Predicting Duration of Traffic Accidents Based on Ensemble Learning

Lina Shan, Zikun Yang, Huan Zhang, Ruyi Shi, and Li Kuang(✉)

Central South University, Changsha 410075, China
kuangli@csu.edu.cn

Abstract. Traffic congestion can be divided into recurrent congestion and accidental congestion, and the latter one is usually caused by traffic accidents. It is of great significance to predict the duration of traffic accidents accurately and transfer the results to drivers on the road in time. Most of the existing works utilize traditional, single machine learning model to predict the duration of accident, while the accuracy is not satisfying. In this paper, we firstly construct and extract features from the accident records including description, location, as well as some external information such as weather. We then divide the duration into multiple periods, corresponding to multiple categories. In order to improve the prediction precision of rare categories, we convert the multi-class classification problem into a binary classification problem, constructing multiple XGBoost binary classifiers which are restricted by F1 (harmonic mean) evaluation index. Finally, in order to improve the overall accuracy further, the classification results are integrated by using artificial neural networks. The experiment is conducted on real datasets in Xiamen and employs mean absolute percentage error (MAPE) and root-mean-square error (RMSE) as indicators. The experimental results show the effectiveness of the proposed method and show better performance in comparison with traditional models.

Keywords: The duration of traffic accidents · XGBoost · Artificial neural networks

1 Introduction

With the continuous growth of the total number of urban vehicles, the frequency of road traffic accidents has also increased. Prediction about traffic accident is an important part of research on intelligent transportation, since it can help us explore the rules of the occurrence of traffic accidents, control road safety and design suitable strategies. It is an important research issue to predict the duration of traffic accidents effectively and accurately, so that we can schedule the traffic scientifically and reduce the probability of second accidents.

At present, most researches mainly focus on predicting and analyzing the quantity of traffic accidents, including the number of accidents per year, the frequency of accidents at certain places, and the frequency of accidents at specific period. There are only a few studies on real-time traffic accident prediction. Initially, scholars use traditional machine learning methods such as Naive Bayes, KNN, multi-variable linear

H. Gao et al. (Eds.): CollaborateCom 2018, LNICST 268, pp. 252–266, 2019.
https://doi.org/10.1007/978-3-030-12981-1_18

regression and decision trees to predict the traffic accident duration. After taking the diversity of feature attributes in traffic accident data into consideration, scholars fuse different traditional models together which allows different models to handle specific feature attributes to achieve better results. In addition, there are some studies aiming at solving the heterogeneity of traffic accident data, the confidence interval of predicted values, and the probability of second accidents.

However, there are still some shortcomings in existing works: (1) The features are not fully excavated and utilized, for example, the feature that is extracted from accident description is just one-dimensional with some discrete values such as scratching, rear-end collision and running into fixtures, as well, the latitude and longitude are used separately so that the combined feature is lost. (2) The duration of traffic accidents has been proved to follow the log normal distribution, i.e., the data are unbalanced, thus the prediction accuracy of rare categories is not so satisfying. (3) In the aspect of modeling, scholars usually just use single traditional model for prediction, and the overall accuracy is limited.

In this paper, we propose that (1) In the extraction of features, the accident description is encoded as a multi-dimensional vector so that more semantics can be captured, the geographical locations are processed via binning and a reasonable interval is determined through visualization, as well, fusion of isomeric characteristics such as temperature and wind direction. (2) After the feature variable is processed, we should classify the traffic accident duration by training the classifier, and choose XGBoost to classify traffic duration. Then, we convert the accident duration prediction into a multi-binary classification task. We use the accuracy of the evaluation index and recall rate to constrain the model and solve the problem of unbalanced distribution of label training samples, which makes the model more convincing in solving the problem. (3) In order to improve the overall prediction accuracy further, we use neural networks to integrate the prediction results of multiple binary classification models and output the average duration of accidents under this label.

The rest of the paper is organized as follows: Sect. 2 describes the related work of the traffic accidents duration. Section 3 outlines the methods for predicting the duration of traffic accidents. Section 4 gives the experimental verification designs and results. Section 5 summarizes the whole paper.

2 Related Work

In recent years, the problem of predicting the duration of traffic accidents has attracted wide attention of scholars. Related works on the research issue focus on the following aspects: (1) investigating the statistical characteristics of traffic accident data (2) exploring the influence factors which affect the length of accident (3) constructing a proper prediction model.

The analysis of the characteristics of the accident duration distribution in different traffic accident data sets is the first step to build the prediction model. Different prediction models are designed for different distribution characteristics, which is beneficial for improving the accuracy of prediction models. Golob and Recker [1–3] analyzed the durations of the incidents involving large trucks, and demonstrated that the accident duration fits a lognormal distribution. Jones and Chung et al. [4–6] has proved that the accident duration follow a bilogarithmic probability distribution for the accidents on Seattle Expressway. Some researchers find out different distributions in different datasets, such as logarithmic normal distribution, log-logical distribution, Weibull distribution, Gamma distribution, generalized distribution, and etc. [7].

Research during the past few decades has demonstrated that various methodologies and techniques have been employed to analyze and model incident duration, mainly on freeways. These models have determined the relationships between incident duration and influencing variables. The main features that they have found include: accident types [8], severity [9], the number of vehicles involved, time characteristics [10], geography characteristics, weather, etc.

In the perspective of constructing a proper prediction model, many classification methods have been developed to predict the duration of traffic accidents, including decision trees [11], Bayesian network [12], clustering [13], GBDT [14], topic models [15], SVM [16], and etc. For example, Zhan [11] utilizes the M5P tree algorithm for lane clearance time prediction, which has an advantage to deal with categorical and continuous variables and variables with missing values. Yang [12] have proved that Bayesian network has the goodness-of-fit results in traffic dataset when compared with other models. Weng [13] develops a cluster-based lognormal distribution model, and the model can predict the mean and the probability of an accident duration from the base accident information.

Many regression models have also been applied in predicting traffic incident duration. For example, Khattak [17] applies a quantile regression to predict the durations of larger incidents more accurately, since quantile regression can estimate the probability of an incident lasting for a specific duration. He [18] proposed a hybrid tree-based quantile regression method to predict the duration time, and the model has the robustness to outliers, and flexibility in combining categorical covariates. Peeta et al. [19] applied a linear regression model to predict incident clearance time with the time-independent variables.

In recent years, in order to improve the accuracy and efficiency of the prediction model further, many scholars start to use neural networks to predict the duration of traffic accidents. For example, Vlahogianni [20] developed a neural network model for incident duration prediction with single and competing uncertainties, and they improved the generalization power of the prediction models. Wei et al. [21] developed two ANN-based models that sequentially predict traffic accident duration, and the results showed that these models achieve a reasonable prediction. Park [22] introduces Bayesian learning to neural networks for accurate prediction of incident duration, and they updated the network parameters using a hybrid Monte Carlo algorithm.

Compared with these related work that already exists, our main innovations include: (1) In this paper, the multidimensional features coding is used to deal with the accident description characteristics, and the geographical location dealt by points box method. (2) The prediction of accident duration is transformed into a learning task of multiple dichotomies, and the prediction accuracy of rare categories improved by using F1 evaluation index. (3) The prediction results of multiple binary classification models are integrated by using neural network to improve the prediction accuracy.

3 Method

3.1 Problem Definition

Given the data set $R = \{r_1, r_2 ... r_n\}$, where r_i denotes the ith traffic accident history, r_i can be represented as a 6-tuple, $r_i = \langle$X, Y, car_number, description, start_time, end_time$\rangle$, where X and Y are the longitude and latitude of the accident location; car_number is the number of vehicles involved in the accident; description is the description of the accident, mainly including the safety and responsibility of the persons involved, and the damage of vehicles; the start_time is the occurrence time of accident; the end_time is the time when the road is clear. The duration of the accident is the end time minus the start time. Some samples are shown in Table 1.

Table 1. Samples of traffic accident dataset

Start time	X	Y	Car number	Accident description	End time
2015/01/01 00:09	118.110169	24.48329	2	A does not ensure safety, B has no fault, the accident results in a right side damage of car A, a left side damage of car B	2015/01/01 00:42
2015/1/1 8:36	118.135313	24.647768	1	A collided with the fixture and an accident occurred, causing damage to the nail tail and right rear corner	2015/1/1 11:09

Given the data set R, and a new accident e = $\langle$start_time, X, Y, car_number, description$\rangle$, based on the historical records, we aim to predict the duration of the new accident, according to the given information including the latitude and longitude of the new accident, the number of vehicles involved, the accident description and the start time.

3.2 Solution

Figure 1 shows a traffic accident duration prediction architecture based on the ensemble learning of XGBoost and ANN, which contains three main steps: **(1) Data preprocessing**: we first process data cleaning, including processing missing data, uniforming format content and removing unreasonable values. Then we analyze the distribution of traffic accident duration and determine reasonable intervals for classification. After that, we construct features for input. **(2) Construct binary classifiers based on XGBoost**: We convert a multi-class classification problem into a multiple binary classification problem. The training set is divided according to the class label, multiple XGBoost binary classifiers are constructed, and parameters for each XGBoost are adjusted by multiple-objective optimization. **(3) Ensemble Learning based on ANN**: Multiple results of binary classification models are integrated based on ANN, so the final category of the predicted duration can be obtained and the average value under the label is output as the prediction value.

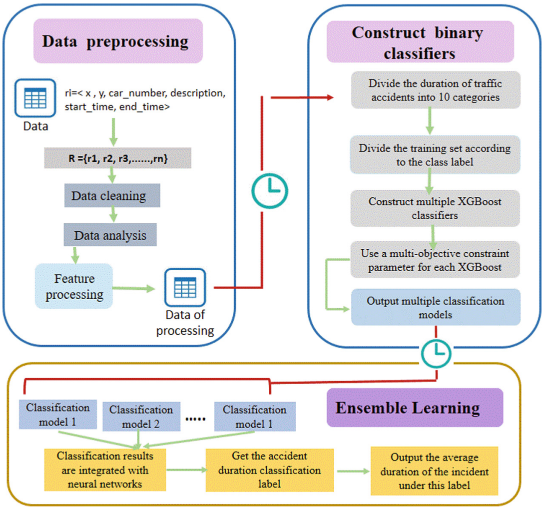

Fig. 1. Architecture of the proposed solution to accident duration prediction

3.3 Data Preprocessing

Data Cleaning and Analysis. The performance of the prediction model is affected by the quality of the data largely. Therefore, we first process the data before constructing features, including removing missing values and unreasonable values, unifying the format content.

The second step is to analyze the distribution of accident durations. The result is shown as Fig. 2. We can see from Fig. 2 that the samples are unbalanced, the accident durations distribute mainly in the range between 6 min to 40 min, accounting for about 70% of the total samples, and the number of samples become fewer and fewer with the increase of duration when the duration is longer than 20 min.

After data cleaning, we can get 39,000 samples approximately. Before we predict the specific minutes of the duration for a target accident, we first try to classify the duration of a target accident into a range, therefore, it is important to divide the durations into multiple ranges reasonably. A primary rule is that we try to make the numbers of accident samples in all ranges are basically equal. We have tried three sizes of each range, which are 5000, 10000 and 20000. For example, when we set the size of each range as 5000, we can divide the durations into [6–13], [14–17], [18–21], [22–25],

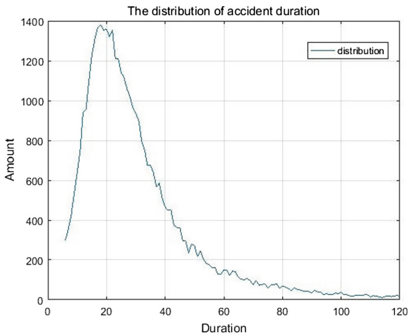

Fig. 2. The distribution of accident duration

[26–30], [31–37], [38–50], [51–109]. We then calculate the average time and variance value in each range. We find that the variance is very large in the last range, since the samples with long duration becomes sparse. Therefore, we aim to divide the last range further, and the division principle is the variances of the subdivisions should be less than a specific threshold value. We have tried three threshold values, which are 10, 15 and 20, and we compare the MAPE values under each threshold respectively, we finally find that the prediction performs the best when the variance threshold is 15. So we divide the last range [51–109] into [51–68], [69–86], and [87–109]. We finally set the division of accident durations as shown in Table 2.

Table 2. Division of traffic accident durations into 10 ranges

Duration time	Amount time	Average	Variance
[6–13]	4984	9.5	4.95
[14–17]	4982	15	2.121
[18–21]	5402	19	2.121
[22–25]	4895	23	2.121
[26–30]	5072	27.5	2.828
[31–37]	4970	33.5	4.243
[38–50]	4885	43.5	8.485
[51–68]	2686	59	12.02
[69–86]	1228	77	12.02
[87–109]	696	97.5	14.09

Feature Processing. In order to prepare for the subsequent algorithms, we then process the original data and integrate external features which affect traffic accidents.

(1) **'start_time':** In addition to the year, month, day, hour, and minute of the accident, we can also get from the original information whether the day is a workday or weekend, whether the time belongs to a peak period or not.

(2) **'car_number'** is the number of cars that are involved in the accident. According to the statistics, the number of accidents involving one car is 3211, that involving 2 cars is 35571, that involving 3 cars and above is 1623. Therefore, the mapping rule is as follows: 1 stands for single-car accident, 2 stands for double-cars accident, 3 stands for multi-cars accident.

(3) **'x', 'y'** are the latitude and longitude where the accident occurred. We point the pairs of latitude and longitude into boxes, and Fig. 3(a)–(d) correspond to the binning results under different block sizes, where the area is divided into 5 * 5, 10 * 10, 15 * 15, 20 * 20 blocks respectively. The color of the block represents the average duration time of accidents that occurred within the region. It can be seen from Fig. 3(a)–(c) that the differences of average duration time in adjacent blocks become significant with the increase of blocks, while the number of blocks in Fig. 3(d) increases further, the difference in adjacent blocks is reduced, since the blocks with average duration between 10 to 20 min are missing. Therefore, it is reasonable to set the division size to 15 * 15.

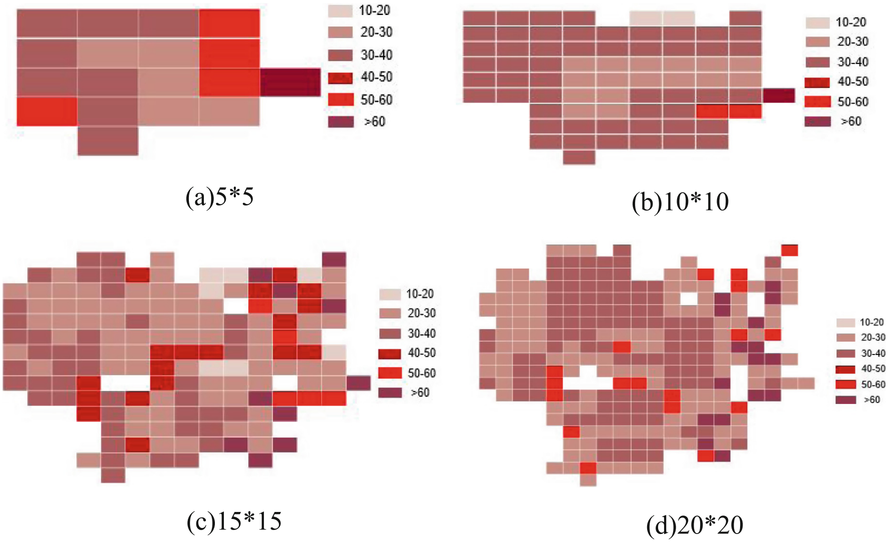

Fig. 3. Binning results under different block sizes

(4) **'weather':** Weather has been proved to be an important factor that has great impact on occurrence and clearance of accidents [10]. Therefore, we further crawled the weather information in Xiamen in 2015, and Table 3 shows the sample data.
We process the obtained information and define 0–10 °C as cold, 10–20 °C as cool, 20–30 °C as comfortable, and 30 °C or higher as hot. In addition, we classify the weather conditions into 5 categories. Table 4 shows the encoding of temperature and weather feature.

(5) **'description'** describes the main information of accident. It briefly explains the responsible party of the traffic accident, the cause of the accident, the casualties

Table 3. The sample of meteorological dataset

Date	Min_ temperature	Max_ temperature	Weather	Wind direction
2015-01-01	8	15	Cloudy	Northeasterly
2015-01-02	9	18	Cloudy	East
2015-01-03	9	19	Cloudy	North
2015-01-04	9	17	Cloudy	East

Table 4. Encoding of temperature and weather

Parameter	Initial value	Parameter significance
Gamma	0.1	Parameters used to control the post pruning
Maxdepth	12	The depth of the tree
Lambda	2	Regularized term parameter
Subsample	0.7	Random sample training sample

and the vehicle collision. From the description, we can extract four kinds of features, which are responsible party, accident characteristic, vehicle damage grade and accident level. Since the characteristics of an accident in each category may occur at the same time, we employ 13-digit 0–1 coding for the features, and 1 represents that the behavior described on this bit has occurred, in contrast, 0 means not. The meaning represented by each digit is shown in Table 5.

Table 5. The feature coding of accident description

Feature	Coding
Accident responsible party coding [1–3]	1: A no-fault behavior
	2: B no-fault behavior
	3: C no-fault behavior
Accident characteristic coding [4–8]	4: Rear-end
	5: Curettage (vehicle to vehicle)
	6: Curettage (vehicle to person.et)
	7: Bumper
	8: Rollover
Vehicle damage grade coding [9–11]	9: Unilateral damage
	10: Bilaterally damage
	11: Multilateral damage
Accident level coding [12, 13]	12: Vehicle loss
	13: Personal injury

For example, <1000100001010> means that A has no-fault behavior; the accident is caused by the curettage between cars; it leads to the damage on both sides of the vehicle, and it just causes vehicle loss.

3.4 Construct Binary Classifiers

Based on the extracted features and processed data, then we aim to classify the traffic accident duration into a reasonable range first. We choose XGBoost as the classifier. The basic idea is to train multiple weak classifiers for the same training set, and then, these classifiers are combined to form a strong classifier.

As explained in Sect. 3.3, the duration of traffic accidents is divided into 10 ranges, which are [6–13], [14–17], [18–21], [22–25], [26–30], [31–37], [38–50], [51–68], [69–86] and [87–109]. In order to construct an effective classifier, we need to consider the size of training samples under the category tag. The sizes are almost the same for the former 7 categories, while the sizes are less for the latter 3 categories, i.e. the samples are unbalanced for the 10 categories. Therefore, we do not try to classify the duration of a new accident into one of the 10 categories directly, instead, we convert the multi-class classification problem into multiple binary classification problem. Figure 4 shows the construction of multiple binary classifiers. We use the accuracy and recall rate to constrain the model, in order to solve the problem of unbalanced distribution of samples.

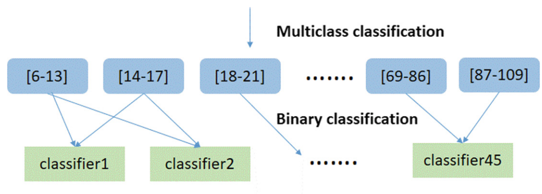

Fig. 4. The process of constructing a binary classifier

As Fig. 4 shows, each classifier performs a binary classification task. For example, we construct the classifier1 which classifies a sample into [6–13] or [14–17], so we pick out the samples in the range of [6–13] and [14–17] to train classifier1. Similarly, we construct the classifier2 which classifies a sample into [6–13] or [18–21], and so on. Every two ranges are formed together as a classification target. In this way, we need to build C(10, 2) classifiers for 10 categories. Then we adjust the parameters of a single XGBoost model to make it perform well in terms of accuracy and recall. Table 6 shows the initial values of the parameters which affect the accuracy and recall rate.

Table 6. Initial setting of parameters

Parameter	Initial value	Parameter significance
Gamma	0.1	Control the post pruning
Maxdepth	12	The depth of the tree
Lambda	2	Regularized term parameter
Subsample	0.7	Proportion of random sampling for training

3.5 Ensemble Learning

After building the 45 binary classifiers, we need to ensemble the results to get the final prediction value. One solution is to take the voting mechanism. For example, when we test a sample, we put it in all the 45 classifiers and count the times it is identified in each of the 10 ranges, then the test sample is deemed to belong to the range with the highest times.

Another solution is that we can train a neural network to ensemble the 45 results. As shown in Fig. 5, the input of neural network is the results of 45 binary classifiers, and n (45)-dimensional feature vector predicted by the learners in first layer is used as the input vector of fully connected neural network classifiers in second layer. According to formula $h = \sqrt{m+n} + \alpha$, we can determine the number of hidden layer nodes in neural network. In the formula, m denotes classifier's output dimension, which is 10 here, and n denotes the input dimension, which is 45, the value of α is generally between 1 to10. The output of this neural network is the label of classification, which is represented as a 10-bit vector, and each bit is 0 or 1. For example, 0100000000 shows the output is the 2^{nd} class. Therefore, we use 45-8-10 three-layer neural network model as the second-layer learner, and the network's activation function is set as sigmod function.

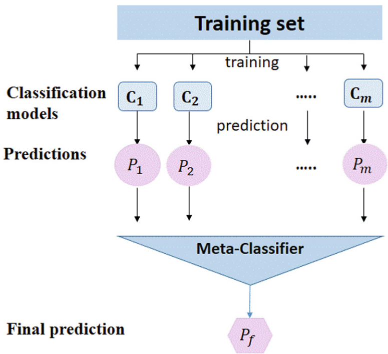

Fig. 5. Process diagram of Stacking

4 Experiments

In this section, we first introduce the evaluation metrics in Sect. 4.1, and then determine the parameters involved in the proposed approach in Sect. 4.2, finally prove its effectiveness and rationality through comparison experiments.

4.1 Evaluation Metrics

We use root mean square error (RMSE) and mean absolute percentage error (MAPE) to evaluate the performance of prediction. The two metrics are defined in formulas (1) and (2), the less RMSE or MAPE, the better performance.

$$RMSE = \sqrt{\frac{\sum(t - \bar{t})^2}{N - 1}} \tag{1}$$

$$MAPE = \frac{1}{N}\sum\frac{|t - \bar{t}|}{t} * 100\% \tag{2}$$

In the formulae, t is the actual duration, $\bar{t}$ is the predicted duration, N is the number of the tests.

For binary classification, we use precision (P) and recall(R) and harmonic mean ($F1$) as the metrics, and their definitions are shown as formulas (3)–(5), where TP is the number of true positive samples, FP is the number of false positive samples, FN is the number of false negative samples. Since P and R are paradoxical sometimes, F1 is used as well, and the larger F1, the better performance.

$$P = TP/(TP + FP) \tag{3}$$

$$R = TP/(TP + FN) \tag{4}$$

$$2/F1 = 1/P + 1/R \tag{5}$$

4.2 Tuning Parameters

Tuning XGBOOST Classifier Parameters. We take the classifiers that classifies the samples into [22–25] or [51–68] as an example to show the way to tune the parameters of each XGBoost classifier. As mentioned above, multiclass classification can be converted into multiple binary classification, and F1 can be used to evaluate the performance of binary classification. Since the samples in different classes are distributed unevenly, using F1 to tune the parameters of models could have a better performance and be more robust. As shown in Fig. 6(a), when *gamma* is 0.2, F1 reaches the maximum value; when *gamma* is between 0.2 and 0.3, F1 starts to decrease; when *gamma* is larger than 0.3, F1 tends to be unchanged. So we can set *gamma* to 0.2.

After determining the value of gamma, we fix it and start to tune parameter *maxdepth*. As shown in Fig. 6(b), when *maxdepth* starts with 7, F1 gradually increases and reaches the highest point at 10, then starts to gradually decrease and finally becomes stable. So we can determine *maxdepth* as 10.

After fixing *gamma* and *maxdepth*, we begin to tune parameter *subsample*. As shown in Fig. 6(c), when *subsample* is between 0.1 and 0.7, F1 is almost rising, and when *subsample* is larger than 0.7, F1 starts to decrease. So we set *subsample* to 0.7.

Tuning Neural Network Parameters. In the process of training neural networks, choosing a proper learning rate *lr* and the number of iterations *te* is very important, which will often affect the prediction ability of whole network. We can randomly select training data, choose different *lr* and *te*, and calculate corresponding MAPE. The results are shown in Fig. 7. We can see that, when learning rate *lr* is too small, for example, *lr*

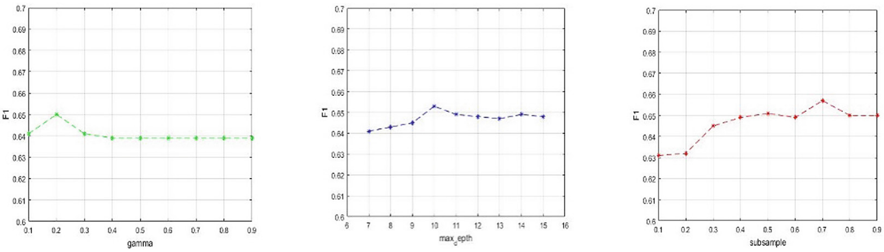

Fig. 6. Tuning parameters in XGBoost

is 0.01 and *te* is 250, or *lr* is too large, for example, *lr* is 1 and *te* is 250, the performance is not so good. When *lr* is 0.5 and *te* is 250, MAPE is the lowest, and the performance of accuracy is the best. So we set *lr* to 0.5 and *te* to 250.

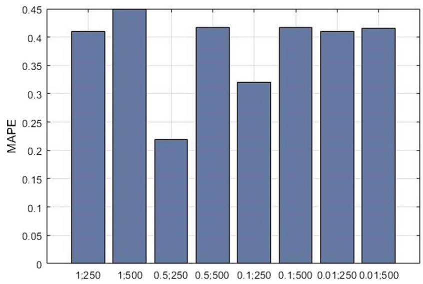

Fig. 7. Tuning parameters in artificial neural network

4.3 Verifying the Effectiveness of the Proposed Solution

Verifying the Effectiveness of the XGBoost Classifiers. In this experiment, we aim to compare our XGBoost classifier with three classical classifiers: random forest, logistic regression and naive Bayes. Figure 8 shows the average MAPE and RMSE of the four models on the whole dataset. It can be seen that the average MAPE of XGBoost is 0.219, and its RMSE is 0.6858. While the MAPE and RMSE values of the other three models are higher than XGBoost, that is to say, their classification performances are worse than XGBoost on our dataset. So we choose XGBoost here.

Verifying the Effectiveness of Converting Multi-class Classification into Binary Classification. In multi-class classification, each class is treated equally. But there are some problems when accuracy is used to evaluate the prediction capability. In our dataset, the data are distributed too unevenly, the accident duration of about 40% training samples falls in the range of [15, 30] minutes. For unbalanced datasets, the prediction accuracy of rare classes is usually more important than that of others. From Fig. 2 we can see, the classes that fall in the range of [0, 15] and [30, +∞] belong to rare classes.

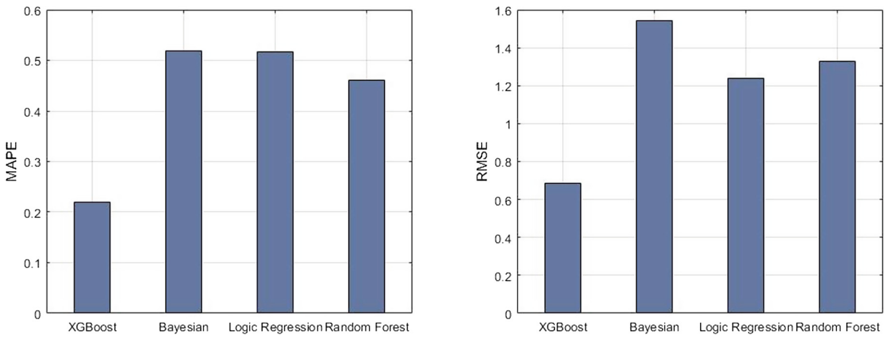

Fig. 8. The comparison of evaluation indexes by different algorithms

In the experiment, we compare the performance of multi-classification and binary classification by using MAPE and RMSE. We divide the duration time into 11 intervals and each interval is 10 min. Figure 9 shows the MAPE and RMSE for each time interval respectively. The red line shows the results for multiclass classification, in which 10 classes are the output in one XGBoost model, and the average value of the predicted class is the predicted duration value; while the black line shows the results for our proposed solution including binary classification and ensemble learning. From Fig. 9 we can see that our proposed solution performs better than multi-classification on the two metrics in each time interval. More importantly, the prediction accuracy of rare classes has been significantly improved.

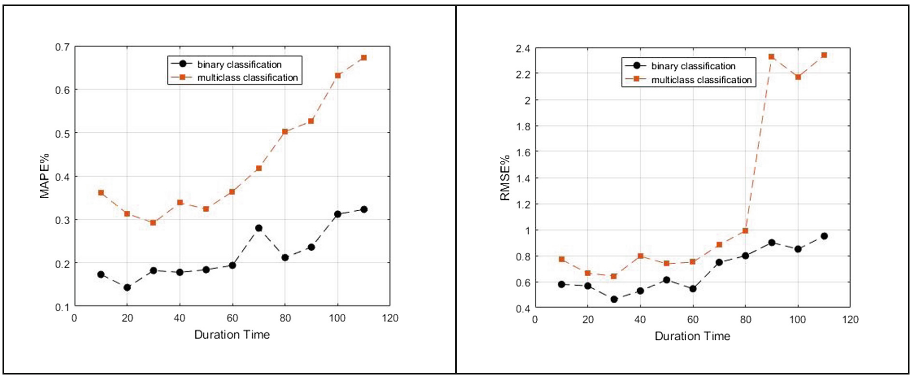

Fig. 9. The comparison between binary classification and multiclass classification

Verifying the Effectiveness of Ensemble Learning. In this experiment, we aim to verify that the neural network is effective for integrating the classification results. We compare our solution with two methods, one is using neural network for prediction directly, and another one is using voting for integrating the classification results. Figure 10 shows the comparison results on MAPE and RMSE for the three solutions.

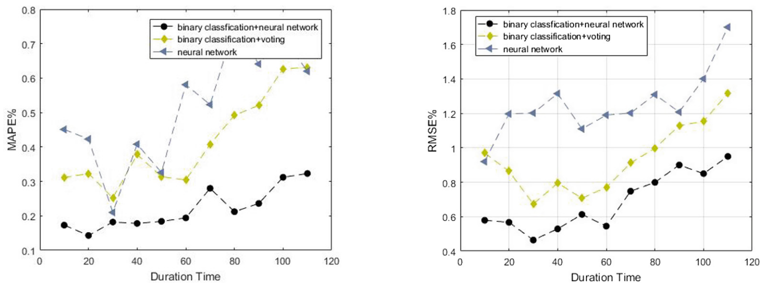

Fig. 10. The comparison between control group and experimental group

From Fig. 10 we can see that the prediction accuracy of our solution is better than the other two methods for each time segment, especially in rare categories, which further indicates the effectiveness of our proposed solution.

5 Conclusions

By using the real accident data in Xiamen, firstly, we deal with the features, such as treatment types of accidents and weather more adequately. Secondly, we convert the multiclass classification problem into multiple binary classification problem based on XGBoost. The conversion can not only reduce the overall prediction error, but also significantly improve the prediction precision of rare categories. Then we use neural network to integrate the prediction results of classifiers in order to reduce the overall prediction error. In the future work, we will study how to deal with the incomplete accident data and predict the reliable accident duration in actual traffic accidents, in addition, we will investigate on predicting traffic flow during accident recovery.

Acknowledgments. The research is supported by National Natural Science Foundation of China (No. 61772560), National Key R&D program of China (No. 2018YFB1003800).

References

1. Giuliano, G.: Incident characteristics, frequency, and duration on a high volume urban freeway. Transp. Res. A: Gen. **23**(5), 387–396 (1989)
2. Golob, T.F., Recker, W.W., Leonard, J.D.: An analysis of the severity and incident duration of truck-involved freeway accidents. Accid. Anal. Prev. **19**(5), 375–395 (1987)
3. Skabardonis, A., Petty, K., Varaiya, P.: Los Angeles I-10 field experiment: incident patterns. J. Transp. Res. Board **1683**, 22–30 (1999)
4. Jones, B., Janssen, L., Mannering, F.: Analysis of the frequency and duration of freeway accidents in Seattle. Accid. Anal. Prev. **23**(4), 239–255 (1991)
5. Nam, D., Mannering, F.: An exploratory hazard-based analysis of highway incident duration. Transp. Res. A: Policy Pract. **34**(2), 85–102 (2000)

6. Chung, Y.: Development of an accident duration prediction model on the Korean freeway systems. Accid. Anal. Prev. **42**(1), 282–289 (2010)
7. Kang, G., Fang, S.-E.: Applying survival analysis approach to traffic incident duration prediction. In: First International Conference on Transportation Information and Safety (ICTIS), Wuhan, China, pp. 1523–1531 (2011)
8. Alkaabi, A.M.S., Dissanayake, D., Bird, R.: Analyzing clearance time of urban traffic accidents in Abu Dhabi, United Arab Emirates, with hazard-based duration modeling method. Transp. Res. Rec. **2229**, 46–54 (2011)
9. Hojati, A.T., Ferreira, L., Washington, S., Charlesa, P.: Hazard based models for freeway traffic incident duration. Accid. Anal. Prev. **52**, 171–181 (2013)
10. Li, R., Guo, M.: Competing risks analysis on traffic accident duration time. J. Adv. Transp. **49**(3), 402–415 (2015)
11. Zhan, C., Gan, A., Hadi, M.: Prediction of lane clearance time of freeway incidents using the M5P tree algorithm. IEEE Trans. Intell. Transp. Syst. **12**(4), 1549–1557 (2011)
12. Yang, H., Shen, L., Xiang, Y., et al.: Freeway incident duration prediction using Bayesian network. In: 2017 4th International Conference on Transportation Information and Safety (ICTIS), pp. 974–980. IEEE (2017)
13. Weng, J., Qiao, W., Qu, X., et al.: Cluster-based lognormal distribution model for accident duration. Transportmetrica **11**(4), 345–363 (2015)
14. Ma, X., Ding, C., Luan, S., et al.: Prioritizing influential factors for freeway incident clearance time prediction using the gradient boosting decision trees method. IEEE Trans. Intell. Transp. Syst. **18**(9), 2303–2310 (2017)
15. Pereira, F.C., Rodrigues, F., Ben-Akiva, M.: Text analysis in incident duration prediction. Transp. Res. Part C **37**(10), 177–192 (2013)
16. Yu, B., Wang, Y.T., Yao, J.B., et al.: A comparison of the performance of ANN and SVM for the prediction of traffic accident duration. Neural Netw. World **26**(3), 271 (2016)
17. Khattak, A.J., Liu, J., Wali, B., et al.: Modeling traffic incident duration using quantile regression. Transp. Res. Rec. J. Transp. Res. Board **2554**, 139–148 (2016)
18. He, Q., Kamarianakis, Y., Jintanakul, K., et al.: Incident duration prediction with hybrid tree-based quantile regression. In: Ukkusuri, S., Ozbay, K. (eds.) Advances in Dynamic Network Modeling in Complex Transportation Systems, pp. 287–305. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-6243-9_12
19. Peeta, S., Ramos, J.L., Gedela, S.: Providing real-time traffic advisory and route guidance to manage Borman incidents on-line using the hoosier helper program. Joint Transportation Research Program, Indiana Dept. Transp., Purdue Univ., West Lafayette, IN, USA. Technical report, 15 2000. http://docs.lib.purdue.edu/jtrp/175/
20. Vlahogianni, E.I., Karlaftis, M.G.: Fuzzy-entropy neural network freeway incident duration modeling with single and competing uncertainties. Comput. Aided Civil Infrastruct. Eng. **28**(6), 420–433 (2013)
21. Lee, Y., Wei, C.-H.: A computerized feature selection method using genetic algorithms to forecast freeway accident duration times. Comput. Aided Civil Infrastruct. Eng. **25**, 132–148 (2010)
22. Park, H., Haghani, A., Zhang, X.: Interpretation of Bayesian neural networks for predicting the duration of detected incidents. J. Intell. Transp. Syst. **20**(4), 385–400 (2016)