

Machine Learning Framework for Multi-level Classification of Company Revenue

Jung-gu Choi^{1*}, Inhwon Ko^{2*}, Jeongjae Kim¹, Yeseul Jeon³, and Sanghoon Han^{1,2}

¹Yonsei graduate program in Cognitive science, Yonsei University, Seoul, Republic of Korea

²Department of Psychology, Yonsei University, Seoul, Republic of Korea

³Department of Applied Statistics, Yonsei University, Seoul, Republic of Korea

Corresponding author: Sanghoon Han (e-mail: sanghoon.han@yonsei.ac.kr).

*These authors contributed equally

This work was supported by the Ministry of Education of the Republic of Korea and National Research Foundation of Korea (NRF-2020S1A5A2A03042694)

ABSTRACT The planning and execution of a business strategy are important aspects of the strategic human resource management of a company. In previous studies, machine learning algorithms were used to determine the main factors correlating employees with company performance. In this study, we introduced a method based on machine-learning algorithms for the classification of company revenue. Both annual and integrated datasets were examined to evaluate the classification performance of the framework under both binary and multiclass conditions. The performance of the proposed method was validated using six evaluation metrics: accuracy, precision, recall, F1-score, receiver operating characteristic curve, and area under the curve. As the experimental results indicate, the XGBoost classifier displayed the best classification performance among the three algorithms (XGBoost classifier, stochastic gradient descent classifier, and logistic regression) used in this study. Moreover, we confirmed the important features of the trained XGBoost model in accordance with variables focusing on human resource management studies. These results demonstrate that the proposed framework has strength in terms of both classification and practical implementation. This study provides novel insights into the relationship between employees and the revenue levels of their employer.

INDEX TERMS Machine learning, company revenue level, classification algorithm, human resource management

I. INTRODUCTION

Estimating the future performance of a company has been a subject of significant interest in various fields. Researchers in the field of human resource management (HRM) have focused on the relationship between organizations and their performance. To reflect various elements of a business practice, strategic human resource management (SHRM) studies have been conducted based on traditional HRM topics. In contrast to previous research focusing mainly on specific subjects, studies on SHRM have emphasized the planning deployment of human resources and activities to achieve the goals of an organization [1].

Many SHRM studies have provided perspectives for both goal-oriented and mutually balanced plans through the alignment of HRM elements [2, 3]. Such perspectives are used to emphasize the role of HRM in supporting business strategies [4]. In this context, researchers have been

interested in the relationships between HRM policy and company-level performance [5–7]. A number of studies conducted in a wide range of industries such as banking [8], steel manufacturing [9, 10], and automotive assembly [11] have suggested an integrative approach with regard to both goals, and their means are essential in HRM practices [12]. In addition, it is important to consider the HR policy from an employee's perspective. Employees in an organization interact with HR systems more than HR managers. In addition, employees are affected by the HRM policy itself [13]. Furthermore, because of individual differences, policies can be interpreted in different ways [14]. For this reason, survey data such as organizational culture, satisfaction, stress level, communication, and trust collected by the members of an organization are particularly useful because they can represent various types of feedback regarding their organization in a more realistic way.

However, the association between revenue at the company level and employee characteristics has yet to be sufficiently studied.

To examine the various components of HRM, a dataset with a large amount of information is essential for research purposes. However, in several HRM studies, researchers have used relatively small datasets because it is extremely difficult to collect and handle datasets collected from many different subjects. Notably, in the field of HRM, researchers are interested in the application of big data. The analysis of big data facilitates better decision-making processes in various HRM practices, including competitiveness, innovation, and efficiency in an organization [15, 16]. In addition, recent studies have shown that the application of various analysis methods has led to the identification of latent patterns or relationships [17, 18].

In previous studies, many researchers used different methods to investigate embedded patterns in various datasets. They used statistical modeling to analyze the relationships in structural data, including employee survey data. Owing to an increase in the amount of data and computing resources, diverse machine learning and deep learning methods have been applied to various domains. Machine learning algorithms have been used to estimate the costs or forecast dangerous events in the economic domain [19–22]. In addition, previous studies have shown that a machine learning model with a time series can be applied to the stock market [23, 24].

In the field of human resource management, previous studies have shown applicable models based on machine learning models. These studies have suggested a model for the prediction of employee turnover using supervised machine-learning algorithms [25]. In addition, an online recruitment system using ensemble machine-learning models was introduced [26]. Several previous studies have focused on performance management related to employee performance evaluation [27]. In these studies, the researchers focused on the relationship between the characteristics of the collaboration process and the teamwork performance [28].

Based on previous studies applying machine learning models, we hypothesize that machine learning algorithms have the potential to classify company revenue levels from employee characteristics. To prove this hypothesis, from a large-scale employee survey dataset, we extracted 312 employee-related variables as features based on the survey categories (from among 1712 variables in the original dataset). The selected features were applied to three classifiers, i.e., logistic regression, stochastic gradient descent (SGD), and XGBoost classifiers. Finally, the performance of the classifiers was evaluated based on six evaluation metrics: accuracy, precision, recall, F1-score, receiver operating characteristic (ROC) curve, and area under the curve (AUC).

The objective of this research is to develop a novel methodology based on a machine learning algorithm for the

classification of company revenue in reference to employee status. The major contributions of this paper are as follows:

- (1) We propose a machine learning based framework for the classification of company revenue based on the characteristics or status of the employees. In addition, we evaluate the performance of the model under various conditions, including both binary and multiclass classification, on a large dataset. Moreover, we compared the classification performance of popular machine learning classification algorithms: the XGBoost classifier, SGD classifier, and logistic regression.
- (2) Advancing from the simple training of machine learning models, our framework provides practical implications in terms of HRM. Because the methodology used in our research can be generally applied to companies in diverse situations, the framework meets the needs of company-level performance managers. In addition, we suggest several insights regarding the HR policy design from the feature importance of the trained XGBoost model corresponding with previous HRM studies.

II. METHODS

A. OVERVIEW

This study consists of four steps. First, we collected and combined three sub-datasets from the Human Capital Corporate Panel (HCCP) dataset and employee-related features selected from it. Second, complete data for all features were selected from the datasets, and a log transform was applied to financial information data, and two types of datasets were generated to validate the framework. Third, three types of machine learning models were constructed from the datasets. Finally, the models were evaluated using the performance indices. The detailed steps are shown in Fig. 1.

B. DATA SOURCES

In this study, we utilized the HCCP dataset released by the Korea Research Institute for Vocational Education and Training (KRIVET) [29]. The HCCP dataset is a longitudinal panel dataset collected to investigate the status of human resources of different companies in Korea since 2005.

This dataset is composed of three types of sub-datasets. First, a dataset of survey results was collected from company employees. Second, a dataset was collected from the management team of each company. Third, financial information data of the companies participating in the survey were gathered. Except for the financial information dataset, all data are from a 7-year period. Each dataset can be combined with the company ID and worker ID columns.

In total, 74 732 employees from 3113 companies participated in the survey. The detailed characteristics of the dataset are shown in Table I.

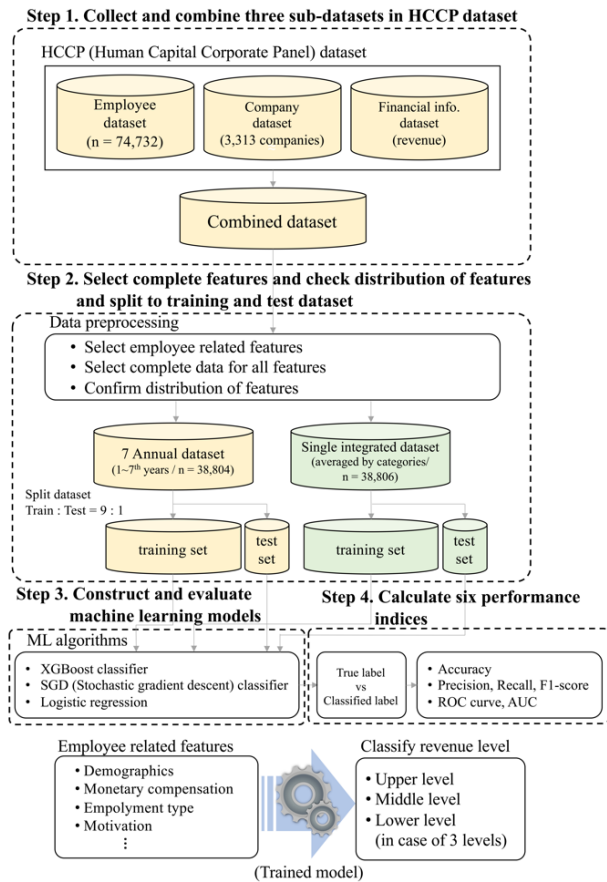


FIGURE 1. Overview of detailed steps of study design.

TABLE I
BASELINE CHARACTERISTICS OF HCCP DATASET.

| Characteristic | HCCP ^a |
|----------------------------------|-----------------------------|
| Age (years), mean (SD) | 46.3 (8.47) |
| Gender, n(%) | |
| Male | 60 232 (80.59%) |
| Female | 14 500 (19.40%) |
| Number of workers | 74 732 |
| Number of companies | 3113 |
| Revenue ^b , mean (SD) | 612 504 310 (2 622 524 467) |

^aHCCP: Human Capital Corporate Panel dataset

^bRevenue in KRW

In addition, the data investigated in the HCCP dataset were organized into 6906 features with 25 categories. There are 24 employee-related categories and a single financial information category. In the dataset, the features include variables from the management team (Nos. 1–6) and employees (Nos. 7–24). Based on the features of the management team, we can check company-related information (e.g., the scale of the company) rather than employee-related information.

More employee-focused information (e.g., demographics, satisfaction, and stress levels) was composed of employee features. The detailed categories of the features in the dataset are shown in Table II.

TABLE II
CATEGORIES OF FEATURES IN HCCP DATASET

| No. | Categories | Types | No. | Categories | Types |
|-----|----------------------------|-------------|-----|--|-------------|
| 1 | General Environment | Categorical | 14 | Teamwork | Categorical |
| 2 | HR department | Categorical | 15 | Workforce Competency | Categorical |
| 3 | Human Resource Management | Categorical | 16 | Job Proficiency | Categorical |
| 4 | Human Resource Development | Categorical | 17 | Organizational Culture | Categorical |
| 5 | Headcount | Categorical | 18 | Job Satisfaction, Commitment, and Stress | Categorical |
| 6 | Research & Development | Categorical | 19 | Career Break | Categorical |
| 7 | Current Job Information | Categorical | 20 | Employment and Work condition | Categorical |
| 8 | Perceived HR activity | Categorical | 21 | Personal Information | Categorical |
| 9 | Employee training | Categorical | 22 | Information System | Categorical |
| 10 | Career History | Categorical | 23 | Performance Evaluation | Categorical |
| 11 | Skills | Categorical | 24 | Senior Support | Categorical |
| 12 | Certification | Categorical | 25 | Financial Information of company | Continuous |
| 13 | Job Understanding | Categorical | | | |

All categorical features consist of 5-point scale questions.

C. DATA PREPROCESSING

1) MERGE SUB-DATASET AND SELECT CATEGORIES OF FEATURES

To organize the data based on a single subject, we combined three sub-datasets. (a) In the company dataset, we selected features of the company to reflect the effect of the company scale. The row and feature dimensions of the original company datasets including the first through the seventh years were respectively (4275, 5043). After extraction of the company-scale features, only datasets with dimensions of (4275, 2) remained (company ID feature and company scale feature). (b) The financial information dataset included 14 130 rows and 150 features. We selected two features from the original financial dataset (company ID, revenue). We set the revenue features as predictors of the machine learning model. (c) Based on the employee dataset, we combined these selected features from two datasets (company and financial datasets). For the company scale and revenue features, there is a difference among the levels of the worker dataset, scale, and revenue features. In this study, we combined datasets based on the company ID to which the subject belongs and analyzed the dataset at the subject level. As a result, an employee dataset merged with company-scale and revenue features showed row and column dimensions of (74 774, 1715), as compared to the original employee datasets (74 774, 1713).

In the HCCP dataset, there are a total of 21 categories of features related to employees and companies. To select features directly related to employees, we selected eight categories (demographics, monetary compensation, employment type, stress, organizational commitment,

motivation, satisfaction, organizational culture, communication and trust, and talent management) among 21 categories. After selecting 8 employee-related categories, only 312 features were selected, and the dimensions of the remaining dataset were (74 774, 312).

2) SELECT COMPLETE DATA FOR ALL FEATURES

In the HCCP dataset, a non-response or unknown data were coded as -9 and -8. We confirmed that a non-response or missing data were included in each variable. To proceed with a precise analysis, all non-responses and missing data included in the dataset were removed. After this step, the complete dataset had dimensions of (38 763, 312) without missing or non-response data.

3) DISTRIBUTION CONFIRMATION OF SELECTED FEATURES

After removing missing and non-response data, we checked the feature distributions. When training machine learning models, the distribution of datasets or features is critical. Features with skewed distributions cause disturbances in the model training and a decline in the model performance. In this step, we confirmed the distributions of the selected features and applied a log transformation to features with skewed distributions.

4) GENERATE ANNUAL AND INTEGRATED DATASET

Two types of datasets were constructed to develop and validate the proposed framework. The first dataset has seven annual datasets from first to the seventh year. From each annual dataset, a total of 312 features were selected from the completed dataset. The dimensions of the first year dataset were (8571, 33), and they were (5897, 40) for the second year, (4837, 41) for the third year, (4477, 49) for the fourth year, (5012, 49) for the fifth year, (4908, 50) for the sixth year, and (5061, 50) for the seventh year. For the seven datasets, the training and test datasets were constructed by dividing them into a ratio of 9:1.

The second is an integrated dataset that combines features belonging to the same category. Except for the revenue feature, the remains are categorical features, and most are on a 5-point scale. To integrate features at scale, binary features were excluded, and only 5-point scale features were averaged. All seven annual datasets were used to construct an integrated dataset. Because there are differences between the features collected each year, some integrated features are missing when constructing a single integrated dataset. To solve this problem, we imputed missing feature values using the average of the existing values in the same feature. Finally, the single integrated dataset consisted of 21 integrated features, and we confirmed the (38 763, 21) dimensions from the integrated dataset. In addition, the integrated dataset was split into training and test datasets at a ratio of 9:1.

To validate the classification performance of the models, we constructed additional datasets with diverse conditions

using different class levels of the same revenue feature. The classification performance was compared under three conditions: binary, three, and four classes.

5) EXPERIMENT FOR CLASSIFICATION PERFORMANCE

In this step, a total of 24 datasets were used to evaluate the performance of the proposed models (seven annual datasets \times three conditions for classification labels + single integrated dataset \times 3 conditions = total 24 datasets). Three machine learning algorithms were used to compare the performance under diverse conditions. After the experiment, we checked the feature importance of the classification algorithm and achieved a superior performance. To validate the framework in terms of HRM, we examined the selected features for both the annual and integrated datasets.

D. MACHINE LEARNING MODELS

1) XGBOOST CLASSIFIER

The XGBoost algorithm is a gradient boosting method based on a supervised learning algorithm [30]. The classification and regression tree (CART) method is used as the basis of the algorithm and is an ensemble model composed of decision tree models, following the principle of additive learning.

The objective function of the XGBoost algorithm using additive learning is as follows.

$$Obj = \sum_{i=1}^n L(y_i, y'_i) + \sum_{k=1}^K \Omega(f_k) \quad (1)$$

$$y'_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (2)$$

In (1), Obj indicates an objective function of the XGBoost algorithm. In this formula, function L is a loss function for the algorithm training. In addition, function Ω is a regularization term and represents the complexity of the tree. The k th decision tree is represented by a function f_k , and the functions that indicate the decision tree are elements of the function space F . The result classified through the function f_k for x_i is y'_i [30]. As a result, this algorithm considers the classification results from several decision trees rather than a single result.

2) STOCHASTIC GRADIENT DESCENT CLASSIFIER

The SGD method is an optimization algorithm used to determine the parameter values that minimize the objective function [31]. This method is extremely similar to the traditional gradient descent algorithm. However, it only calculates the derivative of the error of a random single instance in the data rather than all data. As a result, this makes the algorithm much faster than a gradient descent algorithm [32]. SGD methods can be applied to various machine-learning models. The machine learning

classification algorithm using the SGD method is called an SGD classifier.

3) LOGISTIC REGRESSION

Logistic regression is a supervised learning algorithm that predicts the probability of a data belonging to a certain class as a value of between zero and 1 and classifies it into a category according to the probability [33]. This algorithm calculates the log-odds of each feature and applies the sigmoid function to represent the probability that the data belongs to the corresponding class. A logistic regression algorithm that performs multiple classifications is called multinomial logistic regression or softmax regression [34]. It operates in the same manner as traditional logistic regression, which performs a binary classification, but differs in the number of classes, which is more than three nominal predictors.

E. EVALUATION METRIC

In our framework, we used six performance indices to evaluate the classification performance of the machine learning algorithms. To confirm the classification performance, we calculated the confusion matrix using the true positives (TPs), true negatives (TNs), false positives (FPs), and false negatives (FNs) [35]. Detailed information on the confusion matrix is shown in Figure 2.

| | | Predicted values | |
|---------------|--------------|------------------------|------------------------|
| | | Positive (1) | Negative (0) |
| Actual values | Positive (1) | True Positive (TP) | False Negative (FN) |
| | Negative (0) | False Positive (FP) | True Negative (TN) |

FIGURE 2. Confusion matrix.

We can then obtain four indicators: precision, recall, F1-score, and accuracy.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (6)$$

In addition, to evaluate the performance, we checked the ROC curve and AUC. The ROC curve indicates a trade-off between the true positive rate (TPR) and the false positive rate (FPR). The formulas of TPR and FPR are as follows:

$$True Positive Rate = \frac{TP}{TP + FN} \quad (7)$$

$$False Positive Rate = \frac{FP}{FP + TN} \quad (8)$$

A model with a good performance shows that the ROC curve is closer to the upper-left of the graph, or the AUC is sufficiently close to 1. In addition, to investigate the performance of the algorithms, we averaged the ROC curve and the AUC of each class in a multiclass classification.

F. TOOLS

All codes for data preprocessing, machine learning algorithms, and performance evaluation were written in Python (version 3.6; scikit-learn, version 0.24.1) and R (version 4.0.3).

III. RESULTS

After preprocessing, the annual dataset comprises 38 763 responses, and the integrated dataset comprises 38 763 responses. The detailed numbers of subjects in each dataset are shown in Table III. A log transformation was applied to the revenue feature to convert a skewed distribution into a Gaussian distribution. Figure 3 shows the distribution of the revenue features.

TABLE III
NUMBER OF SUBJECTS IN TRAINING AND TEST DATASETS.

| Dataset | Number of subjects in Training dataset | Number of subjects in Test dataset |
|-----------------------------------|--|------------------------------------|
| Annual dataset ^a (1st) | 7729 | 859 |
| Annual dataset (2nd) | 5311 | 591 |
| Annual dataset (3rd) | 4363 | 485 |
| Annual dataset (4th) | 4032 | 448 |
| Annual dataset (5th) | 4518 | 502 |
| Annual dataset (6th) | 4417 | 491 |
| Annual dataset (7th) | 4554 | 506 |
| Integrated dataset ^b | 34 925 | 3881 |

^aAnnual dataset: sub-dataset collected by year in HCCP dataset.

^bIntegrated dataset: the dataset including features integrated by category from original HCCP dataset.

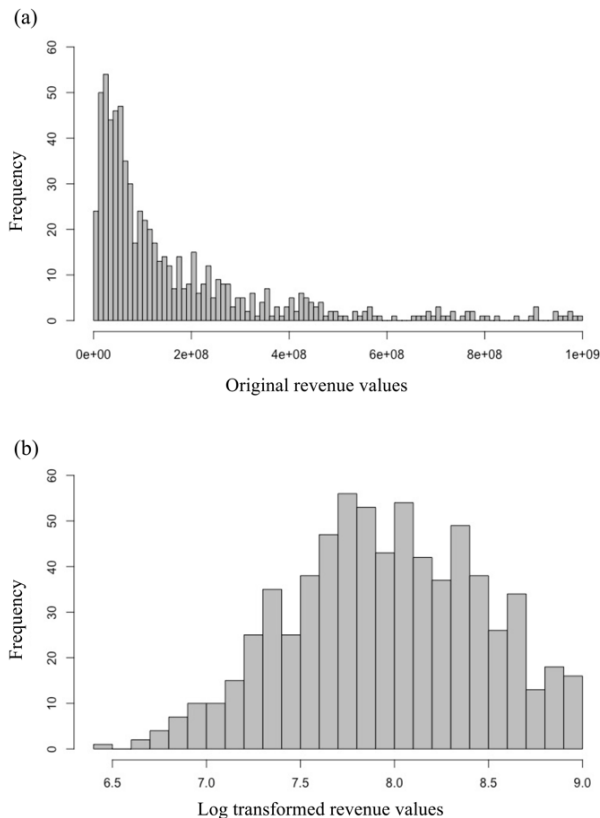


FIGURE 3. Distribution of original and log-transformed revenue features. (a) Distribution of original revenue in HCCP dataset. (b) After applied log transformation, converted distribution of revenue value.

To evaluate the performance of the proposed framework, we applied preprocessed datasets with various class conditions to machine learning algorithms. The hyperparameters for the training algorithms are listed in Table IV.

TABLE IV
HYPERPARAMETERS OF MACHINE LEARNING ALGORITHM.

| Model | Hyperparameter | Value (argument) |
|-----------------------------|------------------|------------------|
| XGBoost classifier | eta | 0.3 |
| | gamma | 0 |
| | max_depth | 6 |
| | min_child_weight | 1 |
| SGD classifier ^a | loss | modified_huber |
| | penalty | L2 |
| | alpha | 0.0001 |
| Logistic regression | penalty | L2 |
| | solver | lbfgs |

^aSGD classifier: Stochastic Gradient Descent classifier

A. BINARY CLASSIFICATION

To evaluate the performance of the model in the binary class condition, we applied an annual dataset and integrated dataset to the XGBoost classifier, SGD classifier, and logistic regression model. We also confirmed the classification performance using several indices. The

experimental results are listed in Table V. The XGBoost classifier showed the best performance among the classification algorithms for the annual and integrated datasets. ROC curves from the integrated dataset are shown in Figure 4. The other ROC curves are presented in Appendix A.

In the binary class condition, the maximum values of the indices were found for the XGBoost classifier for both the annual and integrated datasets. All index values of the trained XGBoost classifier (average precision of 83.2%, average recall of 83.1%, average F1-score of 82.9%, average accuracy of 83.2%, and average AUC of 92.7%) were higher than those of the SGD classifier and logistic regression (Table V).

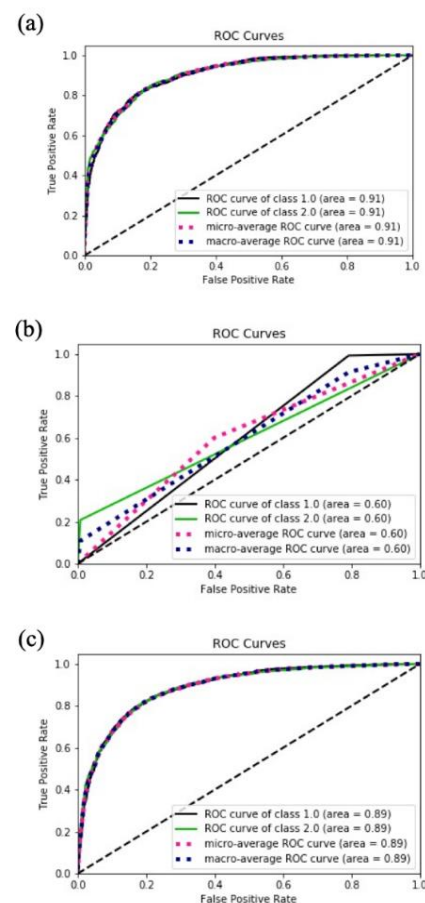


FIGURE 4. ROC curves of classification algorithms from integrated dataset under binary condition: (a) XGBoost classifier, (b) SGD classifier, and (c) logistic regression.

B. MULTICLASS CLASSIFICATION

In addition to the binary class condition, we conducted a multiclass classification using the XGBoost classifier, SGD classifier, and logistic regression model. In addition, we checked the classification performance of each model under various conditions. The experimental results are listed in Tables VI and VII. Under a multiclass condition, the XGBoost classifier showed the best classification performance among the classification algorithms. Figures 5

TABLE V
CLASSIFICATION PERFORMANCE RESULTS FOR MACHINE LEARNING ALGORITHMS UNDER BINARY CLASS CONDITION.

| Dataset | Annual dataset (1st) | | | Annual dataset (2nd) | | | Annual dataset (3rd) | | | Annual dataset (4th) | | |
|------------------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|
| Model | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression |
| Precision | 0.865 | 0.738 | 0.831 | 0.844 | 0.717 | 0.803 | 0.843 | 0.585 | 0.796 | 0.862 | 0.756 | 0.821 |
| Recall | 0.864 | 0.653 | 0.832 | 0.844 | 0.640 | 0.801 | 0.844 | 0.510 | 0.795 | 0.857 | 0.524 | 0.819 |
| F1-score | 0.864 | 0.620 | 0.830 | 0.844 | 0.604 | 0.801 | 0.844 | 0.371 | 0.795 | 0.843 | 0.385 | 0.819 |
| Accuracy | 0.864 | 0.654 | 0.831 | 0.854 | 0.639 | 0.802 | 0.845 | 0.509 | 0.796 | 0.843 | 0.524 | 0.818 |
| AUC ^a | 0.953 | 0.651 | 0.923 | 0.941 | 0.641 | 0.910 | 0.943 | 0.641 | 0.910 | 0.934 | 0.523 | 0.900 |

| Dataset | Annual dataset (5th) | | | Annual dataset (6th) | | | Annual dataset (7th) | | | Integrated dataset | | |
|-----------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|--------------------|----------------|---------------------|
| Model | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression |
| Precision | 0.827 | 0.753 | 0.805 | 0.812 | 0.621 | 0.796 | 0.781 | 0.251 | 0.751 | 0.822 | 0.761 | 0.811 |
| Recall | 0.824 | 0.511 | 0.804 | 0.808 | 0.603 | 0.794 | 0.780 | 0.509 | 0.750 | 0.824 | 0.600 | 0.813 |
| F1-score | 0.823 | 0.359 | 0.804 | 0.808 | 0.587 | 0.793 | 0.782 | 0.389 | 0.751 | 0.823 | 0.527 | 0.811 |
| Accuracy | 0.825 | 0.511 | 0.805 | 0.809 | 0.602 | 0.794 | 0.786 | 0.521 | 0.752 | 0.832 | 0.600 | 0.823 |
| AUC | 0.934 | 0.513 | 0.892 | 0.915 | 0.601 | 0.878 | 0.879 | 0.573 | 0.823 | 0.918 | 0.601 | 0.891 |

^aAUC: Area under the curve

The best classification performance among the three algorithms is emphasized in bold

and 6 depict the ROC curves for the classification algorithms from the integrated dataset with three and four classes, respectively. The results are presented in Appendix 2.

Similarly, under the three class conditions, the XGBoost classifier (average precision of 72.7%, average recall of 71.6%, average F1-score of 71.9%, average accuracy of 71.7%, and average AUC of 89.0%) showed an average value higher than both the SGD classifier and logistic regression (Table VI). Under the four-class condition, the maximum indices were found in the XGBoost classifier (average precision of 63.2%, average recall of 63.4%, average F1-score of 62.6%, average accuracy of 63.4%, and average AUC of 88.0%). XGBoost showed the best

classification performance compared to the SGD classifier and logistic regression under all experimental conditions (Table VII).

C. FEATURE IMPORTANCE OF ALGORITHM

After the experiment on the classification performance, the top-10 most important features from the trained XGBoost classifier were compared for both the annual dataset and the integrated dataset with various class conditions. The scale, monetary compensation, and industry were selected as the top-three most important features from the annual and integrated datasets. Detailed lists of the important features are provided in Tables VIII and IX.

TABLE VI
CLASSIFICATION PERFORMANCE RESULTS FOR MACHINE LEARNING ALGORITHMS UNDER THREE CLASS CONDITIONS.

| Dataset | Annual dataset (1st) | | | Annual dataset (2nd) | | | Annual dataset (3rd) | | | Annual dataset (4th) | | |
|------------------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|
| Model | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression |
| Precision | 0.747 | 0.418 | 0.687 | 0.775 | 0.333 | 0.710 | 0.808 | 0.256 | 0.703 | 0.690 | 0.312 | 0.620 |
| Recall | 0.740 | 0.346 | 0.683 | 0.768 | 0.497 | 0.701 | 0.797 | 0.397 | 0.705 | 0.680 | 0.447 | 0.618 |
| F1-score | 0.743 | 0.232 | 0.687 | 0.770 | 0.398 | 0.711 | 0.797 | 0.299 | 0.704 | 0.684 | 0.353 | 0.617 |
| Accuracy | 0.743 | 0.354 | 0.688 | 0.768 | 0.497 | 0.710 | 0.793 | 0.397 | 0.705 | 0.680 | 0.448 | 0.618 |
| AUC ^a | 0.912 | 0.572 | 0.862 | 0.918 | 0.623 | 0.881 | 0.927 | 0.601 | 0.871 | 0.879 | 0.592 | 0.813 |

| Dataset | Annual dataset (5th) | | | Annual dataset (6th) | | | Annual dataset (7th) | | | Integrated dataset | | |
|-----------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|--------------------|----------------|---------------------|
| Model | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression |
| Precision | 0.743 | 0.312 | 0.621 | 0.728 | 0.376 | 0.631 | 0.603 | 0.332 | 0.587 | 0.721 | 0.366 | 0.634 |
| Recall | 0.740 | 0.447 | 0.627 | 0.709 | 0.373 | 0.629 | 0.591 | 0.445 | 0.582 | 0.706 | 0.548 | 0.636 |
| F1-score | 0.741 | 0.353 | 0.623 | 0.711 | 0.246 | 0.630 | 0.594 | 0.352 | 0.575 | 0.711 | 0.438 | 0.635 |
| Accuracy | 0.743 | 0.448 | 0.627 | 0.708 | 0.374 | 0.629 | 0.591 | 0.445 | 0.584 | 0.706 | 0.548 | 0.636 |
| AUC | 0.893 | 0.602 | 0.820 | 0.884 | 0.534 | 0.813 | 0.824 | 0.582 | 0.783 | 0.883 | 0.53 | 0.828 |

^aAUC: Area under the curve

The best classification performance among the three algorithms is emphasized in bold

TABLE VII
CLASSIFICATION PERFORMANCE RESULTS FOR MACHINE LEARNING ALGORITHMS UNDER FOUR CLASS CONDITIONS.

| Dataset | Annual dataset (1st) | | | Annual dataset (2nd) | | | Annual dataset (3rd) | | | Annual dataset (4th) | | |
|------------------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|
| Model | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression |
| Precision | 0.660 | 0.173 | 0.452 | 0.671 | 0.209 | 0.458 | 0.644 | 0.312 | 0.450 | 0.644 | 0.219 | 0.470 |
| Recall | 0.657 | 0.343 | 0.465 | 0.678 | 0.367 | 0.470 | 0.649 | 0.254 | 0.464 | 0.649 | 0.397 | 0.482 |
| F1-score | 0.653 | 0.229 | 0.455 | 0.670 | 0.253 | 0.462 | 0.637 | 0.108 | 0.448 | 0.637 | 0.271 | 0.467 |
| Accuracy | 0.657 | 0.343 | 0.465 | 0.677 | 0.366 | 0.469 | 0.649 | 0.253 | 0.463 | 0.649 | 0.397 | 0.482 |
| AUC ^a | 0.893 | 0.601 | 0.761 | 0.903 | 0.592 | 0.773 | 0.894 | 0.572 | 0.751 | 0.893 | 0.600 | 0.761 |

| Dataset | Annual dataset (5th) | | | Annual dataset (6th) | | | Annual dataset (7th) | | | Integrated dataset | | |
|-----------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|----------------------|----------------|---------------------|--------------------|----------------|---------------------|
| Model | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression | XGBoost classifier | SGD classifier | Logistic regression |
| Precision | 0.604 | 0.250 | 0.453 | 0.627 | 0.312 | 0.453 | 0.588 | 0.194 | 0.422 | 0.616 | 0.401 | 0.447 |
| Recall | 0.608 | 0.254 | 0.466 | 0.625 | 0.252 | 0.466 | 0.583 | 0.259 | 0.433 | 0.621 | 0.369 | 0.492 |
| F1-score | 0.601 | 0.111 | 0.457 | 0.614 | 0.104 | 0.457 | 0.581 | 0.133 | 0.424 | 0.614 | 0.251 | 0.481 |
| Accuracy | 0.607 | 0.254 | 0.466 | 0.625 | 0.252 | 0.466 | 0.583 | 0.260 | 0.432 | 0.621 | 0.369 | 0.492 |
| AUC | 0.873 | 0.552 | 0.743 | 0.874 | 0.513 | 0.741 | 0.834 | 0.522 | 0.711 | 0.874 | 0.613 | 0.783 |

^aAUC: Area under the curve

The best classification performance among the three algorithms is emphasized in bold

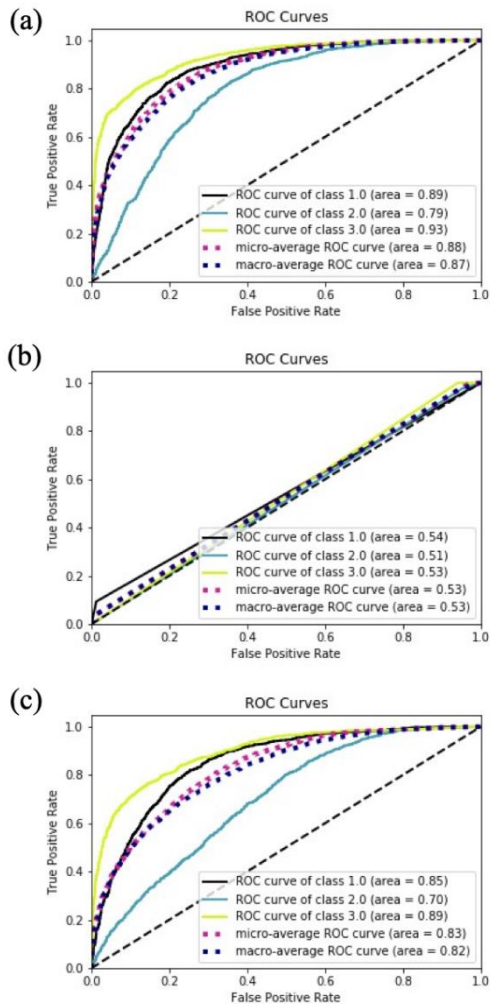


FIGURE 5. ROC curves of classification algorithms from integrated dataset under three class conditions: (a) XGBoost classifier, (b) SGD classifier, and (c) logistic regression

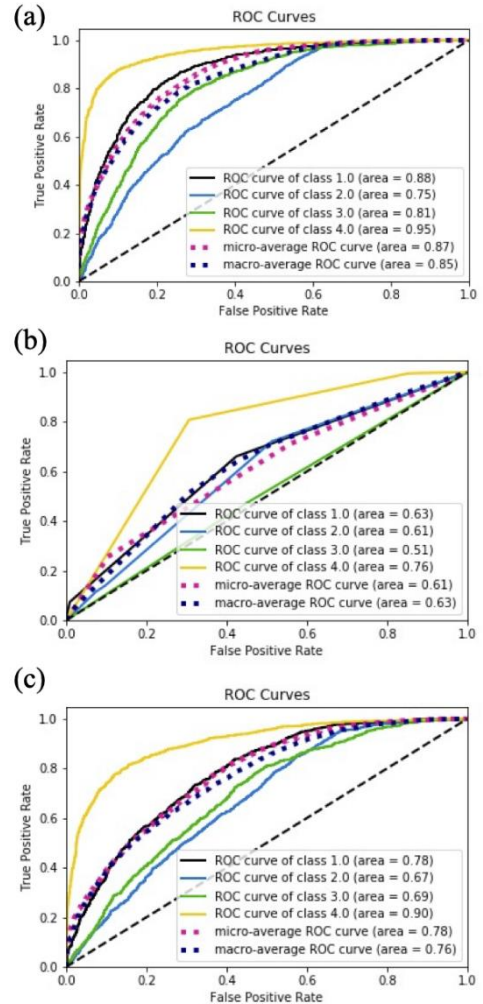


FIGURE 6. ROC curves of classification algorithms from integrated dataset under four class conditions: (a) XGBoost classifier, (b) SGD classifier, and (c) logistic regression

TABLE VIII
LIST OF TOP-10 MOST IMPORTANT FEATURES OF TRAINED XGBOOST CLASSIFIER UNDER EXPERIMENTAL CONDITIONS IN ANNUAL DATASET.

| Classes | Dataset | Top 10 feature importance | | | | | | | | | |
|---------|---------|--------------------------------|--------------------------------|------------------------------------|------------------------------------|------------------------------|-----------------------------|-----------------------------|-----------------------------|---------------------------|-----------------------------|
| 2 | 1st | W137_01 (Monetary) | W1_ind1 (Industry) | Scale (Company Size) | W1_birthy (Demographics) | W1_edu (Demographics) | W134 (Contract Cond.) | W124_08 (Culture) | W1_sex (Demographics) | W124_06 (Motivation) | W124_15 (Satisfaction) |
| | 2nd | W243_01 (Monetary) | Scale (Company Size) | W2_ind1 (Industry) | W2_birthy (Demographics) | W2_birthm (Demographics) | W2_edu (Demographics) | W229_01 (Comm.&Trust) | W228_02 (Talent Mgmt.) | W2_sex (Demographics) | W232_05 (Commitment) |
| | 3rd | W337_01 (Monetary) | Scale (Company Size) | W3_ind1 (Industry) | W3_birthy (Demographics) | W335 (Contract Cond.) | W329_03 (Satisfaction) | W328_02 (Culture) | W327_01 (Comm.&Trust) | W330_04 (Commitment) | W330_07 (Commitment) |
| | 4th | W437_01 (Monetary) | Scale (Company Size) | W4_ind1 (Industry) | W4_birthy (Demographics) | W430_03 (Commitment) | W435 (Contract Cond.) | W428_04 (Culture) | W428_01 (Culture) | W429_02 (Satisfaction) | W430_02 (Commitment) |
| | 5th | W539_01 (Monetary) | Scale (Company Size) | W5_ind1 (Industry) | W529_05 (Comm.&Trust) | W5_edu (Demographics) | W532_03 (Commitment) | W529_06 (Comm.&Trust) | W532_06 (Stress) | W532_07 (Stress) | W528_01 (Talent Mgmt.) |
| | 6th | W636_01 (Monetary) | Scale (Company Size) | W6_birthy (Demographics) | W6_ind1 (Industry) | W634 (Contract Cond.) | W627_02 (Culture) | W6_edu (Demographics) | W629_03 (Commitment) | W629_06 (Stress) | W627_03 (Culture) |
| | 7th | W736_01 (Monetary) | Scale (Company Size) | W7_birthy (Demographics) | W729_03 (Commitment) | W7_ind1 (Industry) | W7_edu (Demographics) | W734 (Contract Cond.) | W729_06 (Stress) | W728_02 (Satisfaction) | W727_01 (Culture) |
| 3 | 1st | W137_01 (Monetary) | Scale (Company Size) | W1_ind1 (Industry) | W1_birthy (Demographics) | W1_edu (Demographics) | W124_01 (Comm.&Trust) | W124_13 (Commitment) | W1_birthm (Demographics) | W124_14 (Satisfaction) | W124_16 (Commitment) |
| | 2nd | W243_01 (Monetary) | Scale (Company Size) | W1_ind1 (Industry) | W2_birthy (Demographics) | W2_edu (Demographics) | W241 (Contract Cond.) | W229_01 (Comm.&Trust) | W232_02 (Motivation) | W229_03 (Comm.&Trust) | W230_03 (Culture) |
| | 3rd | W337_01 (Monetary) | Scale (Company Size) | W3_ind1 (Industry) | W3_birthy (Demographics) | W335 (Contract Cond.) | W328_02 (Culture) | W330_02 (Motivation) | W328_01 (Culture) | W328_03 (Culture) | W3_edu (Demographics) |
| | 4th | W437_01 (Monetary) | Scale (Company Size) | W4_birthy (Demographics) | W4_ind1 (Industry) | W435 (Contract Cond.) | W4_birthm (Demographics) | W427_01 (Comm.&Trust) | W426_01 (Talent Mgmt.) | W428_05 (Culture) | W4_edu (Demographics) |
| | 5th | W539_01 (Monetary) | Scale (Company Size) | W5_birthy (Demographics) | W5_ind1 (Industry) | W537 (Contract Cond.) | W5_edu (Demographics) | W532_03 (Commitment) | W529_03 (Comm.&Trust) | W531_02 (Satisfaction) | W532_07 (Stress) |
| | 6th | W636_01 (Monetary) | Scale (Company Size) | W6_birthy (Demographics) | W6_ind1 (Industry) | W634 (Contract Cond.) | W6_edu (Demographics) | W6_birthm (Demographics) | W627_12 (Culture) | W628_02 (Satisfaction) | W632 (Contract Cond.) |
| | 7th | W736_01 (Monetary) | Scale (Company Size) | W7_birthy (Demographics) | W7_ind1 (Industry) | W7_edu (Demographics) | W7_birthm (Demographics) | W734 (Contract Cond.) | W727_01 (Culture) | W729_03 (Commitment) | W729_06 (Stress) |
| 4 | 1st | Scale (Company Size) | W137_01 (Monetary) | W1_ind1 (Industry) | W1_birthy (Demographics) | W1_edu (Demographics) | W124_01 (Comm.&Trust) | W124_13 (Commitment) | W124_07 (Commitment) | W134 (Contract Cond.) | W1_birthm (Demographics) |
| | 2nd | W243_01 (Monetary) | Scale (Company Size) | W2_ind1 (Industry) | W2_birthy (Demographics) | W2_edu (Demographics) | W232_02 (Motivation) | W232_05 (Commitment) | W231_03 (Satisfaction) | W241 (Contract Cond.) | W230_03 (Culture) |
| | 3rd | W337_01 (Monetary) | Scale (Company Size) | W3_ind1 (Industry) | W3_birthy (Demographics) | W335 (Contract Cond.) | W330_01 (Motivation) | W3_edu (Demographics) | W3_birthm (Demographics) | W327_01 (Comm.&Trust) | W328_02 (Culture) |
| | 4th | W437_01 (Monetary) | Scale (Company Size) | W4_birthy (Demographics) | W4_ind1 (Industry) | W435 (Contract Cond.) | W427_04 (Comm.&Trust) | W426_02 (Talent Mgmt.) | W430_06 (Stress) | W4_edu (Demographics) | W430_05 (Stress) |
| | 5th | W539_01 (Monetary) | Scale (Company Size) | W5_ind1 (Industry) | W5_birthy (Demographics) | W532_07 (Stress) | W5_birthm (Demographics) | W537 (Contract Cond.) | W530_02 (Culture) | W5_edu (Demographics) | W532_01 (Commitment) |
| | 6th | W636_01 (Monetary) | Scale (Company Size) | W6_birthy (Demographics) | W6_ind1 (Industry) | W6_edu (Demographics) | W627_02 (Culture) | W634 (Contract Cond.) | W6_birthm (Demographics) | W629_03 (Commitment) | W625_03 (Talent Mgmt.) |
| | 7th | W736_01 (Monetary) | Scale (Company Size) | W7_birthy (Demographics) | W7_ind1 (Industry) | W727_01 (Culture) | W7_edu (Demographics) | W729_01 (Commitment) | W7_sex (Demographics) | W734 (Contract Cond.) | W725_01 (Talent Mgmt.) |

Features that appear many times are emphasized in bold

TABLE IX
LIST OF TOP-10 MOST IMPORTANT FEATURES IN TRAINED XGBOOST CLASSIFIER UNDER EXPERIMENTAL CONDITIONS FOR INTEGRATED DATASET.

| Classes | Top 10 feature importance | | | | | | | | | |
|---------|-------------------------------|--------------------------------|-----------------------------|-----------------------------------|---------------------------------|-------------------------|----------------------------|---------------------------------|--------------------------------|--------------------------------|
| 2 | Monetary (Monetary) | Scale (Company Size) | W_ind1 (Industry) | W_birthy (Demographics) | Stress (Stress) | W_edu (Demographics) | Motivation (Motivation) | Labor union (Contract Cond.) | Sex (Demographics) | Culture (Culture) |
| 3 | Monetary (Monetary) | Scale (Company Size) | W_ind1 (Industry) | W_birthy (Demographics) | Labor union (Contract Cond.) | W_edu (Demographics) | Stress (Stress) | Culture (Culture) | Comm.&Trust (Comm.&Trust) | Satisfaction (Satisfaction) |
| 4 | Monetary (Monetary) | Scale (Company Size) | W_ind1 (Industry) | W_birthy (Demographics) | Culture (Culture) | Stress (Stress) | W_edu (Demographics) | Labor union (Contract Cond.) | Talent Mgmt. (Talent Mgmt.) | Comm.&Trust (Comm.&Trust) |

Features that appear numerous times are emphasized in bold

IV. DISCUSSION

In this study, we proposed a framework for classifying company revenue using machine learning algorithms. For this purpose, we designed an experimental paradigm including various class conditions and compared the important features of the trained models. To validate the robustness of our framework, we applied three different classification algorithms with multiple datasets and checked the feature importance to confirm the adaptability in the HRM field.

The findings regarding the classification performance and top-10 important features of the trained models are used to evaluate the experimental paradigm, as shown in Tables V–IX.

A. FRAMEWORK PERFORMANCE

To classify the revenue, the experiment was conducted under three conditions. Under the first condition, the binary levels of the revenue were identified using classification algorithms. In other words, a company can be split into two groups based on its revenue. We conducted a multiclass classification using three and four classes under the second and last conditions for comparison with the first condition. All three conditions were cross-validated using both annual and integrated datasets.

The evaluation indices used in this study were the accuracy, precision, recall, F1-score, ROC curve, and AUC. In the comparison of the classification performance, we considered the recall (i.e., the ratio of misclassification) and precision (i.e., the ratio of correct classification). To compare the classification performance, the values of the ROC curve and AUC indices were averaged for every class. As a result, micro-and macro-averaged indices were used for comparison.

We selected three classification algorithms in our research, i.e., the XGBoost classifier, logistic regression, and SGD classifier. Under every condition, the XGBoost classifier showed the best classification performance among the three algorithms. In addition, we found similar performance results in related studies predicting employee turnover based on the XGBoost model (AUC of 87.0%) used in predictions of employee performance based on machine learning algorithms (J48 algorithm: precision of 72.7%, recall of 70.8%, F1-score of 70.8%, and AUC of 85.9%) [36, 37].

B. EMPHASIZED FEATURES

In traditional HRM studies, researchers have investigated a wide range of variables in an organizational environment. In this study, we examined both the company- and employee-level variables and confirmed the top-10 most important features. The result suggests several insights from an academic and practical perspective.

First, monetary reward, scale of the company, industrial categories of the company, and demographics (year of birth) were the most important features for the classification of revenue. All results showed the same trends for both the annual and integrated datasets. In relation to these results, monetary compensation has been considered a primary incentive for employees and a determinant of a company's productivity. Furthermore, empirical evidence in previous research has suggested that companies can utilize monetary compensation as a management tool to enhance the performance in terms of short- and long-term revenue growth [38–40]. In addition, the scale of the company and industrial categories have been regarded as influential components of performance. Researchers interested in the high performance of HR practices and various components of an organization have focused on the size of the organization as an element affecting the financial performance owing to the market power and economics based on scale [41].

Second, four additional features (organizational culture, satisfaction, stress level, and communication and trust) were selected as features that are less important than the first four features. Organizational culture among employees has been a popular topic in studies on how the culture of an organization influences its outcome [42–44]. In addition, the satisfaction and stress levels of employees are considered to have a strong association with the performance of an organization [45–48]. Communication and trust between employees produce several benefits for both employees and an organization [49–51]. In the last aspect of our feature importance, contract conditions and industrial relations (e.g., labor unions) were determined. Industrial relations cover various research areas, such as the industrial relation system, the type of workplace system, and the development of a business strategy within the company [52, 53].

In summary, we confirmed that the features selected from our framework correspond with the points emphasized in previous studies on HRM. This reveals the close relation of employee- and company-level features to the organizational performance.

C. PROPOSED FRAMEWORK FOR HRM PRACTICE

The classification model also suggests some practical implications to HRM practitioners. As a strategic partner, HR managers are often asked to facilitate the entire company-level performance. The purpose of this study is to provide a useful tool for such managers that will help them make better decisions in a complex business environment. Our framework is suitable for this purpose because the survey items used (features) can generally be applied to companies under various situations. HR managers can set benchmarks on our research methodology and plan the measurements of their own

organizational status. In addition, the XGBoost classifier identified the company revenue with an average accuracy of 83.2% under a binary class condition. Considering the complexity of the organizational phenomena, a specific model performance is a useful navigator in estimating the performance of a company in advance.

In terms of HR policy making, the feature importance showed some key elements of consideration. First, monetary compensation was robustly found among the top-10 most important features. Several types of cultural associations between employees and an organization have been studied [54], and the proportion of salary in relation to the total compensation can be determined by the associations. Notably, studies have shown that employees in east Asian countries have formed relational models in comparison to transactional models found in Western countries [55]. However, the feature importance of monetary reward in our research suggests that such a cultural difference may have also transformed during the last 20 years in Korea.

Second, variables from a company-level strategy (e.g., size and industry) are found among the top-10 most important features. Compared to the separative approach during the early stage of HRM research, recent studies emphasized the integration of HR and a corporate strategy. HR practitioners also need to renew their focus from administrative supporters to strategic partners in the development of a business strategy. Finally, organizational culture, satisfaction, stress level, communication, and trust repeatedly appeared among the list of the top-10 most important features. A comprehensive HRM system can be achieved by focusing on specific features rather than relying on intuition.

V. CONCLUSION

The estimation and classification of revenue is critical for planning the business strategy in a company. In this study, we proposed a machine-learning framework for the classification of company revenue. We also designed an experimental paradigm to investigate the performance of our model under various conditions. The proposed framework identified revenue levels with a maximum AUC of 95.3% under the binary-class conditions, 92.7% for the three-class conditions, and 90.3% for the four-class conditions. The results suggest that our framework based on a dataset of employee surveys has the potential to reliably classify company revenue. In addition, more diverse variables from the HRM field (e.g., personality, leadership, and decision making), as well as a different approach with deep learning methods, need to be considered in a future study. We expect our framework to contribute to the practice of HRM.

AUTHOR CONTRIBUTIONS

Conceptualization: Inhwan Ko, Jung-gu Choi, Sanghoon Han
Data curation: Jung-gu Choi, Inhwan Ko
Investigation: Inhwan Ko, Jung-gu Choi
Project administration: Jung-gu Choi, Inhwan Ko, Sanghoon Han
Validation: Jeongjae Kim, Yeseul Jeon, Sanghoon Han
Writing – original draft: Inhwan Ko, Jung-gu Choi

APPENDIX A

The ROC curves from seven annual datasets under binary class conditions

APPENDIX B

The ROC curves from seven annual datasets under three-class conditions

APPENDIX C

The ROC curves from seven annual datasets under four-class conditions

REFERENCES

- [1] P. M. Wright, G. C. McMahan, "Theoretical Perspectives for Strategic Human Resource Management," *Journal of Management*, vol. 18, no. 2, pp. 295–320, Jun. 1992.
- [2] K. A. Golden, V. Ramanujam, "Between a Dream and a Nightmare: On the Integration of the Human Resource Management and Strategic Business Planning Processes," *Human Resource Management*, vol. 24, no. 4, pp. 429–452, Winter, 1985.
- [3] J. Milliman, M. A. von Glinow, M. Nathan, "Organizational Life Cycles and Strategic International Human Resource Management in Multinational Companies: Implications for Congruence Theory," *Academy of Management Review*, vol. 16, no. 2, pp. 318–339, Apr. 1991.
- [4] P. M. Wright, B. B. Dunford, S. A. Snell, "Human resources and the resource based view of the firm," *Journal of Management*, vol. 27, no. 6, pp. 701–721, Nov. 2001.
- [5] B. Becker, B. Gerhart, "The Impact of Human Resource Management on Organizational Performance: Progress and Prospects," *Academy of Management Journal*, vol. 39, no. 4, pp. 779–801, Aug. 1996.
- [6] D. E. Guest, "Human resource management and performance: a review and research agenda," *International Journal of Human Resource Management*, vol. 8, no. 3, pp. 263–276, Jun. 1997.
- [7] J. Pfeffer, "Seven Practices of Successful Organizations," *California Management Review*, vol. 40, no. 2, pp. 96–124, Winter, 1998.
- [8] J. E. Delery, D. H. Doty, "Modes of Theorizing in Strategic Human Resource Management: Tests of Universalistic, Contingency, and Configurational Performance Predictions," *Academy of Management Journal*, vol. 39, no. 4, pp. 802–835, Aug. 1996.
- [9] J. B. Arthur, "Effects of Human Resource Systems on Manufacturing Performance and Turnover," *Academy of Management Journal*, vol. 37, no. 3, pp. 670–687, Jun. 1994.
- [10] M. A. Youndt, S. A. Snell, J. W. Dean, Jr., D. P. Lepak, "Human Resource Management, Manufacturing Strategy, and Firm Performance," *Academy of Management Journal*, vol. 39, no. 4, pp. 836–866, Aug. 1996.
- [11] J. P. MacDuffie, "Human Resource Bundles and Manufacturing Performance: Organizational Logic and Flexible Production Systems in the World Auto Industry," *Industrial and Labor Relations Review*, vol. 48, no. 2, pp. 197–221, Jan. 1995.

- [12] M. A. Huselid, "The Impact of Human Resource Management Practices on Turnover, Productivity, and Corporate Financial Performance," *Academy of Management Journal*, vol. 38, no. 3, pp. 635–672, Jun. 1995.
- [13] R. A. Guzzo, K. A. Noonan, "Human Resource Practices as Communications and the Psychological Contract," *Human Resource Management*, vol. 33, no. 3, pp. 447–462, Fall, 1994.
- [14] D. N. den Hartog, P. Boselie, J. Paauwe, "Performance Management: A Model and Research Agenda," *Applied Psychology*, vol. 53, no. 4, pp. 556–569, Oct. 2004. DOI: <https://doi.org/10.1111/j.1464-0597.2004.00188.x>.
- [15] J. Garcia-Arroyo, A. Osca, "Big data contributions to human resource management: a systematic review," *The International Journal of Human Resource Management*, 2019. DOI: 10.1080/09585192.2019.1674357.
- [16] A. Braganza, L. Brooks, D. Nepelski, M. Ali, R. Moro, "Resource management in big data initiatives: Processes and dynamic capabilities," *Journal of Business Research*, vol. 70, pp. 328–337, Jan. 2017.
- [17] T. L. Case, A. Gardiner, P. Rutner, J. N. Dyer, "A LinkedIn Analysis of Career Paths of Information Systems Alumni," *Journal of the Southern Association for Information System*, 2012. DOI: 10.3998/jsais.11880084.0001.102.
- [18] F. J. García-Peñalvo, J. Cruz-Benito, M. Martín-González, A. Vázquez-Ingelmo, J. C. Sánchez-Prieto, R. Therón, "Proposing a machine learning approach to analyze and predict employment and its factors," *IJIMAI*, vol. 5, no. 2, pp. 39–45, 2018.
- [19] M. H. Rafiei, H. Adeli, "Novel machine-learning model for estimating construction costs considering economic variables and indexes," *Journal of construction engineering and management*, 2018, DOI: 10.1061/(ASCE)CO.1943-7862.0001570.
- [20] B. Park, J. K. Bae, "Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data," *Expert systems with applications*, 2015.
- [21] G. Kou, X. Chao, Y. Peng, F. E. Alsaadi, E. Herrera-Viedma, "Machine learning methods for systemic risk analysis in financial sectors," *Technological and Economic Development of Economy*, vol. 25, no. 5, pp. 716–742, 2019.
- [22] X. Ma, S. Lv, "Financial credit risk prediction in internet finance driven by machine learning," *Neural Computing and Applications*, vol. 31, pp. 8359–8367, 2019.
- [23] J. Patel, S. Shah, P. Thakkar, K. Kotecha, "Predicting stock market index using fusion of machine learning techniques," *Expert Systems with Applications*, vol. 42, no. 4, pp. 2162–2172, 2015.
- [24] J. Patel, S. Shah, P. Thakkar, K. Kotecha, "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques," *Expert systems with applications*, vol. 42, no. 1, pp. 259–268, 2015.
- [25] Y. Zhao, M. K. Hryniewicki, F. Cheng, B. Fu, X. Zhu, "Employee turnover prediction with machine learning: A reliable approach," In *Proceedings of SAI intelligent systems conference*, vol. 869, pp. 737–758, 2018.
- [26] E. Faliagka, K. Ramantas, A. Tsakalidis, G. Tzimas, "Application of machine learning algorithms to an online recruitment system," In *Proc. International Conference on Internet and Web Applications and Services*, 2012, pp. 215–220.
- [27] S. Garg, S. Sinha, A. K. Kar, M. Mani, "A review of machine learning applications in human resource management," *International Journal of Productivity and Performance Management*, 2021, DOI 10.1108/IJPPM-08-2020-0427.
- [28] S. Fan, X. Li, J. L. Zhao, "Collaboration process pattern approach to improving teamwork performance: a data mining-based methodology," *INFORMS Journal on Computing*, Vol. 29, No. 3, pp. 438–456, 2017.
- [29] Human Capital Corporate Panel (HCCP). Korea Research Institute for Vocational Education and Training. 2017. URL: <https://www.krivnet.re.kr/ku/ha/kuCCFDs.jsp> [accessed 2020-11-06]
- [30] T. Chen, C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.
- [31] T. Zhang, "Solving large scale linear prediction problems using stochastic gradient descent algorithms," in *Proceedings of the twenty-first international conference on Machine learning*, 2004, pp. 116.
- [32] W. A. Gardner, "Learning characteristics of stochastic-gradient-descent algorithms: A general study, analysis, and critique," *Signal processing*, vol. 6, no. 2, pp. 113–133, 1984.
- [33] C. M. BISHOP, "Pattern recognition and machine learning," springer, 2006, pp.205–206.
- [34] D. Böhning, "Multinomial logistic regression algorithm," *Annals of the institute of Statistical Mathematics*, vol. 44, no. 1, pp. 197–200, 1992.
- [35] C. Yin, Y. Zhu, J. Fei, X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, pp. 21954–21961, 2017.
- [36] J. M. Kirimi, C. A. Moturi, "Application of data mining classification in employee performance prediction," *International Journal of Computer Applications*, vol. 146, no. 7, 2016.
- [37] P. Ajit, "Prediction of employee turnover in organizations using machine learning algorithms," *algorithms*, 2016.
- [38] G. P. Latham, E. A. Locke, "Goal Setting—A Motivational Technique That Works," *Organizational Dynamics*, vol. 8, no. 2, pp. 68–80, Autumn, 1979.
- [39] E. P. Lazear, "Performance pay and productivity," *American Economic Review*, vol. 90, no. 5, pp. 1346–1361, Dec. 2000.
- [40] H. S. Kim, S. Jang, "The effect of increasing employee compensation on firm performance: Evidence from the restaurant industry," *International Journal of Hospitality Management*, vol. 88, article number 102513, Jul. 2020.
- [41] L. Y. Sun, A. Samuel, S. L. Kenneth, "High-performance human resource practices, citizenship behavior, and organizational performance: A relational perspective," *Academy of management Journal*, vol. 50, no. 3, 2007.
- [42] E. H. Schein, "The Three Levels of Culture," in *Organizational Culture and Leadership*, 4th ed. San Francisco, California, Jossey-Bass Publishers, 2010, pp. 7–196.
- [43] S. J. Hogan, L. V. Coote, "Organizational culture, innovation, and performance: A test of Schein's model," *Journal of Business Research*, vol. 67, no. 8, pp. 1609–1621, Aug. 2014.
- [44] G. A. Marcoulides, R. H. Heck, "Organizational Culture and Performance: Proposing and Testing a Model," *Organization Science*, vol. 4, no. 2, pp. 209–225, May, 1993.
- [45] T. A. Judge, C. J. Thoresen, J. E. Bono, G. K. Patton, "The Job Satisfaction-Job Performance Relationship: A Qualitative and Quantitative Review," *Psychological Bulletin*, vol. 127, no. 3, pp. 376–407, May, 2001.
- [46] J. P. Meyer, D. J. Stanley, L. Herscovitch, L. Topolnytsky, "Affective, Continuance, and Normative Commitment to the Organization: A Meta-analysis of Antecedents, Correlates, and Consequences," *Journal of Vocational Behavior*, vol. 61, no. 1, pp. 20–52, Aug. 2002.
- [47] J. Cook, T. Wall, "New work attitude measures of trust, organizational commitment and personal need non-fulfilment," *Journal of Occupational Psychology*, vol. 53, no. 1, pp. 39–52, Mar. 1980.
- [48] J. R. Hackman, G. Oldham, R. Janson, K. Purdy, "A New Strategy for Job Enrichment," *California Management Review*, vol. 17, no. 4, pp. 57–71, Summer, 1975.
- [49] L. Rhoades, R. Eisenberger, "Perceived organizational support: A review of the literature," *Journal of Applied Psychology*, vol. 87, no. 4, pp. 698–714, Aug. 2002.
- [50] R. M. Kramer, "Trust and distrust in organizations: Emerging perspectives, enduring questions," *Annual Review of Psychology*, vol. 50, pp. 569–598, 1999.
- [51] A. Zaheer, B. McEvily, V. Perrone, "Does Trust Matter? Exploring the Effects of Interorganizational and Interpersonal Trust on Performance," *Organization Science*, vol. 9, no. 2, pp. 141–159, 1998.
- [52] J. B. Arthur, "The Link between Business Strategy and Industrial Relations Systems in American Steel Minimills," *Industrial and Labor Relations Review*, vol. 45, no. 3, pp. 488–506, Apr. 1992.

- [53] C. Doucouliagos, P. Laroche, "What do unions do to productivity? A meta-analysis," *Industrial Relations*, vol. 42, no. 4, pp. 650–691, Oct. 2003.
- [54] D. M. Rousseau, "Contracting: A Modern Dilemma," in *Psychological contracts in organizations: Understanding written and unwritten agreements*, Thousand Oaks, California, SAGE Publications, Inc., 1995, pp. 6–21.
- [55] W. G. Ouchi, "The Z Organization," in *Theory Z: How American Business Can Meet The Japanese Challenge*, Addison-Wesley, 1981, pp. 71–94.



SANGHOON HAN was born in 1977. He is currently a professor in the Department of Psychology, Yonsei University. His research areas are decision making and cognitive science.



JUNG-GU CHOI was born in 1993. He is currently pursuing a PhD degree in cognitive science through the Applied Brain Cognitive Laboratory of the Yonsei University graduate program. His research areas are brain science and deep reinforcement learning.



INHWANKO is in an MS course in psychology at Yonsei University. He has diverse working experience, from a human resource management practitioner in multi-national companies to a certified public labor attorney. His research focuses on converging human resource management and cognitive neuroscience.



JEONGJAE KIM is currently an MS student in cognitive science at Yonsei University. His research interest is in neuroscience, machine learning, and data mining.



YESEUL JEON was born in 1992. She is currently pursuing a PhD degree in applied statistics at Yonsei University. Her research areas are spatial-temporal data analysis and network modeling.

APPENDIX A. ROC curves from seven annual datasets under binary-class conditions

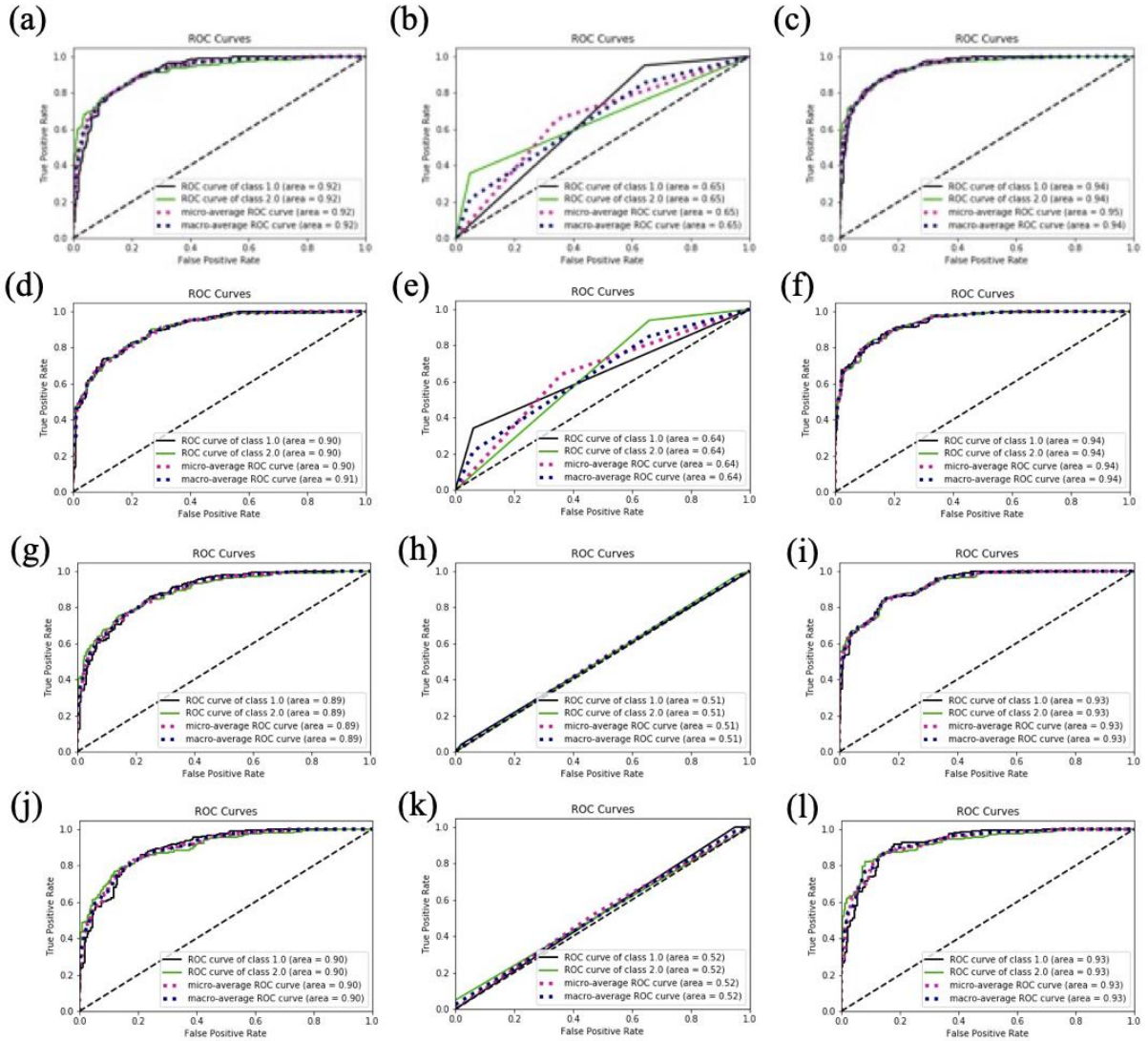


FIGURE A1. ROC curves of classification algorithms from seven annual datasets under binary class conditions. (a), (d), (g), (j) logistic regression, (b), (e), (h), (k) SGD classifier, (c), (f), (i), (l) XGBoost classifier. (a), (b), (c) 1st year dataset, (d), (e), (f) 2nd year dataset, (g), (h), (i) 3rd year dataset, and (j), (k), (l) 4th year dataset

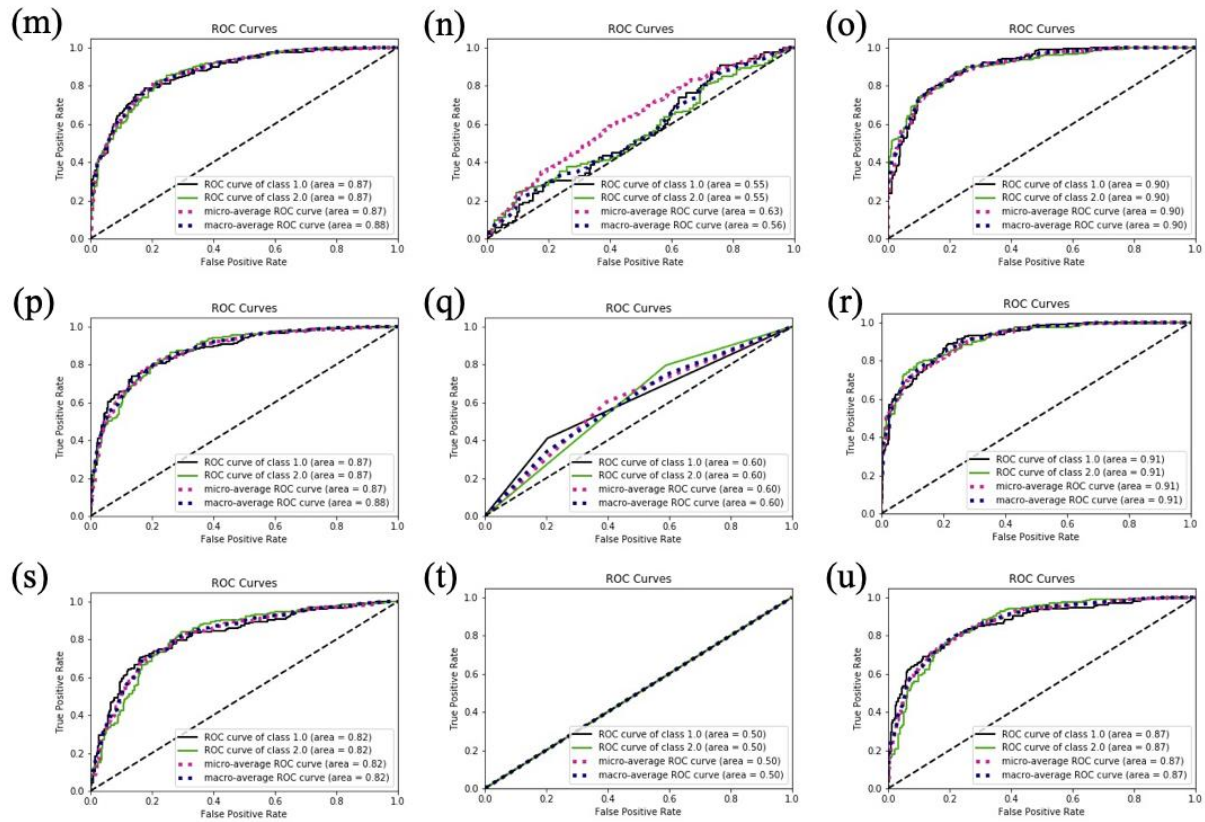


FIGURE A2. ROC curves of classification algorithms from seven annual datasets under binary class conditions. (m), (p), (s) logistic regression, (n), (q), (t) SGD classifier, (o), (r), (u) XGBoost classifier. (m), (n), (o) 5th year dataset, (p), (q), (r) 6th year dataset, and (s), (t), (u) 7th year dataset

APPENDIX B. ROC curves from seven annual datasets under three-class conditions

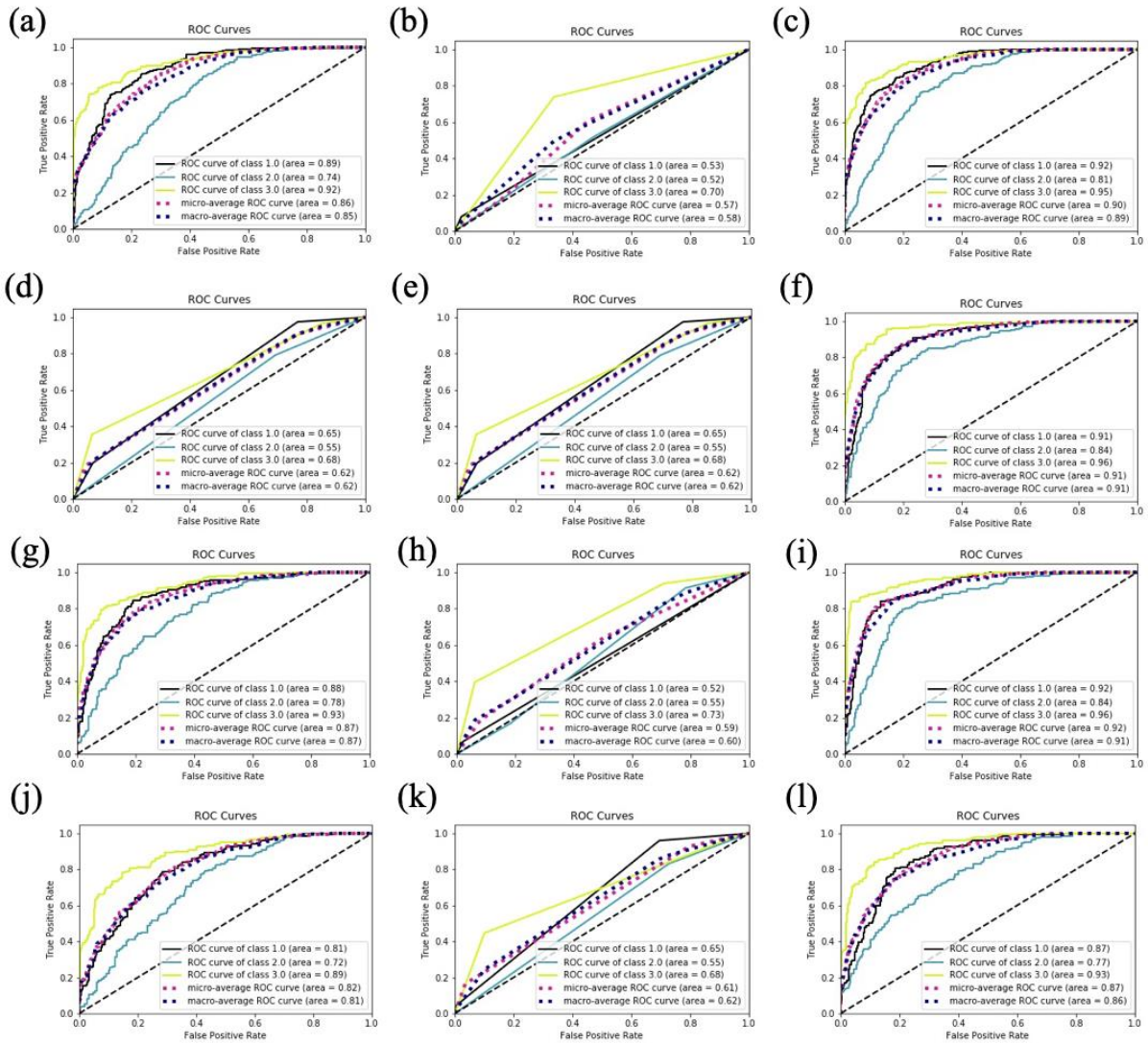


FIGURE B1. ROC curves of classification algorithms from seven annual datasets under three class conditions. (a), (d), (g), (j) logistic regression, (b), (e), (h), (k) SGD classifier, (c), (f), (i), (l) XGBoost classifier. (a), (b), (c) 1st year dataset, (d), (e), (f) 2nd year dataset, (g), (h), (i) 3rd year dataset, and (j), (k), (l) 4th year dataset

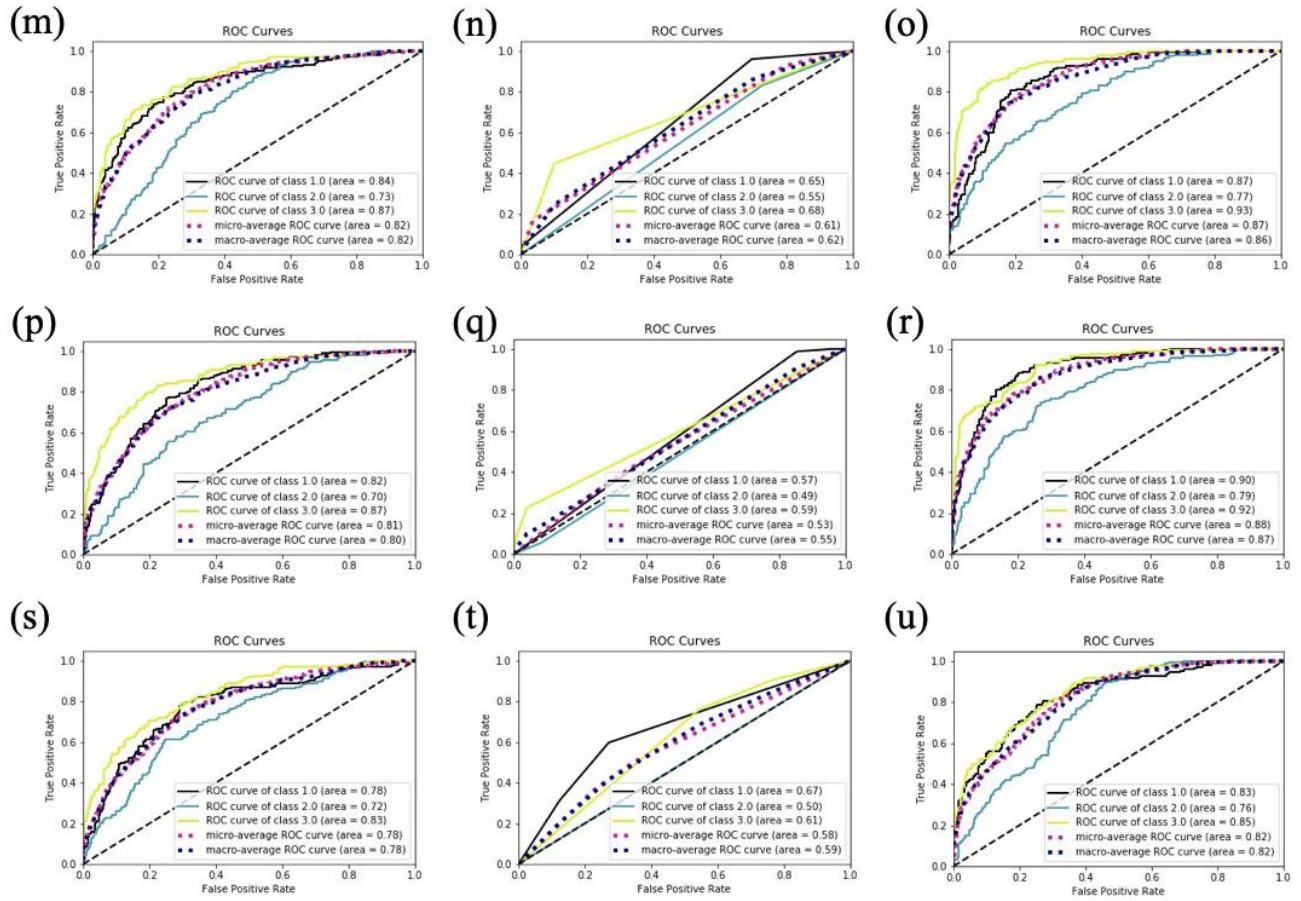


FIGURE B2. ROC curves of classification algorithms from seven annual datasets under three class conditions. (m), (p), (s) logistic regression, (n), (q), (t) SGD classifier, (o), (r), (u) XGBoost classifier. (m), (n), (o) 5th year dataset, (p), (q), (r) 6th year dataset, and (s), (t), (u) 7th year dataset

APPENDIX C. ROC curves from seven annual datasets under four-class conditions

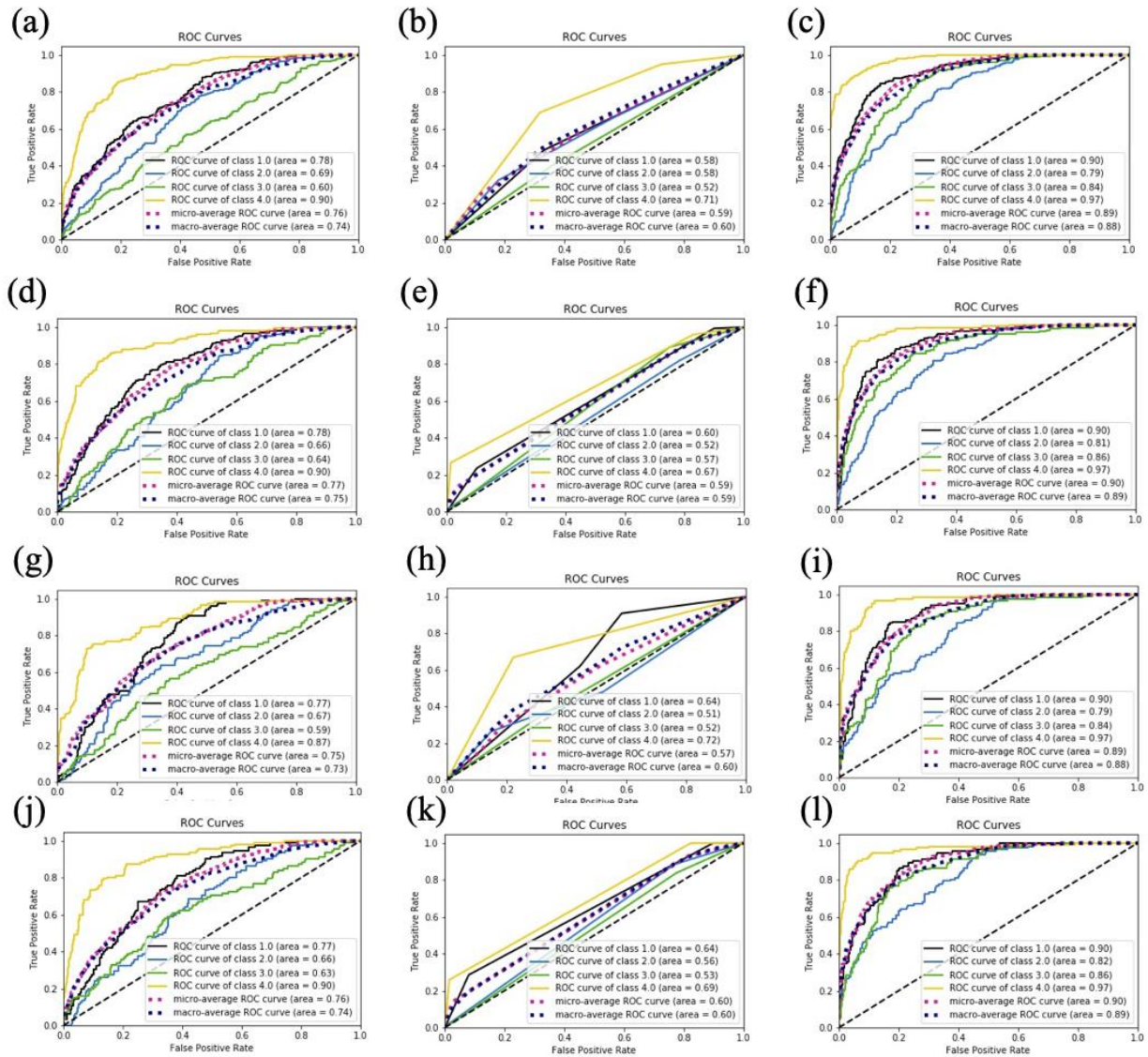


FIGURE C1. ROC curves of classification algorithms from seven annual dataset in four class conditions. (a), (d), (g), (j) logistic regression, (b), (e), (h), (k) SGD classifier, (c), (f), (i), (l) XGBoost classifier. (a), (b), (c) 1st year dataset, (d), (e), (f) 2nd year dataset, (g), (h), (i) 3rd year dataset, and (j), (k), (l) 4th year dataset

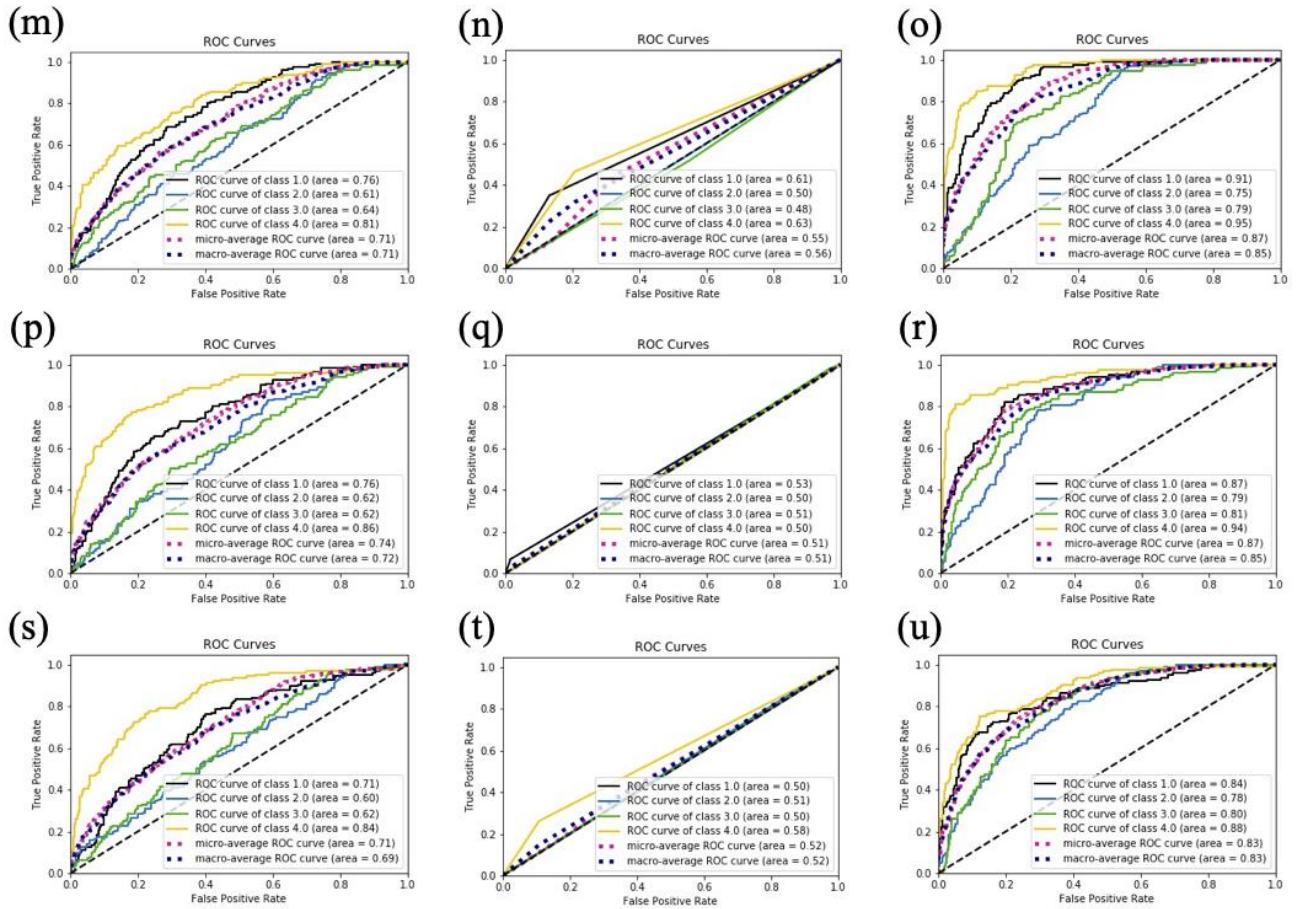


FIGURE C2. ROC curves of classification algorithms from seven annual datasets under four class conditions. (m), (p), (s) logistic regression, (n), (q), (t) SGD classifier, (o), (r), (u) XGBoost classifier. (m), (n), (o) 5th year dataset, (p), (q), (r) 6th year dataset, and (s), (t), (u) 7th year dataset