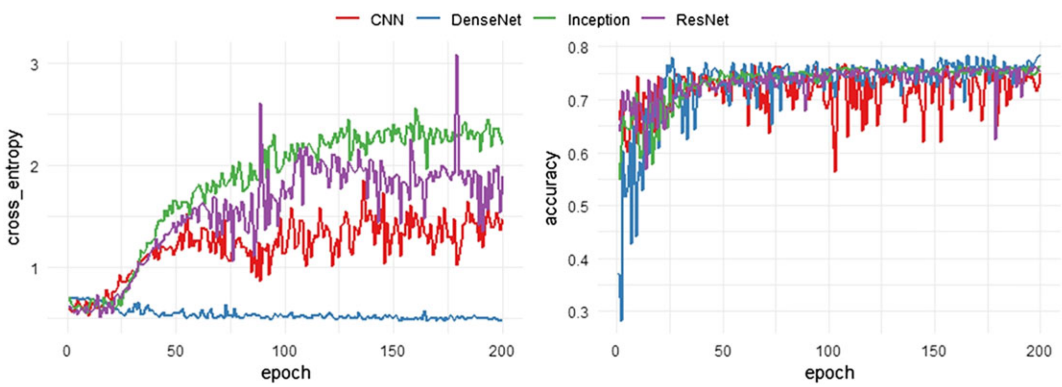

FIGURE 3 Graph of cross-entropy and accuracy with respect to the number of epochs for the four models

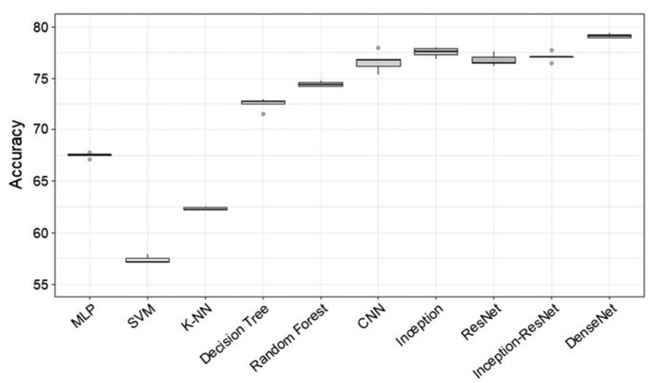

FIGURE 4 Comparison of accuracy from the fivefold cross-validation

TABLE 6 Paired *t*-test results for comparison of accuracy

Index	Model	Mean (±*SD*)	*t*	*P* value
1	MLP	0.675 (±0.241)	60.94	<0.001***
2	SVM	0.574 (±0.337)	99.71	<0.001***
3	k-NN	0.623 (±0.162)	113.49	<0.001***
4	Decision tree	0.725 (±0.576)	23.85	<0.001***
5	Random forest	0.744 (±0.211)	34.87	<0.001***
6	CNN	0.766 (±0.958)	6.29	0.002***
7	Inception	0.775 (±0.473)	6.85	0.001***
8	ResNet	0.767 (±0.580)	9.94	<0.001***
9	Inception-ResNet	0.771 (±0.469)	8.25	0.001***
10	DenseNet	0.791 (±0.198)	-	-

$^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$

TABLE 7 Confusion matrix of the proposed model

Predicted\true	Fully paid	Charged off
Fully paid	32,395 (TP)	7,516 (FP)
Charged off	1,295 (FN)	1,939 (TN)

4.4 | Discussion

Various features of P2P social lending are related to the borrowers' ability to repay (Ye, Dong, & Ma, 2018). Many feature extraction and feature selection studies have been conducted for the repayment prediction owing to the influence of the features used in the repayment prediction on the classification performance. Although the proposed method achieves high performance by automatically extracting powerful and diverse features, some constraints still exist. We consider two aspects of P2P lending and the model.

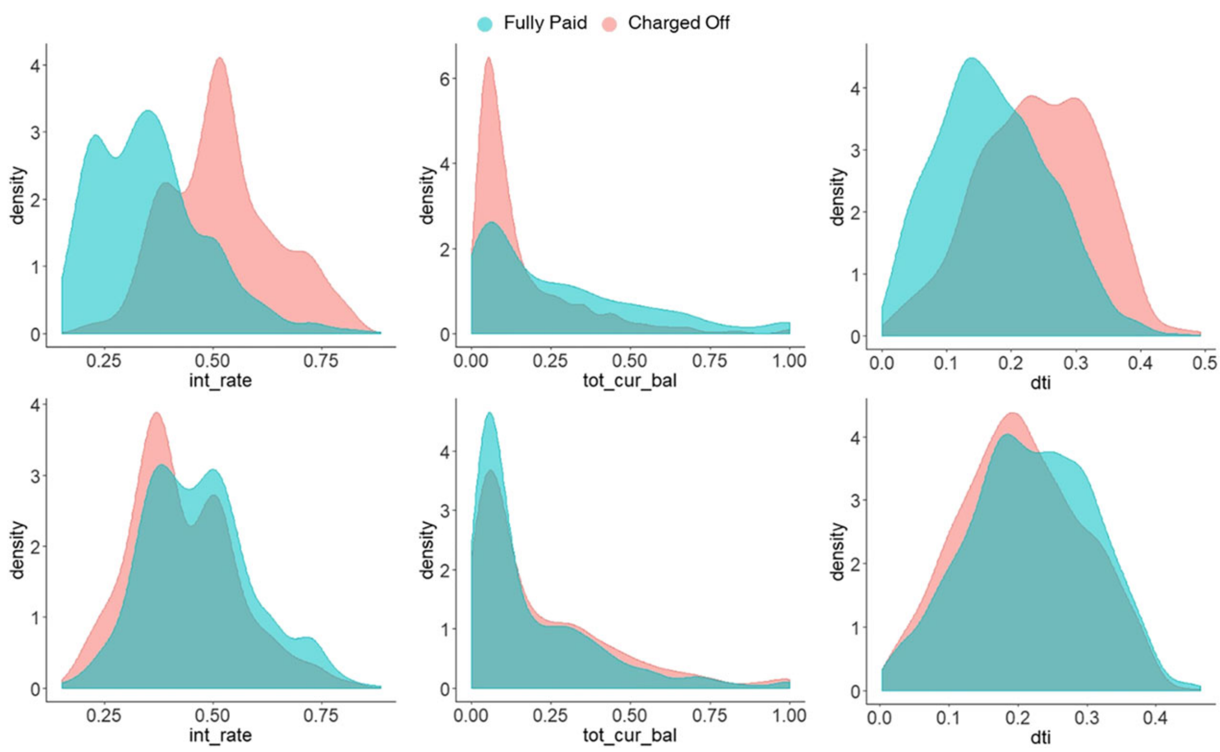

FIGURE 5 Comparison of distribution from classified samples

First, diverse models extract different features according to the parameters. It is possible to design a classifier for a better repayment prediction by combining diverse classifiers that can extract different features. Hence, it is necessary to verify which features have a significant impact on which structure. However, it is difficult to understand the internal operation of models accurately.

Second, P2P social lending data have a class imbalance problem. We presented the F1-score as a measure to address this problem, but we could not solve it completely. For data with class imbalances, many researchers changed the training data by oversampling, undersampling, or adjusting the class weights in the algorithm. Hence, a robust model is required to extract discriminative features efficiently from data with class imbalances.

5 | CONCLUSIONS

In this study, we proposed a structure of DenseNet for predicting repayment in social lending. DenseNet maintains low- and high-level semantic information of lending data with dense connectivity and automatically extracts relevant and powerful features. In the experiments with the Lending Club dataset, we demonstrated that the proposed DenseNet model achieved the highest performance without overfitting compared with CNN, Inception, and ResNet, which are the most advanced CNN models. The analysis of misclassified instances verified the extracted feature. Our model was very effective in predicting repayment; thus, our proposed model can help classify the loan status of the borrower.

Because P2P social lending is traded online, one can collect a variety of information about borrowers. For example, it is possible to obtain social network data from a borrower's Facebook or Twitter account. This may have a significant potential for prediction models that are designed by combining loan information from borrowers with additional information from social networks. P2P social lending data also account for more than half of the data whose repayment has not expired. We expect that using this data to design a repayment prediction model will achieve good performance.

P2P lending platforms are increasing around the world. In addition to the Lending Club in the United States, it also provides available data on certain lending platforms. The representative social lending platforms are Prosper, Kiva, and Zidisha. They operate differently and provide data with different attributes. Unlike Lending Club, Kiva and Zidisha are nonprofit systems that do not charge interest and lend small amounts of money. We need to consider experiments with other social lending datasets to evaluate the generalization ability. We assume that borrowers using each platform will have discriminative features, and our model expects to extract these features automatically, which constitutes the future work to be addressed. In the future, we will also study an automatic deep learning system that extracts features from loan application in real-time transactions and predicts defaults.

ACKNOWLEDGEMENT

This research was supported by Korea Electric Power Corporation (Grant: R18XA05).

ORCID

Sung-Bae Cho https://orcid.org/0000-0002-7027-2429

REFERENCES

Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010* (pp. 177–186). https://doi.org/10.1007/978-3-7908-2604-3_16

Brown, I., & Mues, C. (2012). An experimental comparison of classification algorithms for imbalanced credit scoring data sets. *Expert Systems with Applications*, *39*(3), 3446–3453. https://doi.org/10.1016/j.eswa.2011.09.033

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19*(2), 171–209. https://doi.org/10.1007/s11036-013-0489-0

Chen, X.-W., & Lin, X. (2014). Big data deep learning: Challenges and perspectives. *IEEE Access*, *2*, 514–525. https://doi.org/10.1109/ACCESS.2014.2325029

Chen, Y. (2017). Research on the credit risk assessment of Chinese online peer-to-peer lending borrower on logistic regression model. *DEStech Transactions on Environment, Energy and Earth Science*, 216–221.

Emekter, R., Tu, Y., Jirasakuldech, B., & Lu, M. (2015). Evaluating credit risk and loan performance in online peer-to-peer (P2P) lending. *Applied Economics*, *47*(1), 54–70. https://doi.org/10.1080/00036846.2014.962222

Fu, Y. (2017). Combination of random forests and neural networks in social lending. *Journal of Financial Risk Management*, *6*(4), 418–426. https://doi.org/10.4236/jfrm.2017.64030

Guo, Y., Zhou, W., Luo, C., Liu, C., & Xiong, H. (2016). Instance-based credit risk assessment for investment decisions in P2P lending. *European Journal of Operational Research*, *249*(2), 417–426. https://doi.org/10.1016/j.ejor.2015.05.050

He, K., & Sun, J. (2015). Convolutional neural networks at constrained time cost. IEEE conference on computer vision and pattern recognition, 5353–5360.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *IEEE conference on computer vision and pattern recognition* (pp. 770–778).

Huang, G., Liu, Z., Matten, L., & Weingerger, K.-Q. (2017). Densely connected convolutional networks. In *IEEE conference on computer vision and pattern recognition*, New York, USA: Springer International Publishing, (pp. 4700–4708).

Huang, G., Sun, Y., Liu, Z., Sedra, D., & Weinberger, K.-Q. (2016). Deep networks with stochastic depth. In *European conference on computer vision* (pp. 646–661).

Jiang, C., Wang, Z., Wang, R., & Ding, Y. (2017). Loan default prediction by combining soft information extracted from descriptive text in online peer-to-peer lending. *Annals of Operations Research*, 1–19.

Kantardzic, M. (2011). *Data mining: Concepts, models, methods, and algorithms*John Wiley & Sons. https://doi.org/10.1002/9781118029145

Kim, T.-Y., & Cho, S.-B. (2018). Web traffic anomaly detection using C-LSTM neural networks. *Expert Systems with Applications*, *106*, 66–76. https://doi.org/10.1016/j.eswa.2018.04.004

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceeding of the IEEE*, *86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Lin, X., Li, X., & Zheng, Z. (2017). Evaluating borrower's default risk in peer-to-peer lending: Evidence from a lending platform in China. *Applied Economics*, *49*(35), 3538–3545. https://doi.org/10.1080/00036846.2016.1262526

Malekipirbazari, M., & Aksakalli, V. (2015). Risk assessment in social lending via random forests. *Expert Systems with Applications*, *42*(10), 4621–4631. https://doi.org/10.1016/j.eswa.2015.02.001

Milne, M., & Parboteeah, P. (2016). The business models and economics of peer-to-peer lending. *European Credit Research Institute*, (17), 1–31.

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. In *International conference on machine learning*, Haifa, Israel: International Machine Learning Society, (pp. 807–814).

Najafabadi, M. M., Villanustre, F., Khoshgoftaar1, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of Big Data.*, *2*(1), 1. https://doi.org/10.1186/s40537-014-0007-7

Namvar, A., Siami, M., Rabhi, F., & Naderpour, M. (2018) Credit risk prediction in an imbalanced social lending environment. arXiv preprint arXiv:1805.00801.

Philip Chen, C. L., & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on big data. *Information Sciences*, *275*, 314–347. https://doi.org/10.1016/j.ins.2014.01.015

Polena, M., & Regner, T. (2016). Determinants of borrowers' default in P2P lending under consideration of the loan risk class. *Jena Economic Research Papers*, (23), 1–30.

Qiu, J., Wu, Q., Ding, G., Xu, Y., & Feng, S. (2016). A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, *2016*(1), 67. https://doi.org/10.1186/s13634-016-0355-x.

Ronao, C. A., & Cho, S.-B. (2016a). Anomalous query access detection in RBAC-administered databases with random forest and PCA. *Information Sciences*, *369*, 238–250. https://doi.org/10.1016/j.ins.2016.06.038

Ronao, C. A., & Cho, S.-B. (2016b). Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Systems with Applications*, *59*, 235–244. https://doi.org/10.1016/j.eswa.2016.04.032

Serrano-Cinca, C., & Gutiérrez-Nieto, B. (2016). The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending. *Decision Support Systems*, *89*, 113–122. https://doi.org/10.1016/j.dss.2016.06.014

Serrano-Cinca, C., Gutiérrez-Nieto, B., & López-Palacios, L. (2015). Determinants of default in P2P lending. *PLoS ONE*, *10*(10), e0139427. https://doi.org/10.1371/journal.pone.0139427

Sohangir, S., Wang, D., Pomeranets, A., & Khoshgoftaar, T. M. (2018). Big data: Deep learning for financial sentiment analysis. *Journal of Big Data*, *5*(1), 3.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*(1), 1929–1958.

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In *The AAAI conference on artificial intelligence*, Palo Alto, USA: AAAI Press, (Vol. 4) (p. 12).

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. IEEE conference on Computer Vision and Pattern Recognition, 1–9.

Tieleman, T., & Hinton, G. (2012). Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. COURSERA: Neural networks for. *Machine Learning*, *2*(4), 26–31.

Vinod Kumar, L., Natarajan, S., Keerthana, S., Chinmayi, K. M., & Lakshmi, N. (2016). Credit risk analysis in peer-to-peer lending system. IEEE International Conference on Knowledge Engineering and Applications, 193–196.

Xie, S., Girshick, R., Dollar, P., Tu, Z., & He, K. (2017). Aggregated residual transformations for deep neural networks. IEEE conference on Computer Vision and Pattern Recognition, 5987–5995.

Xu, J., Chen, D., & Chau, M. (2016). Identifying features for detecting fraudulent loan requests on P2P platforms. IEEE Conference on Intelligence and Security Informatics, 79–84.

Yan, J., Yu, W., & Zhao, J. L. (2015). How signaling and search costs affect information asymmetry in P2P lending: The economics of big data. *Financial Innovation*, *1*(1), 19. https://doi.org/10.1186/s40854-015-0018-1

Ye, X., Dong, L., & Ma, D. (2018). Loan evaluation in P2P lending based on random forest optimized by genetic algorithm with profit score. *Electronic Commerce Research and Applications*, *32*, 23–36. https://doi.org/10.1016/j.elerap.2018.10.004

Zagoruyko, S., & Komodakis, N. (2016). Wide residual networks. *arXiv preprint arXiv*:1605.07146.

Zeiler, M. D. (2012). ADADELTA: An adaptive learning rate method. *arXiv preprint arXiv:1212.5701*.

Zhang, Y., Li, H., Hai, M., Li, J., & Li, A. (2017). Determinants of loan funded successful in online P2P lending. *Procedia Computer Science*, *122*, 896–901. https://doi.org/10.1016/j.procs.2017.11.452

Zhao, H., Ge, Y., Liu, G., Wang, G., Chen, E., & Zhang, H. (2017). P2P lending survey: Platforms, recent advances and prospects. *ACM Transactions on Intelligent Systems and Technology*, *8*(6), 72.

AUTHOR BIOGRAPHIES

Ji-Yoon Kim received her M.S. degree in computer science from Yonsei University. Her research interests include probabilistic recognition models, and deep learning.

Sung-Bae Cho received the B.S. degree in computer science from Yonsei University, Seoul, Korea, and the M.S. and PhD degrees in computer science 1from KAIST (Korea Advanced Institute of Science and Technology), Taejeon, Korea. He was an Invited Researcher of Human Information Processing Research Laboratories at ATR (Advanced Telecommunications Research) Institute, Kyoto, Japan, from 1993 to 1995 and a Visiting Scholar at University of New South Wales, Canberra, Australia, in 1998. He was also a Visiting Professor at University of British Columbia, Vancouver, Canada, from 2005 to 2006. Since 1995, he has been a Professor in the Department of Computer Science, Yonsei University. His research interests include neural networks, pattern recognition, intelligent man-machine interfaces, evolutionary computation, and artificial life. Dr Cho was awarded outstanding paper prizes from the IEEE Korea Section in 1989 and 1992, and another one from the Korea Information Science Society in 1990. He was also the recipient of the Richard E. Merwin prize from the IEEE Computer Society in 1993. He was listed in Who's Who in Pattern Recognition from the International Association for Pattern Recognition in 1994 and received the best paper awards at International Conference on Soft Computing in 1996 and 1998. Also, he received the best paper award at World Automation Congress in 1998 and listed in Marquis Who's Who in Science and Engineering in 2000 and in Marquis Who's Who in the World in 2001. He is a Senior Member of IEEE and a Member of the Korea Information Science Society, INNS, the IEEE Computational Intelligence Society, and the IEEE Systems, Man, and Cybernetics Society.

How to cite this article: Kim J-Y, Cho S-B. Predicting repayment of borrows in peer-to-peer social lending with deep dense convolutional network. *Expert Systems*. 2019;e12403. https://doi.org/10.1111/exsy.12403

APPENDIX A

TABLE A1 Comparison of performance by hyperparameters in the validation set

Structure	Growth rate	Compression	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
DenseNet-B	32	0.25	73.3	85.1	80.2	82.5
		0.5	74.8	84.3	83.5	83.9
		0.75	73.8	84.9	81.2	83.0
	64	0.25	69.7	86.3	73.1	79.2
		0.5	70.7	86.1	74.9	80.1
		0.75	73.1	85.5	79.3	82.3
	128	0.25	75.4	84.3	84.5	84.4
		0.5	75.9	84.7	84.8	84.7
		0.75	75.0	84.5	83.5	84.0
DenseNet-C	32	0.25	74.9	84.4	83.6	84.0
		0.5	75.0	84.3	83.9	84.1
		0.75	71.7	85.9	76.6	81.0
	64	0.25	72.9	85.3	79.1	82.1
		0.5	67.2	87.2	68.2	76.6
		0.75	72.9	85.1	79.3	82.1
	128	0.25	73.9	84.9	81.3	83.0
		0.5	75.8	83.7	86.0	84.8
		0.75	69.9	86.4	73.2	79.3
DenseNet-BC	32	0.25	78.8	79.1	99.0	88.0
		0.5	78.7	82.1	93.1	87.2
		0.75	78.5	80.6	95.5	87.4
	64	0.25	78.8	79.9	97.5	87.8
		0.5	79.3	83.6	91.5	87.4
		0.75	79.2	81.5	95.1	87.7
	128	0.25	79.0	79.6	98.5	88.0
		0.5	79.6	82.4	94.1	87.9
		0.75	79.0	79.9	97.8	87.9