pISSN : 2508-3384 eISSN : 2508-3392
https://doi.org/10.22640/lxsiri.2019.49.1.65
Journal of Cadastre & Land InformatiX Vol.49 No.1 (2019)
한국국토정보공사「지적과 국토정보」제49권 제1호 2019년 6월 p. 65-84

Prediction of Citizens' Emotions on Home Mortgage Rates Using Machine Learning Algorithms

기계학습 알고리즘을 이용한 주택 모기지 금리에 대한 시민들의 감정예측

김윤기*
Kim, Yun-Ki

Abstract

This study attempted to predict citizens' emotions regarding mortgage rates using machine learning algorithms. To accomplish the research purpose, I reviewed the related literature and then set up two research questions. To find the answers to the research questions, I classified emotions according to Akman's classification and then predicted citizens' emotions on mortgage rates using six machine learning algorithms. The results showed that AdaBoost was the best classifier in all evaluation categories. However, the performance level of Naive Bayes was found to be lower than those of other classifiers. Also, this study conducted a ROC analysis to identify which classifier predicts each emotion category well. The results demonstrated that AdaBoost was the best predictor of the residents' emotions on home mortgage rates in all emotion categories. However, in the sadness class, the performance levels of the six algorithms used in this study were much lower than those in the other emotion categories.

Keywords: Machine Learning, Algorithm, Mortgage Rates, Akman, Emotion Classification, AdaBoost, Classifier, Naive Bayes, Performance Level

1. Introduction

Human beings usually feel various emotions after big events or accidents. When people around them are promoted or married, they express their positive emotions through different types of media. However, when their loved ones die, or their sons don't pass the college entrance exams, people often feel negative emotions such as sadness and surprise. As we enter the information age, we are free to express our feelings about specific political issues and events through social media as well as traditional media. In particular, Twitter is an excellent medium for users to easily access anytime, anywhere and express their views on specific issues (Neethu et al. 2013). Today,

* 청주대학교 지적학과 교수 Professor, Department of Land Management, Choengju University (kim2875@cju.ac.kr)

discussions on various social issues are being actively conducted on Twitter. In the course of these discussions, Twitter users can express their emotions about specific topics without any restriction (Go et al. 2009). Citizens, for example, can express their views on issues that are very closely related to their lives such as housing rents, home mortgage rates, public transportation fees, parking problems, and smog. Also, citizens may express their emotions on the tweets that other people post.

As such, certain events or accidents can have a direct impact on citizens' emotional changes. It is therefore imperative for governments to identify the emotions of citizens on specific policies before formulating and implementing them. In the early 1970s, some psychologists identified six basic emotions that humans experience in every culture. They selected a tribe of New Guinea, which had been completely isolated from the outside, as experimental groups and analyzed how they felt (Ekman et al. 1971). Research showed that primitive tribes also experienced anger, disgust, fear, joy (happiness), sadness, and surprise, as do humans in other cultures. Since then, many researchers have analyzed human emotional changes on specific issues based on the types of emotions Ekman et al. (1971) classified.

Recently, with the help of machine learning algorithms, scholars have been able to analyze human emotions more scientifically and systematically (Wang et al. 2012). In particular, researchers are using multiple machine learning algorithms simultaneously to predict and classify emotions that people feel on a specific issue, rather than using only one algorithm. Through the evaluation of performance metrics such as accuracy and recall, these scholars have tried to identify the best machine learning algorithm for classifying and predicting emotions (Colneriĉ et al. 2018). Although machine learning algorithms have been used to predict and classify citizen emotions about specific issues in a variety of fields, few studies have ever used machine learning algorithms to analyze citizens' emotions on home mortgage rates. Home mortgage rate fluctuations generally can have a significant influence on the housing market (Gerardi et al. 2007; Cannuscio et al. 2012). If the mortgage rate rises, citizens will reduce their investment in housing, which ultimately undermines the housing market itself. On the other hand, if the mortgage rate decreases, it can stimulate the investment sentiment of the citizens and eventually increase the demand for housing. Therefore, to accurately predict the future housing market, it is necessary for policymakers and researchers to predict citizen's emotions on mortgage rates accurately. Thus, this study attempts to predict citizen's emotions on mortgage rates using machine learning algorithms.

2. Literature Review

2.1. Emotion Classification

Emotions refer to a mental state that is variously related to thoughts, feelings, behavioral responses, and preferences (WIKIPEDIA 2019). Humans express various emotions when an event occurs, or an unexpected accident happens. But it is not easy for us to correctly classify the type of emotion because it is the mental state that exists within our inner world. Nonetheless, many scholars have analyzed how people express emotions when faced with certain situations. Ekman et al. (1971) conducted experiments on the indigenous people in New Guinea to verify that facial expressions of emotion are universal. In this study, the researchers divided human emotions into six categories: anger, disgust, fear, joy, sadness, and surprise, and then told a story to the natives and then examined how their facial expressions changed. Their findings support the hypothesis that facial expressions of emotion are universal (Ekman et al. 1971).

On the other hand, Plutchick (1980) classifies human emotions expressed in adaptive biological processes into eight categories: acceptance, anger, anticipation, disgust, joy, fear, sadness, and surprise. We can confirm that six of emotion types classified by Plutchick (1980) are identical to those of emotion classified by Ekman et al. (1971). Plutchick (1980) used the wheel of emotions to explain human emotions. Frijda (1986) classified human emotions based on forms of readiness. To explain the emotional behavior of human beings, he classified the emotions into six types: desire, happiness, interest, surprise, wonder, and sorrow. When we compare the types of emotions classified by Frijda with those proposed by Ekman et al. (1970), we can see that there is a clear difference between them. Among the six emotional types presented by Frijda (1986), happiness, surprise, wonder, and sorrow are similar to the emotion types proposed by Ekman etal., but desire and interest are significantly different from those presented by Ekman et al.

The rapid development of information and communication technology over the past two decades has had a profound effect on patterns of emotion research. In the past, scholars attempted to distinguish human emotion types by observing human facial expressions or gestural patterns. Today, however, researchers are trying to classify people's emotions through analysis of speech manuscripts or articles posted on social media such as Twitter, Facebook, YouTube. Especially Casale et al. (2008) used two speech corpora, written in German and English, to analyze human emotions included in the speech. They classified the emotions in the speech into seven categories: anger, boredom, anxiety, happiness, sadness, disgust, and neutral. We confirmed that the types of emotions Casale et al. (2008) proposed are very similar to those of emotions Ekman et al. (1971) classified. People are expressing their feelings quite honestly in the form of comments on videos they watched on YouTube. The emotion analysis of these YouTube users can be a great help in identifying the overall trends of YouTube. Chen et al. (2017) categorized the emotions that YouTube users feel while watching videos into six types: happiness, anger, disgust, fear, sadness, and surprise. The types of emotions that these researchers used to classify the emotions of YouTube users are almost identical to those proposed by Ekman et al. (1971).

Today, many Twitter users post tweets about diverse social issues. These tweets are likely to contain honest feelings about specific issues of Twitter users themselves. For this reason, many researchers are trying to analyze the emotions of Twitter users on particular topics through tweet analysis. In particular, Roberts et al. (2012) used the six emotion types (anger, disgust, fear, joy, sadness, and surprise) proposed by Ekman et al. (1971) and love to analyze the emotions of Twitter users. And Colneriĉ et al. (2018) used both the six emotion types proposed by Ekman et al. (1971) and the eight emotion types suggested by Plutchick (1980) to analyze the emotions contained in tweets. The researchers collected the data using the hashtags that included Ekman's emotional type as search terms and analyzed it. The researchers also collected data using eight hashtags including Plutchick's emotion types and then examined the emotion distribution of Twitter users.

2.2. Machine Learning Algorithms

Machine learning technology has been used by many researchers to classify and predict phenomena in various fields. Depending on their situation, scholars use supervised machine learning algorithms or non-supervised learning algorithms to solve their problems. Recently, researchers have used different supervised machine learning algorithms to classify and predict emotions contained in speeches and tweets. The supervised machine learning classifiers that researchers use to classify emotions are as follows.

2.2.1. AdaBoost

Boosting means an ensemble technique that creates a strong classifier from a few weak classifiers (Hastie et al. 2009). In general, boosting is done through complicated steps. The first step in boosting is to build the first model from the training data. In the second stage of boosting, the second model is constructed to correct the error of the first model (Hu et al. 2008). The model addition process continues until the training data error is minimized, that is until the training data is completely predicted (Wei et al. 2004). AdaBoost is a boosting algorithm optimized for binary classification. Today, many researchers are using the AdaBoost algorithm to improve the performance of other algorithms in binary classification situations. It has been found that the processing speed of AdaBoost significantly improves if simple weak classifiers such as Decision Tree or Decision Stump are used (Hu et al. 2008).

2.2.2. Decision Tree

The Decision Tree algorithm is a kind of supervised learning algorithm. Unlike other supervised learning algorithms, the Decision Tree algorithm can be used to predict and classify a phenomenon (Du et al. 2002). The Decision Tree algorithm attempts to solve the problem using a decision tree that is shaped like an upside-down tree. In the decision tree, each internal node represents an attribute, and each leaf node represents a class label. At the top of the decision tree is the root node, and at the bottom are the terminal nodes (Safavian et al. 1991). The root node is the whole training data set from which the decision tree algorithm begins. As we go down from the root node, the number of data belonging to each node decreases. Finally, when we reach the terminal nodes, the number of data is reduced to one. In the

Decision Tree, it is important to select the attributes that will be used for analysis. Information gain and Gini index are often used to select attributes in the model.

2.2.3. Random Forest

Random Forest is based on the decision trees. Just as the forest consists of many trees, the Random Forest is made up of numerous decision trees (Pal 2005). The most crucial task in research using the Random Forest algorithm is to choose rational decision tree attributes and pruning methods. In general, many scholars use the Information Gain Ratio Criterion and Gini Index to select attributes of the decision tree. Another challenge faced by researchers is identifying how decisions are made in these complex decision trees. The decision-making process in Random Forest is not much different from the decision-making process in the real world. That is, the majority rule applies to the decision making process in the Random Forest (Rodriguez-Galiano et al. 2012). Suppose we have a Random Forest with 100 decision trees. If 65 of the decision trees have indicated a favorable opinion on a matter, the Random Forest algorithm will make a decision based on the majority rule in favor of the issue.

2.2.4. Logistic Regression

Logistic regression algorithm is a widely used algorithm in machine learning today. Logistic regression is so named because it uses logistic functions as the basis of analysis. The logistic function, also called the sigmoid function, was developed by statisticians to explain population growth in ecology. The logistic function takes an s-shape and has a value between 0 and 1 (Ng et al. 2002). The logistic regression algorithm imposes weaker constraints than general linear regression analysis. In other words, the logistic regression classifier does not assume that there is a linear relationship between the independent variables and the dependent variable, unlike the general linear regression analysis (Subasi et al. 2005). The logistic regression algorithm calculates the changes in the logarithm of odds of the dependent variable rather than the changes of the dependent variable itself, as in the linear regression analysis. In logistic regression analysis, the relationship between independent variables and dependent variables is not linear, as the logarithm of odds and the independent variables have a linear relationship (Alkan et al. 2005).

2.2.5. KNN

The KNN algorithm means a K-Nearest Neighbor algorithm. The KNN algorithm is a method of predicting or classifying new data by using information of k neighboring data closest to existing data when new data is given (Tan 2006). KNN generally does not have a procedure called training. In other words, KNN is also called Lazy model because it does not construct a separate model for analysis. The results of KNN vary greatly depending on the distance measurement method used. To measure the distance to the nearest neighbor, KNN uses Euclidean Distance, Manhatten Distance, Mahalanobis Distance, Correlation Distance, and Rank Correlation Distance (Peterson 2009). Today, researchers are using

Euclidean Distance to predict and classify social phenomena. The KNN algorithm is known to be a very effective algorithm when the data size is large.

2.2.6. Naive Bayes

Today, researchers are using the Naive Bayes algorithm to classify various phenomena in their fields (Sitthi et al. 2016). The Naive Bayes algorithm is based on the Bayesian theorem and is very effective when the dimension of the input is high (Vatsavai et al. 2011). Although this algorithm is simple in structure, it is faster and more efficient than other more sophisticated machine learning algorithms (Dong et al. 2014). The Naive Bayes algorithm can train data more effectively under supervised environments. The Naive Bayes algorithm can also estimate parameters more accurately with a small amount of training data.

2.3. Emotion Prediction Using Machine Learning Algorithms

Today, many researchers are using a variety of machine learning algorithms to classify and predict emotions in postings on social media. First, Casale et al. (2008) extracted the emotions of speakers included in the speech using two speech corpus: Berin Database of Emotional Speech (EMO-DB) and Speech Under Simulated and Actual Stress (SUSAS). Their findings reveal that the Support Vector Machine (SVM) algorithm is the most effective algorithm for predicting the emotions of speakers. Yadav et al. (2015) also classifies the emotions of speakers in the speech using SMO, SVM, Decision Tree, and KNN algorithms, and then compares the performance of each classifier. The results of the analysis show that SVM and SMO are the most powerful algorithms for classifying and predicting emotions in speech.

Quite a few researchers attempted to classify and predict emotions about specific issues based on tweets posted by citizens. First, Wang et al. (2012) used two classifiers, the LIBLINEAR algorithm, and the Multinomial Naive Bayes (MNB) algorithm, to identify the emotions of users in the tweets. The results of the study indicate that classifiers perform better when combined with unigram and bigram. And Sidrov et al. (2012) used the Spanish Emotion Lexicon to extract the emotions contained in tweets. The researchers investigated how the n-gram size and corpus size affect the precision of machine learning algorithms. Their findings reveal that the predictive power of the model is the most excellent when they use unigram, at least 3000 tweets, and the SVM algorithm simultaneously. Bravo-Marquez et al. (2013) also attempted to improve emotion classification of Twitter by using different emotional dimensions as meta-level features. The researchers used the Stanford Twitter Sentiment (STS) dataset and the Sanders dataset for the experiment. Each of these datasets contains positive, negative, and neutral tags. The researchers examined the performance of multiple machine learning algorithms in emotion classification for STS and Sanders datasets. The results show that SVM outperforms other classifiers in accuracy and F1. Also, Colneriĉ et al. (2018) tried to classify and predict Twitter users' emotions using several machine learning algorithms. The researchers collected data from Twitter using the hashtags that contain Ekman's emotion classification and Plutchick's emotion classification and then used such classifiers as SVM, Naive Bayes, Logistic Regression, Random Forest, RNN (Recurrent Neural Networks), and CNN (Convolutional Neural Networks) to classify emotions. Their findings indicate that RNN is superior to other algorithms in predicting emotions.

3. Research Questions

The primary purpose of this study was to predict citizen's emotions about home mortgage rates using various machine learning algorithms. To lay the groundwork for achieving these research objectives, I reviewed the literature on emotion classification and machine learning algorithms. As a result of the literature review, I have confirmed that the following gaps exist in the related research. First, existing studies on most emotion classification do not consider which algorithm is superior in predicting each emotion class. Second, there is little research using the AdaBoost algorithm to classify and predict emotions. Third, few studies have attempted to predict emotions on home mortgage rates using machine learning algorithms. To fill these research gaps, this study set the following research questions.

Research Question 1. Which machine learning algorithm is best for predicting citizens' emotions on home mortgage rates?

Research Question 2. Which machine learning algorithm best classifies each emotion class?

4. Methodology

4.1. Data Collection and Preprocessing

First, this study collected 3,000 tweets related to mortgage rates using the Twitter API on February 7, 2019. Because the purpose of this study is to predict the emotions felt by New Yorkers in the United States on mortgage rates, I downloaded and analyzed only the tweets that were uploaded within a radius of 300 miles from Downtown, New York City (coordinates: 40.730610,-73.935242). In the pre-processing stage, retweet, @, punctuation, numbers, HTML links, unnecessary spaces, emojis, and special characters were removed.

4.2. Emotion Extraction from Tweets

I used R's Sentiment Package to extract emotions about New Yorkers' home mortgage rates. Sentiment Package is very effective in classifying emotions, even though it is an old emotion analysis package. Sentiment Package uses WordNet-Affect Lexicon to classify emotions into six types according to Ekman's emotion classification. The WordNet-Affect Lexicon contains 1,536 words related to the six Ekman emotions of anger, disgust, fear, joy, sadness, surprise (Strapparava et al. 2004). The emotions extracted from the tweets were stored in the CSV file format for later analysis.

4.3. Emotion Classification Using Machine Learning Algorithms

As described above, this study extracts one emotion type for each tweet using WordNet-

Affect Lexicon and stores it as a CSV file. Also, for the convenience of analysis, I removed all fields except the Emotion field and Text field from the corpus. First, the pre-processed data was analyzed using Orange, a machine learning program based on Python. First, the stored CSV file was loaded using the Corpus widget of Orange. Then, the Emotion field was set as the target variable and the Text field the meta attribute, respectively. To analyze citizens' emotions about home mortgage rates, I used six machine learning classifiers: AdaBoost, Logistic Regression, Random Forest, Decision Tree, KNN, and Naive Bayes. The 10-fold cross-validation method was used instead of the train/test split method as a machine learning method. In general, the train/test split method can cause overfitting problem. Researchers are using a 10-fold cross-validation method to avoid this problem (Gislason et al. 2006). In 10-fold cross-validation, the original data is randomly divided into ten sub-data of the same size. One of the ten sub-data is used as a sample to test the model, and the remaining nine sub-data are used as data to train the model. In 10-fold cross-validation, this process is repeated 10 times, and the final value is obtained by averaging the values obtained for each fold (Pal 2006). This study also used metrics such as AUC (Area Under Curve), Accuracy, F1, Precision, and Recall to evaluate the performance of each machine learning classifier. Finally, this study identified the machine learning algorithm that classifies each emotion class most accurately using ROC (Receiver Operating Characteristics) analysis.

5. Results

5.1. Research Question1

Research Question 1 was about which machine learning algorithm best predicted citizens' emotions about housing mortgage rates. To accomplish the purpose, this study evaluated the performance of the models using five metrics: AUC, Accuracy, F1, Precision, and Recall. This study used Orange's Test & Score widget and Confusion Matrix widget to evaluate the performance of classifiers. Table 1 shows the performance of each classifier in classifying citizen's emotions about the mortgage rates.

5.1.1. AUC

The AUC measures the two-dimensional area under the entire ROC curve from the coordinates (0,0) to the coordinates (1,1). The AUC value is between 0 and 1. The closer the AUC value is to 1, the better the performance of the algorithm (Wu et al. 2016). The lower the AUC value of any algorithm is, the worse the algorithm is. As shown in Table 1, AdaBoost was the best classifier in terms of AUC. The AUC value of AdaBoost is 0.969, which is higher than those of other algorithms. Also, the AUC values of Random Forest, Logistic Regression, and KNN were all above 0.9, indicating that the performance of those classifiers was high. However, the AUC values of Tree and Naive Bayes are lower than those of the above four algorithms, but they are not very low. Overall, the AUC values of all classifiers remain high, making them excellent

Table 1. Performance Evaluation Results

Method	AUC	CA	F1	Precision	Recall
AdaBoost	0.969	0.914	0.912	0.913	0.914
Random Forest	0.962	0.914	0.910	0.909	0.914
Logistic Regression	0.953	0.870	0.855	0.862	0.870
kNN	0.926	0.872	0.871	0.871	0.872
Tree	0.865	0.885	0.883	0.881	0.885
Naive Bayes	0.820	0.294	0.435	0.946	0.294

algorithms for predicting data.

5.1.2. Classification Accuracy (CA)

Classification Accuracy is one of the most popular metrics used to evaluate classification models. This metric shows how accurate the prediction of a machine learning algorithm is. That is, Classification Accuracy* can be obtained by dividing the number of accurate predictions by the total number of predictions.

1) Adaboost

As shown in Table 1, AdaBoost was the best algorithm in terms of classification accuracy. This classifier accurately predicts 96.69% of New Yorkers' emotions about home mortgage rates. Looking at the Confusion Matrix for Adaboost in Table 2, we can more easily understand the prediction accuracy of the classifier. As can be seen from this figure, the AdaBoost algorithm accurately predicts the emotions of the citizens' home mortgages against all the emotion classes.

Table 2. Confusion Matrix for Adaboost

Actual \ Predicted	Anger	Disgust	Fear	Joy	Sadness	Surprise	Σ
Anger	31	0	2	3	0	3	39
Disgust	0	10	0	2	0	2	14
Fear	0	0	167	33	0	20	220
Joy	0	1	3	2171	0	77	2252
Sadness	0	0	0	3	0	7	10
Surprise	1	0	6	91	3	364	465
Σ	32	11	178	2303	3	473	3000

Table 3. Confusion Matrix for Random Forest

Actual \ Predicted	Anger	Disgust	Fear	Joy	Sadness	Surprise	Σ
Anger	29	0	2	6	0	2	39
Disgust	0	10	0	2	0	2	14
Fear	1	0	170	33	0	16	220
Joy	0	1	5	2194	0	52	2252
Sadness	0	0	0	2	0	8	10
Surprise	0	0	12	114	1	338	465
Σ	30	11	189	2351	1	418	3000

2) Random Forest

The Random Forest Algorithm has also demonstrated the excellence in predicting New Yorkers' emotions about home mortgage rates. That is, the Classification Accuracy value of Random Forest classifier is the same as that of AdaBoost, and it is confirmed that it has excellent prediction ability. The predictability of this classifier can be seen in the Confusion Matrix for Random Forest in Table 3. Random Forest accurately predicts the citizens' emotions about home mortgage rates in almost all emotion classes except surprise.

* Here, Classification Accuracy (CA) can be obtained by using the following equation. CA = TP + TN / TP + TN + FP + FN Where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

Table 4. Confusion Matrix for Decision

Actual \ Predicted	Anger	Disgust	Fear	Joy	Sadness	Surprise	Σ
Anger	29	0	2	5	1	2	39
Disgust	0	8	2	2	0	2	14
Fear	1	0	168	30	0	21	220
Joy	6	3	30	2140	0	73	2252
Sadness	1	0	0	4	0	5	10
Surprise	7	0	29	116	2	311	465
Σ	44	11	231	2297	3	414	3000

3) Decision Tree

The Decision Tree algorithm also showed relatively high accuracy in predicting citizens' emotions about home mortgage rates (Table 1). As can be seen from Table 4, the Decision Tree classifier showed a very high predictive power of 95% for the joy class. However, it showed low Classification Accuracy for other emotion classes such as anger, fear, disgust, sadness, surprise among emotion classes. Notably, in case of sadness emotion class, Classification Accuracy value is 0. Based on these results, we can confirm that the prediction accuracy is low in the emotion classes with low frequency.

4) KNN

The KNN algorithm was also somewhat accurate in predicting New Yorkers' emotions about home mortgage rates (Table 1). First, in the case of the joy class, the algorithm predicts 2,086 out of a total of 2,252, which shows a relatively high Classification Accuracy of 88%. However, in the case of the sadness class, this classifier accurately predicts 0 out of 10, showing a very low prediction capability of 0%.

Table 5. Confusin Matrix for KNN

Actual \ Predicted	Anger	Disgust	Fear	Joy	Sadness	Surprise	Σ
Anger	25	0	1	9	0	4	39
Disgust	0	8	1	5	0	0	14
Fear	1	1	178	22	0	18	220
Joy	5	5	36	2086	0	120	2252
Sadness	0	0	0	3	0	7	10
Surprise	2	2	27	114	0	320	465
Σ	33	16	243	2239	0	469	3000

Table 6. Confusion Matrix for Logistic Regression

Actual \ Predicted	Anger	Disgust	Fear	Joy	Sadness	Surprise	Σ
Anger	20	0	2	16	0	1	39
Disgust	0	0	0	12	0	2	14
Fear	0	0	151	63	0	6	220
Joy	0	0	13	2210	0	29	2252
Sadness	0	0	0	5	0	5	10
Surprise	0	0	12	223	0	230	465
Σ	20	0	178	2529	0	273	3000

5) Logistic Regression

As can be seen in Table 1, the Logistic Regression classifier showed relatively high accuracy in predicting citizens' emotions about mortgage interest rates. In particular, this algorithm has proven to be very good at predicting the joy class. In the case of the joy class, this algorithm correctly predicts 2,210 out of a total of 2,252, which shows a high Classification Accuracy of about 97%. However, this algorithm showed very low Classification Accuracy in disgust class and sadness class.

6) Naive Bayes

As can be seen in Table 1, Naive Bayes was the least accurate algorithm for predicting New

Table 7. Confusion Matrix for Naive Bayes

Actual \ Predicted	Anger	Disgust	Fear	Joy	Sadness	Surprise	Σ
Anger	21	0	2	0	15	1	39
Disgust	0	12	0	0	2	0	14
Fear	0	7	104	0	107	2	220
Joy	0	8	13	679	1546	6	2252
Sadness	0	0	0	0	9	1	10
Surprise	0	7	3	12	387	56	465
Σ	21	34	122	691	2066	66	3000

Yorkers' emotions about home mortgage rates. This algorithm predicts 29.4% of the citizens' emotions about the mortgage rates, so the prediction ability of this algorithm seems to be low. Looking more closely at the Confusion Matrix for Naive Bayes in Table 7, we can see this phenomenon in more detail. The emotion class with the lowest Classification Accuracy was the surprise class. In the case of the surprise class, Naive Bayes algorithm only predicts 56 out of 465.

5.1.3. Precision

Precision refers to the proportion of accurately predicted positive observations among the total predicted positive observations. As shown in Table 1, the classifier with the highest precision value was Naive Bayes, and the algorithm with the lowest precision value was Logistic Regression. However, precision values did not show a significant difference between classifiers.

5.1.4. Recall

Recall means the ratio of the number of true positives to the sum of the number of true positives and the number of false negatives (Moore-Kochlacs et al. 2016). Recall is also called sensitivity. The best classifiers in terms of recall value turned out to be AdaBoost and Random Forest. These high recall values show that the ratio of actually detected true to existing true is very high. Also, algorithms such as Logistic Regression, KNN, and Decision Tree have comparatively high recall values, indicating that they are useful algorithms for predicting citizens' emotions about home mortgage interest rates.

5.1.5. F1 Score

F1 score is also called F-measure or balanced F score. The F1 score is obtained by using the harmonic mean value of precision and recall and has a value between 1 and 0. The F1 score is also used as one of the crucial metrics for evaluating the performance of machine learning algorithms. As shown in Table 1, AdaBoost was the best classifier in terms of F1 score. The Random Forest also has a relatively high F1 value of 0.910, indicating that this algorithm is useful in predicting citizens' emotions about home mortgage rates. However, Naive Bayes was identified as the worst classifier in terms of F1 score. As can be seen in Table 1, Naive Bayes F1 score is 0.294, demonstrating that the algorithm does not play a role in predicting citizens' emotions about home mortgage rates.

5.2. Research Question2

Research Question 2 was about which machine learning algorithm best predicts each emotion class. This study used Ekman's emotion classification to categorize citizens' emotions about home mortgage rates. ROC analysis was carried out to find out which algorithm is best for predicting each emotion class. Orange's ROC Analysis widget was used for this analysis. The ROC curve is a curve showing the performance of the classifier at all classification thresholds. The ROC curve displays true positive rate (TPR) on the y-axis and false positive rate (FPR) on the x-axis (Zhang et al. 2018). That is, the ROC curve shows the combination of TPR and FPR of various classification thresholds. If the classifier's ROC curve is above the baseline, it means good performance level. If the classifier's ROC curve is below the baseline, it indicates poor performance level. It also means random performance level if the ROC curve of any classifier matches the baseline. If the ROC curve of a classifier connects the origin (0,0), the point (0,1), and the point (1,1) in straight lines, the performance level of this algorithm is considered perfect. The closer the ROC curve of a classifier is to the perfect curve, the higher the performance level of the classifier is.

5.2.1. Anger

Figure 1 shows the ROC curves for anger emotion class. As can be seen in this figure, the ROC curves for all six algorithms used in this study are located above the baseline (dashed line), indicating that all classifiers predict the anger class well. However, as can be seen in this figure, the performance level of each classifier shows a marked difference. Among the ROC curves of the six classifiers, the closest to the perfect ROC curve was the ROC curve of AdaBoost. Therefore, we can confirm that the AdaBoost algorithm is most effective in predicting anger among the emotion classes for home mortgage rates, and the ROC curve of the logistic regression is higher than those of the other classifiers, thus showing a higher performance level. However, the ROC curves of Naive Bayes and KNN are located lower than those of other algorithms, indicating that the performance levels are relatively low.

5.2.2. Disgust

Figure 2 shows the ROC curves for the disgust emotion class. As can be seen in this figure, the ROC curves of all classifiers are above the baseline.

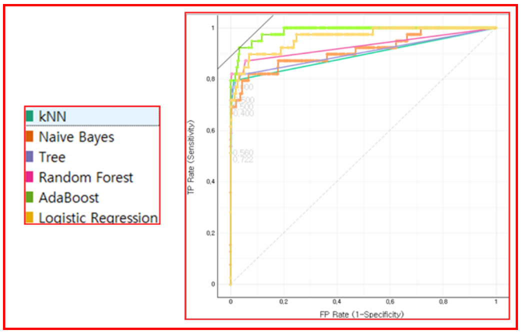

Figure 1. ROC curves for anger emotion class

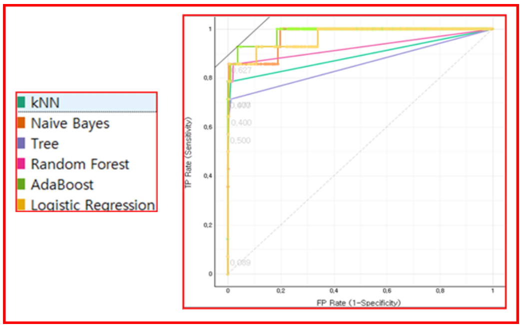

Figure 2. ROC curves for disgust emotion class

Therefore, we find that all the algorithms used in this study are useful in predicting disgust among emotion classes for home mortgage rates. However, the performance level shows a big difference between the algorithms. Among the six algorithms used in this study, AdaBoost was the classifier with the highest performance level. The ROC curves of the Logistic Regression and the Random Forest are also located above those of the other classifiers. Thus, we can confirm that Logistic Regression and Random Forest are effective in predicting disgust among the emotion classes. However, the ROC curves of Decision Tree and KNN are located below those of other classifiers. Therefore, the performance levels of Decision Tree and KNN seem to be lower than those of other algorithms.

5.2.3. Fear

Figure 3 shows the ROC curves for the fear emotion class. As can be seen in this figure, the closest to the perfect curve was the ROC curve of AdaBoost. Thus, we can see that AdaBoost is the most effective algorithm for predicting the fear class among the classes of emotions for home mortgage rates. However, the ROC curves of Decision Tree and KNN are located below those of other algorithms. Therefore, we can recognize that the performance levels of KNN and Decision Tree are lower than those of other classifiers.

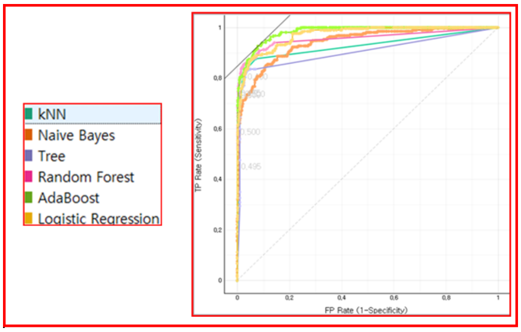

Figure 3. ROC curves for fear emotion class

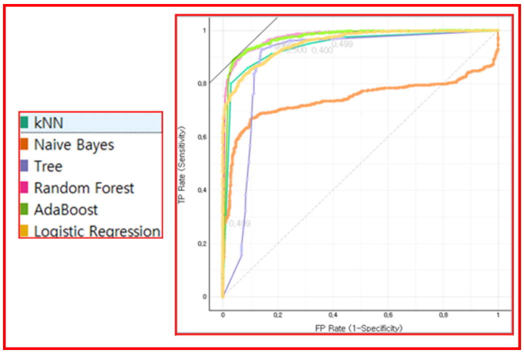

Figure 4. ROC curves for joy emotion class

5.2.4. Joy

Figure 4 shows the ROC curves for the joy emotion class. As can be seen in this figure, the ROC curves located at the top of the graph were those of AdaBoost and Random Forest. Therefore, AdaBoost and Random Forest seem to be very useful algorithms for predicting the joy emotion

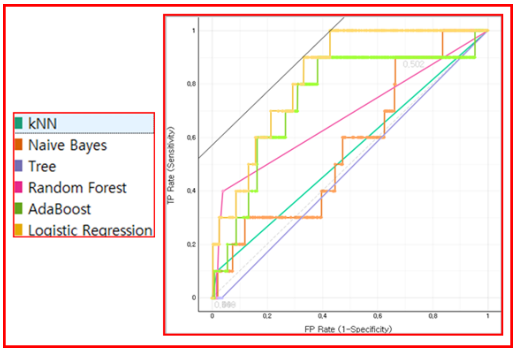

Figure 5. ROC curves for sadness emotion class

class. However, the ROC curve of Naive Bayes was lower than those of other algorithms in all sections except for the initial part. So this reveals that the performance level of Naive Bayes is lower than those of other classifiers.

5.2.5. Sadness

Figure 5 shows the ROC curves for the sadness emotion class. As can be seen in this figure, the performance levels of all classifiers in predicting the sadness class are much lower than those in the other emotion classes. Although AdaBoost and Logistic Regression were somewhat effective in predicting sadness class, Decision Tree, Naive Bayes, and KNN did not predict the sadness class accurately.

5.2.6. Surprise

Even in the surprise emotion class, there is a big difference in performance levels among classifiers. First, AdaBoost, Random Forest, and Logistic Regression seem to predict the surprise emotion class more accurately than other algorithms. However, as in other emotion classes, Naive Bayes

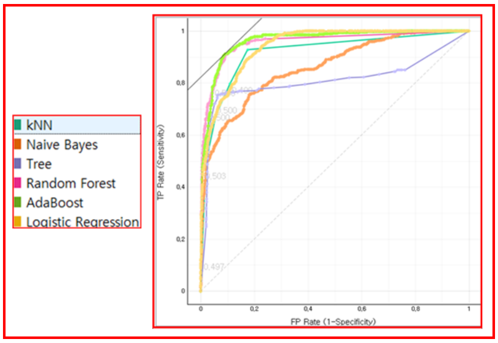

Figure 6. ROC curves for surprise emotion class

and Decision Tree seems to have lower performance levels than other classifiers.

6. Discussion

6.1. Research Question 1

Home mortgage rates play an important role in housing policy-making process. If mortgage rates rise, housing demand generally declines and the housing market eventually stagnates. In addition, if the home mortgage rates fall, it stimulates the citizens' desire to buy the housing, which eventually contributes to the activation of the housing market. To forecast the housing market tends, it is necessary to predict citizens' emotions about mortgage rates accurately. However, the emotions of residents can change over time, and the performance levels vary according to the classifiers used by researchers. From this point of view, the results of this study suggest a great deal in predicting the residents' emotions about home mortgage rates. This study used six classifiers to predict citizens' emotions about home mortgage rates. However, performance levels differed significantly among classifiers. The best-performed algorithm was AdaBoost. AdaBoost outperformed other algorithms in almost all emotion categories. In the future, I would like to strongly recommend the use of this algorithm to classify and predict citizens' emotions about specific issues using Twitter data or text data. The performance levels of classifiers also depend on the data size. Therefore, it can be a useful research experience to examine how the performance level changes when data size is different

6.2. Research Question 2

In this study, citizens' emotions regarding home mortgage rates were classified into six classes: anger, disgust, fear, joy, sadness, and surprise. In almost all emotion classes AdaBoost outperformed all other algorithms. However, the difference in performance levels was not greater than expected. In most emotion classes, the ROC curves of all algorithms are located above the baseline, confirming that all algorithms have high predictive power. However, the performance levels of classifiers in predicting the sadness emotion class were not high. The ROC curves of some classifiers in the sadness category were located near the baseline. Sadness generally falls into the category of negative emotions. It is essential to identify positive emotions as well as negative feelings at the same time to pinpoint the problems that the housing market faces. Therefore, the use of a new machine learning algorithm should be attempted to predict the sadness emotion class accurately. ROC curves can be more effective when used with AUC. Often, the ROC alone is difficult to determine which classifier is more predictive. For example, the ROC curve of a classifier may be located at the bottom in the early part but may be in the upper portion after the middle section. In this case, it is desirable to compare the performance levels of the classifiers by using the ROC curve and the AUC value at the same time.

7. Conclusion

The primary purpose of this study was to predict the citizens' emotions about home mortgage rates using machine learning algorithms. To accomplish the research objective, I reviewed the literature on emotion classification and machine learning algorithms. The literature review confirmed that there were several gaps in existing research and then this study set up two research questions to fill these gaps. Also, to find the answers to the research questions, this study downloaded 3,000 tweets that include citizen's emotions regarding home mortgage rates, classified emotions according to Akman's classification, and then used six machine learning algorithms to predict residents' emotions concerning home mortgage rates.

As a result of the analysis, AdaBoost was confirmed as the best classifier in all evaluation categories. Also, Random Forest and Logistic regression were found to be algorithms that predict residents' emotions regarding home mortgage rates well. However, the performance level of Naive Bayes was significantly lower than those of other classifiers. Also, this study conducted a ROC analysis to determine which classifier predicts each emotion class accurately. The results of the analysis showed that AdaBoost was the best algorithm for predicting citizens' emotions about home mortgage rates in all categories of anger, disgust, fear, joy, sadness, surprise. However, in the sadness class, the performance levels of the six algorithms used in this research were not much higher than those in the other emotion categories.

This study can contribute to the related research fields in several aspects. First, this research was the first attempt to use the machine learning algorithm to analyze citizens' emotions concerning home mortgage rates. Second, this study was the first effort to use the Akman emotion classification method to predict the citizens' emotions regarding home mortgage rates. Third, this research is meaningful in that ROC analysis is used to compare the performance levels of the algorithms by emotion classes.

Despite its usefulness and differentiation, this study has the following limitations. First, because this research used only 3,000 tweet data, it may have derived only limited research results. In future studies, if we increase the number of data to 10,000 or more, more objective and accurate research results will be obtained. Second, in this research, the classifiers used to predict residents' emotions regarding home mortgage rates did not perform well in predicting the sadness among the emotion classes. Since sadness is an essential part of the emotion classes, it is necessary to predict the sadness class more accurately in future studies. New classifiers such as SVM or MLP should be used for this purpose. Finally, this study conducted only cross-sectional analysis to predict New Yorkers' emotion about home mortgage rates. However, emotions regarding home mortgage rates can be changed by time and place. Therefore, in future research, time series analysis or regional comparative analysis should be carried out to predict citizens' emotions regarding home mortgage rates accurately.

참고문헌
References

Alkan A, Koklukaya E, Subasi A. 2005. Automatic seizure detection in EEG using logistic regression and artificial neural network. *Journal of Neuroscience Methods*, 148(2):167-176.

Almeida AM, Cerri R, Paraiso EC, Mantovani RG, Junior SB. 2018. Applying multi-label techniques in emotion identification of short texts. *Neurocomputing*, 320:35-46.

Arnold MB. 1960. Emotion and personality. Vol. I. Psychological aspects.

Avetisyan H, Bruna O, Holub J. 2016. Overview of existing algorithms for emotion classification. Uncertainties in evaluations of accuracies. *Journal of Physics: Conference Series*. 772(1): 012039.

Bravo-Marquez F, Mendoza M, Poblete B. 2013. Combining strengths, emotions and polarities for boosting Twitter sentiment analysis. *Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining*.

Breiman L. 1999. *Random forests*. UC Berkeley TR567.

Burnap P, Colombo W, Scourfield J. 2015. Machine classification and analysis of suicide-related communication on twitter. *Proceedings of the 26th ACM conference on hypertext & social media* p. 75-84.

Burnap P, Williams ML. 2015. Cyber hate speech on twitter: An application of machine classification and statistical modeling for policy and decision making. *Policy & Internet*. 7(2):

223-242.

Cannuscio CC, Alley DE, Pagán JA, Soldo B, Krasny S, Shardell M, Lipman TH. 2012. Housing strain, mortgage foreclosure, and health. *Nursing outlook,* 60(3):134-142.

Casale S, Russo A, Scebba G, Serrano S. 2008. Speech emotion classification using machine learning algorithms. *The IEEE International Conference on Semantic Computing*, p. 158-165. IEEE.

Chen YL, Chang CL, Yeh CS. 2017. Emotion classification of YouTube videos. *Decision Support Systems*. 101:40-50.

Colnerič N, Demsar J. 2018. Emotion Recognition on Twitter: Comparative Study and Training a Unison Model. *IEEE Transactions on Affective Computing*.

Damrongsakmethee T, Neagoe VE. 2017. Data Mining and Machine Learning for Financial Analysis. *Indian Journal of Science and Technology*. 10(39).

Dong L, Li X, Xie G. 2014. Nonlinear methodologies for identifying seismic event and nuclear explosion using random forest, support vector machine, and naive Bayes classification. *Abstract and Applied Analysis*. Vol. 2014. Hindawi.

Du W, Zhan Z. 2002. Building decision tree classifier on private data. *Proceedings of the IEEE international conference on Privacy, security and data mining*. 14:1-8.

Ekman P. 1992. An argument for basic emotions. *Cognition & emotion*. 6(3-4):169-200.

Ekman P, Friesen WV. 1971. Constants across cultures in the face and emotion. *Journal of personality and social psychology*. 17(2):124.

Frijda NH. 1986. *The emotions*. Cambridge University Press.

Gievska S, Koroveshovski K. 2014. The impact of affective verbal content on predicting personality impressions in youtube videos. *Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition*. p. 19-22.

Hastie T, Rosset S, Zhu J, Zou H. 2009. Multi-class adaboost. S*tatistics and its Interface*. 2(3): 349-360.

Hu W, Hu W, Maybank S. 2008. Adaboost-based algorithm for network intrusion detection. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*. 38(2): 577-583.

Gerardi K, Shapiro AH, Willen P. 2007. Subprime outcomes: Risky mortgages, homeownership experiences, and foreclosures.

Gil GB, de Jesús AB, Lopéz JMM. 2013. Combining machine learning techniques and natural language processing to infer emotions using Spanish Twitter corpus. *International Conference on Practical Applications of Agents and Multi-Agent Systems*. p. 149-157.

Gislason PO, Benediktsson JA, Sveinsson JR. 2006. Random forests for land cover classification. *Pattern Recognition Letters*, 27(4):294-300.

Go A, Bhayani R, Huang L. 2009. Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford*. 1(12).

Jang E, Rak B, Kim S, Sohn J. 2012. Emotion classification by machine learning algorithm using physiological signals. *Proc. of Computer Science and Information Technology. Singapore*, 25:1-5.

Lim JS, Kim JM. 2014. An empirical comparison of machine learning models for classifying emotions in Korean Twitter. *Journal of Korea Multimedia Society*. 17(2):232-239.

Luo C, Wu D, Wu D. 2017. A deep learning approach for credit scoring using credit default swaps. *Engineering Applications of Artificial Intelligence*. 65:465-470.

Mohammad SM. 2012. Emotional tweets. *Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation*. p. 246-255.

Moore-Kochlacs CE. 2016. Extracellular electrophysiology with close-packed recording sites: spike sorting and characterization [dissertation].

Neethu MS, Rajasree R. 2013. Sentiment analysis in twitter using machine learning techniques. *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*. p.1-5.

Ng AY, Jordan MI. 2002. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in neural information processing systems*. p. 841-848.

Ortony A, Turner TJ. 1990. What's basic about basic emotions?. *Psychological review*. 97(3): 315.

Pal M. 2005. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*. 26(1):217-222.

Pal M. 2006. Support vector machine-based feature selection for land cover classification: a case study with DAIS hyperspectral data. *International Journal of Remote Sensing*. 27(14):2877-2894.

Peterson LE. 2009. K-nearest neighbor. *Scholarpedia*. 4(2):1883.

Plutchik R. 1990. Emotions and psychotherapy: A psychoevolutionary perspective. In *Emotion, psychopathology, and psychotherapy*. p. 3-41.

Plutchik R. 1991. *The emotions*. University Press of America.

Roberts K, Roach MA, Johnson J, Guthrie J, Harabagiu SM. 2012. EmpaTweet: Annotating and Detecting Emotions on Twitter. *LREC*. 12:3806-3813.

Rodriguez-Galiano VF, Ghimire B, Rogan J, Chica-Olmo M, Rigol-Sanchez JP. 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. I*SPRS Journal of Photogrammetry and Remote Sensing*. 67:93-104.

Safavian SR, Landgrebe D. 1991. A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*. 21(3):660-674.

Sitthi A, Nagai, M., Dailey, M., & Ninsawat, S. 2016. Exploring land use and land cover of geotagged social-sensing images using naive bayes classifier. *Sustainability*. 8(9):921.

Strapparava C, Valitutti A. 2004. Wordnet affect: an affective extension of wordnet. *Lrec*. 4: 1083-1086.

Subasi A, Ercelebi E. 2005. Classification of EEG signals using neural network and logistic regression. *Computer methods and programs in biomedicine*. 78(2):87-99.

Tan S. 2006. An effective refinement strategy for KNN text classifier. *Expert Systems with Applications*. 30(2):290-298.

Tang D, Qin B, Liu T, Li Z. 2013. Learning sentence representation for emotion classification on microblogs. *Natural Language Processing and Chinese Computing*. p. 212-223. Springer, Berlin, Heidelberg.

Tang D, Wei F, Qin B, Liu T, Zhou M. 2014. Coooolll: A deep learning system for twitter sentiment classification. *Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014)*. p. 208-212.

Vatsavai RR, Bright E, Varun C, Budhendra B, Cheriyadat A, Grasser J. 2011. Machine learning approaches for high-resolution urban land cover classification: a comparative study. *Proceedings of the 2nd International Conference on Computing for Geospatial Research & Applications*. p.11.

Velásquez F, Gordon J. 2012. Empirical study of machine learning based approach for opinion mining in tweets. *Mexican international conference on Artificial intelligence*. p. 1-14. Springer, Berlin, Heidelberg.

Wang W, Chen L, Thirunarayan K, Sheth AP. 2012. Harnessing twitter "big data" for automatic emotion identification. *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*. p. 587-592. IEEE.

Wei Y, Bing X, Chareonsak C. 2004. FPGA implementation of AdaBoost algorithm for detection of face biometrics. *Biomedical Circuits and Systems, 2004 IEEE International Workshop on*. p. S1-6.

Wen S, Wan X. 2014. Emotion classification in microblog texts using class sequential rules. In *Twenty-Eighth AAAI conference on artificial intelligence*.

Wu Y, Zhang T, Hou X, Xu C. 2016. New Blind Steganalysis Framework Combining Image Retrieval and Outlier Detection. *KSII Transactions on Internet &Information Systems*. 10(12):6206-6212.

Yadav P, Aggarwal G. 2015. Speech Emotion Classification using Machine Learning. *International Journal of Computer Applications*. 118(13):44-47.

Zhang Y, Liu S. 2018. Analysis of structural brain MRI and multi-parameter classification for Alzheimer's disease. *Biomedical Engineering/Biomedizinische Technik*. 63(4):427-437.

2019년 04월 16일 원고접수(Received)
2019년 05월 17일 1차심사(1st Reviewed)
2019년 06월 03일 2차심사(2st Reviewed)
2019년 06월 18일 게재확정(Accepted)

초 록

본 연구의 목적은 기계학습 알고리즘을 이용하여 주택모기지 금리에 대한 시민들의 감정을 예측하는 것이었다. 연구목적을 달성하기 위해 본 연구는 관련문헌을 검토한 다음 두개의 연구 질문을 설정하였다. 또한 연구 질문에 대한 답을 구하기 위해 본 연구는 Akman의 분류에 따라 감정을 분류 한 다음 여섯 가지 기계학습 알고리즘을 이용하여 모기지 금리에 대한 시민들의 감정을 예측하였다. 분석결과 AdaBoost가 모든 평가범주에서 가장 우수한 분류기로 확인되었다. 그러나 Naive Bayes의 성능수준은 다른 분류기들의 성능수준보다 낮은 것으로 밝혀졌다. 또한 본 연구는 어느 분류기가 각 감정범주를 잘 예측해주는지를 파악하기 위해 ROC 분석을 실시하였다. 분석결과, AdaBoost가 모든 감정범주에서 주택모기지 금리에 대한 주민들의 감정을 가장 잘 예측해주는 것으로 확인되었다. 그러나 슬픔범주에서 여섯 가지 알고리즘의 성능수준은 다른 감정범주보다 훨씬 낮게 나타났다.

주요어 : 기계학습, 알고리즘, 모기지 금리, Akman, 감정분류, AdaBoost, 분류기, Naive Bayes, 성능수준