ABSTRACT

Flight delay is inevitable and it plays an important role in both profits and loss of the airlines. An accurate estimation of flight delay is critical for airlines because the results can be applied to increase customer satisfaction and incomes of airline agencies. There have been many researches on modeling and predicting flight delays, where most of them have been trying to predict the delay through extracting important characteristics and most related features. However, most of the proposed methods are not accurate enough because of massive volume data, dependencies and extreme number of parameters. This paper proposes a model for predicting flight delay based on Deep Learning (DL). DL is one of the newest methods employed in solving problems with high level of complexity and massive amount of data. Moreover, DL is capable to automatically extract the important features from data. Furthermore, due to the fact that most of flight delay data are noisy, a technique based on stack denoising autoecoder is designed and added to the proposed model. Also, Levenberg-Marquart algorithm is applied to find weight and bias proper values, and finally the outpu has been optimized to produce high accurate results. In order to study effect of stack denoising autoencoder and LM algorithm on the model structure, two other structures are also designed. First structure is based on autoencoder and LM algorithm (SAE-LM), and the second structure is based on denoising autoencoder only (SDA). To investigate the three models, we apply the proposed model on U.S flight dataset that it is imbalanced dataset. In order to create balance dataset, undersampling method are used. We measured precision, accuracy, sensitivity, recall and F-measure of the three models on two cases. Accuracy of the proposed prediction model analyzed and compared to previous prediction method. results of three models on both imbalanced and balanced dataset shows that precision, accuracy, sensitivity, recall and F-measure of SDA-LM model with imbalanced and balanced dataset is improvement than SAE-LM and SDA models. The results also show that accuracy of the proposed model in forecasting flight delay on imbalanced and balanced dataset respectively has greater than previous model called RNN.

INTRODUCTION

As the air travels have a significant role in economy of agencies and airports, it is necessary for them to increase quality of their services. One of the important modern life challenges of airports and airline agencies is flight delay. In addition, delay in flight makes passengers concerned and this matter causes extra expenses for the agency and the airport itself. In 2007, U.S government had endured 31–40 billion dollar downsides due to flight delays . In 2017, 76% of the flights arrived on time. Where, in comparison to 2016, the percentage of on time flights decreased by 8.5%. As some of the reasons of flight delays the following can be mentioned: security, weather conditions, shortage of parts and technical and airplane equipment issues and flight crew delays. Delay in flight is inevitable , which has too much negative economic effects on passengers,agencies and airport. Furthermore, delay can damage the environment through fuel consumption increment and also leads to emission of pollutant gases. In addition, the delay affects the trade, because goods' transport is highly dependent to customer trust, which can increase or decrease the ticket sales, so that on time flight leads to customer confidence.

So that, flight prediction can cause a skillful decision and operation for agencies and airports, and also a good passenger information system can relatively satisfy the customer.

According to abundant and diversity of reasons for flight delays, We are faced with a massive amount of data which is not possible to be processed through previous methods of data analysis like classification [1], or the decision tree and machine learning based method to process this volume of data are not proper, because characteristics of older intelligent system has been designed by human and usually were personalized, also people rarely perceive some features and usually neglect these matters. On the other hand, in older learning process, as the number of categories available for classification increases, the level of difficulty increases and extraction of important and

effective features becomes relatively impossible. Due to complexity and effect of parameters on each other, the problem of flight delay prediction is considered as NP Complete. Furthermore, the problem essentially is accompanied by oscillation and also these are considered as non linear problems . On the other hand, applied data includes noise and error that should be handled to cope with the problem .

There have been too many studies in this area. For example, older Regression method has been used to compute delay propagation. For this model, the destination delay is highly dependent to arrival flights and the effective factors include; day, time, airport capacity and some factors are related to passenger loads. In addition, as the problem neglects the weather conditions, this model shows inefficiency in U.S.A but it is suitable for Europe. Where, only 1–4% of the Europe flights delayed due to weather condition, this value for U.S.A is between 70 and 75% in an intelligent neural network has been designed which estimated the destination delay for actual applications in controlling traffic progress. This model employs factors of airport type, airplane type, date,time, flight path, flight frequency for network training and non-linear and linear for data analysis. As it is difficult to interpret neural network parameters, the way factor behavior and most important verification of the most important factors in flight is extremely difficult. Furthermore, older intelligent algorithm usually uses shadow learning models to solve conditions with a big data in complicated classifications. However, results of this analysis are very different with respect to ideal condition. Although model design can have a good or bad situation, response is highly dependent to experience and even hap penstance and this procedure requires too much time. Therefore, traditional simulation and modelling techniques is not suitable or even efficient for such problems. There is an ongoing subject of study which solves this problem and this paper also has tried to use that subject in modelling.

One of the newest modern methods in solving such extended and complicated accompanied by bulky data that has been concerned by many scientists is deep neural networks. The design of learning technology is taken from human neural network learning is a branch of machine learning and collection of algorithms that trying to model such high-level abstract contents through application learning in different layers and levels. Therefore, this subject enables the deep learning to process a bulky data volume in complicated data classification . Moreover, this structure is proper for extracting some the characteristics, so that learning is capable to extract maximum number of possible characteristics . Layered network structure and capability of computation for each data scale has led to progressing application of techniques. This networks have different types including convolutional Neural Network , Autoencoders , Restricted Boltzmann Machine and Sparse coding based method that each of them is applicated for specified problem.

One of the recently presented works in solving problem employs the recurrent Deep Neural Network and its results has a high accuracy in flight delay prediction. How- ever, this model has drawbacks of overfitting, that researchers have solved that through typical data dropout technique for each step of repeated training procedure. Moreover, application of this method decreases the computation time and memory space during the training.

The next drawback is the noise of input data. However, the researcher neglects the noise during prediction.

This paper tries to represent a model based on deep learning, which considers the effective factors in the delay. Moreover, noisy data requires utility of stack denoising autoencoder (SDA) in designing the model. Afterwards, optimized structure of the flight delay forecasting model with Levenberg Marquart (LM) algorithm. In addition, in this paper by developing a deep learning-based model, the accuracy of flight delay predictions can be increased.

Finally, we review previous work related to our topic in “Literature review” section, a complete description of research process and also the holistic structure of the designed model is represented in the third section.

Fourth section evaluates the determined results from the previous methods. Fifth section presents a conclusion and an overall view about the study.

LITERATURE REVIEW

Nowadays, service quality plays an important role in attracting customers. Among these, air travels have their special customers and the most important matter in these travels is the flight time, on time arrival at destination for passengers such those who have an important meeting, that has been leading to high expenses for the passenger until get to their destination on time. Flight delay has negative economic effects on the passenger, agencies and airports. Therefore, any reduction of these effect requires decreasing postponed flight price, so that prediction or estimation has a great significance and numerous studies has been to dedicated this subject. Correspondingly, all the scientists have tried to design a model that understands effective factors and computes effect of each factor and their relation. Overall, the prediction methods are classified into five groups including Statistical Methods, Probability methods, network-based methods, operational methods and machine learning methods. In one of the best studies that has been performed based on statistics delay time has been considered to be reduced. Their study has investigated important factors before fly and those which occur on the ground. In the next step, it has predicted the delay at destination based on factors that occur in the vicinity of arrival time at destination. Eventually, results have shown that whenever, the delay is correctly predicted, pas- senger disaffection and fuel consumption decrease and consequently number of flight increases. Moreover, it is possible to increase the agencies' benefits through reducing number of passengers who wrongly selected their routs or specifying the probabilities for some flights and optimizing delay time prediction. Another prominent investigation based on Probability has been done and the author believes that huge storm in U.S.A has highly affected the flight delay. This study has been devoted to predict delay based on mathematical calculations and through con- sidering delay time duration of the flights that had been engaged to storm in the same day. Metrological reports have shown the effect of storm one hour before and after event cause ephemeral climate at the region. In the next step, Monte Carlo simulation has been used to estimate the airport runway capacity, so that traffic of each runway would have been estimated. As the research has employed only one factor, the model has not enough accuracy, but it is possible to increase region air capacity path structure. A model has been presented in , which is one of the best network-based mod- els. The researchers have presented a model based on Bayesian and Gaussian mixture model- expectation maximization (GMM-EM) algorithm to predict and analyze the factors affecting the flight delay in Brazil for several point along the path. At the first stage of model, the degree of effectiveness for each factor is specified and then it has speci- fied investigated whether the delay had happened in a greater domain or no. the next delay probability is computed using GMM-EM and EM algorithm which are speci- fied based on similarity. The result has shown that it is possible to predict the probability of delay in higher levels through specifying low level factors. Moreover, GMM EM similarity function has more values rather than EM algorithm in each step, so that the results would have been converged sooner. In addition, the model accuracy is increased, so that the prediction is more trustable. One of the best studies in the area of operating method has been presented. Studied the effects of capacity and damage on different levels of delay in American airports. Other simulations focus on stability and reliability during the delay and its propaga- tion. For instance, inthe problems of congestion were studied. Then, a queue-based model was presented for analyzing delay propagation in consecutive flights in the Los Angeles airport.

One of the best studies in the area of machine learning method has been presented by a model which applicate machine learning techniques to investigate delay in arrival flights. This research firstly has extracted important characteristics and then has been used for both neural networks and deep believe network through arbitrary sam- ples to train the model. The model utilizes Memento and Resilient Back Optimized Propagation that the Resilient back propagations quicker than back propagation and as a result the model training and consequently has been increased. Deep believe networks is based on a few Boltzmann machine that each com- munication layer receives data from the previous layer and in each step a Boltzmann machine is added to Believe Network overall,

training time reduced using param- eter adjustment operation and learning rate, false classified error rate. As each layer has convergence at the output, training speed is reduced and the gradient approaches zero. In addition, a relatively small data base is used for the model because of limited sys- tem capacity. So that this problem leads to a noticeable reduction in prediction precision whenever it is not at database.A model has been presented which was one of the machine learning method. The researcher has presented a modelbased on support vector regressor (SVR) algorithm to predict flight delay in U.S.A airports. Due to the large amount of data, the data was grouped and sampled by month. At the first stage for categorical variables, cat-boost used the ordered boosting method. Because cat-boost itself had the effect of scoring fea- tures, it was possible to select parameters that were more important to the model when the threshold was unknown, so cat-boost was used to evaluate the features of each fea- ture to select features, and finally 15 features were selected to build a training model. Then has beenused several common regression prediction algorithms to predict the delay at the same time for the round-trip flight between John F. KennedyInternational

Airport and O'Hare International Airport.Finally, the specific delay time was predicted. The results have shown SVR has the best prediction result for the flight delay time with the best accuracy value was 80.44%. Also, the time characteristics had a large impact on the mode performance. The air time and flight distance would also have a greater impact on on-time perfor- mance of specific flight; Different carriers and specific aircraft would also have a slight influence of on time performance. Accuracy of this model is low because detailed weather and aircraft data could not be collected. A research analyzes flight information of U.S domestic flight operated by Ameri- can Airlines, covering top 5 busiest airports of US and predicting possible arrival delayof the flight using Data Mining and Machine Learning Approaches. Due to the imbal- anced data, Over-Sampling technique, Randomized SMOTE was applied for Data Bal- ancing. The Gradient Boosting Classifier Model was deployed by training and then Grid Search on Gradient Boosting Classifier Model on flight data, caused hyper-parameter tuned and achieving amaximum accuracy of 85.73%. Result showed that deleting some features affected the value of accuracy and reduced it. A group of researchers [127] have designed 5 models to predict flight delay based on machine learning models such as Logistic Regression, Decision Tree Regression, Bayesian Ridge, Regression and Gradient Boosting Regression. They collected data from Bureau of Transportation, U.S. Statistics of all the domestic flights taken in 2015 and predicted whether the arrival of a particular flight would be delayed or not. The metrics to evaluate the performance of the models were: Mean squared error (MSE), Mean Absolute Error (MAE), Explained Variance Score, Median Absolute Error and R2 Score. Due to the used of imbalanced data sets, the amount of calcu- lated error was high. Based on the results, Random Forest Regressor was observed as the best model in prediction of arrival and departure delay. One of the newest studies in the area of machine learning method has been pre- sented by a model which applicate supervised learning methods to aggregate flight departure delays in china airports . The expected departure delays in airports was selected as the prediction target while four popular supervised learning methods: multiple linear regression, support vector machine, extremely randomized trees and LightGBM were investigated to improve the predictability and accuracy of the model. Of special note was that the model performances with local weather characteristics was not as good as those without Sustainability meteorological data. They measured accuracy, MSE and MAE for evaluating 4 methods and result has shown LightGBM model could provide the best result, giving 0.86 accuracy. A group of Researchers [129] designed a framework to integrate multiple data sources to predict the departure delay of a scheduled flight and discuss the details ofthe data pipeline. They were the first group, to take advantage of airport situational awareness map, which was defined as airport traffic complexity (ATC), and combined the proposed ATC factors with weather conditions and flight information. In the first stage, historical data, weather condition data, and tarmac aircraft and vehicles GPS data were collected from different data sources. After that the feature extraction stage, was applied principal component analysis to weather data, and were extracted ATC features from tarmac aircraft and vehicle trajectory data, also utilize the historical

scheduling table data. It seems that except for the extracted featuresmore potentially useful features can be explored from the airport situational awareness map. Then in the modelling stage, multiple datasets were combined and various data combinations were used to train a regressor model that could be used for pre- dicting departure delay time. Authors selected four popular regressors from different families (linear regression, SVR, ANN, and regression trees) to show the robustness of their proposed approach to different regressors. Finally, has been evaluated the prediction results using Root Mean Square Error (RMSE) to measure the performance of flight delay time prediction using different models and different combinations of data sources. Result has shown LightGBM regressor outperforms other conventional regressors with exten- sive experiments on a large real-world dataset. Although Other works which have been done in recent years is not in the scope of this article, it is still related to the topic in a way that contributes to the progress of this article, so here we have included studies that employed a support vector machine (SVM) model to explore the non-linear relationship between flight delay outcomes and another model that explored a broader spectrum of factors. This model could potentially affect the flight delay and proposed a gradient boosting deci- sion tree (GBDT) based models for generalized flight delay prediction. The presented techniques are faced to limitations, because these techniques cannot resist against the massive data volume and complicated computations. For example, in some of these studies, the model is designed based on the specifications andconditions of a special country . Some other consider weather conditions in their prediction , next group has considered the special situation like or destination.

CONCLUSION

Predicting flight delays is on interesting research topic and required many attentions these years. Majority of research have tried to develop and expand their models in order to increase the precision and accuracy of predicting flight delays. Since the issue of flights being on-time is very important, flight delay prediction models must have high precision and accuracy. In this study, we proposed a novel optimized forecasting model based on deep learning which engages LM algorithm. Afterwards, two other structures are created to study and validate the positive effect of denoising autoencoder and LM algorithm, which one has deleted denoising autoencoder and the other has omitted LM algorithm. Moreover, we have imbalanced dataset which should be balanced. We used undersampling and upsampling technique to balance the data. However, results show that upsampling leads to overfitting. Therefore, under sampling is used for balancing. Comparing the three models for two of imbalanced and balanced datasets shows that accuracy of SDA-LM model with imbalanced dataset respectively is greater by 8.2 and 11.3% Than SAE-LM and SDA models. On the other hand, these values for balanced datasets are respectively as 10.4 and 7.3%. Therefore, using stack denoising autoencoderand LM algorithm in optimizing the results, and also balancing the dataset, has positive effect on delay forecasting and leads to increment in accuracy and precision of SDA-LM model with imbalanced dataset is greater by 6.1 and 5.4% than SAE-LM and SDA models. Whereas, the accuracy of the SDA-LM model with balanced dataset is greater by10% than SAE-LM and SDA models and the amount of precision is the same for all three models with balance dataset. At the next stage, the model has been evaluated and computed for subjects of discarding with a standard deviation for all evaluation parameters during 30 times of model run. The results, shows that standard deviation for all balanced evaluation parameters is lower than the imbalanced form. Therefore, data balance leads to lower standard deviation. amount of model standard deviation for imbalanced dataset is 0.045 while this value is reported 0.21 for balanced dataset which is a small value and means that scattering results are low and close to average. Finally, we compared the accuracy of the proposed Model against SAE-LM, SDA and RNN models. Using our experimental results, we show that accuracy of the model on imbalanced dataset is 92.1% and for balanced dataset is 96.2%, which is respectively greater by 4.1 and 8.2% respectively. Therefore by proposed

model has greater accuracy in forecasting flight delay compared to previous model called RNN . The next step would be to apply this technique on other data sets or on other sampling data and investigate the accuracy. Table 10 Standard deviation calculated for all evaluation parameters for the three structures using the balanced dataset.