International Journal of Engineering Research and Technology. ISSN 0974-3154 Vol.13, No.4 (2020), pp. 814-822
 https://dx.doi.org/10.37624/IJERT/13.4.2020.814-822

A Hybrid Artificial Intelligence Approach to Predict Flight Delay

Bashayer Alharbi[1] and Master Prince[2,*]

[1]MS Student, Department of Computer Science, College of Computer, Qassim University, Qassim, Saudi Arabia.

[2]Assistant Professor, Department of Computer Science, College of Computer, Qassim University, Qassim, Saudi Arabia.

**Corresponding author: Master Prince*

[1]ORCID: 0000-0002-9092-3364, [2]ORCID: 0000-0002-2703-4580

Abstract

Commercial airlines and the passengers suffer from flight delay. Flight delay causes huge loss for the airlines and unsatisfied passengers. The researchers attempt to solve this problem through prediction extensively by machine learning approach and data mining tools. Accurate and robust performance is still to get through existing models. Our proposed hybrid approach is intended to use the power of machine learning as data mining tool and to predict the delay using classification algorithm of deep learning. An extensive evaluation of the proposed method is carried out by comparing the performance by using two data sets: one is local and the other is benchmark from Kaggle to obtain the best performing classifier. Three predictive models were applied on the datasets: logistic regression, decision tree and the proposed approach. The result shows that the proposed method performed well as comparing to the existing state-of-the art.

Keywords: Deep Learning, Delay Prediction, Flight Delay, Machine Learning, Classification.

I. INTRODUCTION

Nowadays the importance of air travels is increasing as it saves time and money. Airline companies take care about scheduling the flights for the stratification of the passengers. It is the most important issue to be loyal to the airline company. One of the problems that passengers suffer from is the delay; flight delay is unappreciable from passengers because they may have another transit flight, appointment in hospital or other important businesses. For that the announcement about the delay have to be earlier than usual because if the passengers have an urgent issue, he/she can take an immediate action either going by car or train and they appreciate the company that give them a value. Priority Pass Consumer Reports surveyed 2,000 air travelers in June 2010 and asked them to rate the 12 most annoying things about air travel on a scale of 1 (least annoying) to 10 (most annoying). Flight delay was 7th on the list with a score of 6.8 on their scale [1]. In addition, it saves huge losses for the airline company [2], [3]. Predicting flight delay is one of the biggest problems for major airline. Nearly 30% of the jet operator flights for United States airlines were delayed in 2000, and almost 3.5% of these flights were cancelled [4]. Therefore, to optimize the performance of current airports, predicting the probable delay of a given flight can be very useful as the unused airspace and airport capacity can be assigned to a different flight [5],[6].

Several models have been developed to solve this problem based on probability, statistics, and operations research [5],[6],[7],[8]. Predication is considered one of the most important machine learning techniques that helps to predict the delay before it happens. The flight delay prediction is considered as an important research topic due to the dynamics of the flight operating process[1]. The advanced machine learning techniques and associated data mining tools can help to understand and predict several complex phenomena, the approach is used in enabling businesses and research collaborations alike to make informed decisions flight delay prediction [9]. In addition, every year approximately 20% of airline flights are delayed or canceled mainly due to bad weather, carrier equipment, security or technical airport problems [9], [2]. These delays result in significant cost to both airlines and passengers [9], [2], [1]. The costs related to the airline company when flight connections are missed, or flight crews and aircraft need to be reallocated due to maintenance problems or crew duty time limits [1], this reallocation reduces the delay propagation effects [3]. Flight delays effect the successive flight when the delayed flight aircraft and/or flight crew causes a delay for successive flights scheduled that day for the aircraft, airport and crew [1], [3].

Our motivation in terms of this problem is to apply both of the above tools-machine learning and deep learning-in order to predict the flight delay more precisely. In particular, this paper proposes a hybrid approach which mines the data by Principal Component Analysis (PCA) as machine learning tool and deep learning tool for prediction of the flight delay as it has been proved its potential.

The methodology we applied for this involves three steps: 1. Pre-process the dataset 2. Exploratory Data Analysis and 3. Apply the predictive models Decision Tree (DT), logistic Regression (LR) and the hybrid approach. Finally the results are evaluated and it shows better performance for the hybrid approach than the other models.

The paper has been structured as follows: Section. 2 provides a brief review of the literature to understand the problem and to build the model. Section 3 describes the proposed methodology for the model. In Section 4, performance of the classifiers are compared and concluded in terms of training accuracy, testing accuracy, precision, recall, ROC, F1score and confusion matrix. Finally, itis concluded with future work in Sect. 5.

II. RELATED WORKS

Extensive research has been done on flight delay prediction .Airline operations are highly complex processes that are intended to regulate many expensive, tightly constrained and interdependent resources such as: crews, aircraft, airports and maintenance facilities. Many studies have been carried out on airline planning problems but only a few have been performed on the characteristics of airline delays and the prediction of delay statistics. Delays occur when an event takes place later than the time at which it is planned, scheduled, or expected to happen [10].

In a study carried out by N.Kuhn and N. Jamadagniy [11] applied DT, LR and neural networks classifiers on a public available in Kaggle.com for United States domestic air-traffic for the year 2015. Their goal is to predict if a given flight's arrival will be delayed or not. The results showed test accuracy of approximately 91% for all three classifiers.

However, considering the nature of this problem using the artificial neural network (ANN) techniques can be beneficial as artificial neural networks are very practical in solving nonlinear problems. Also, due to their supervised learning capability they can easily adapt to the dynamics of air traffic capacity and demand [12].

Rahul Nigam and G.K [13] used a machine learning technique called LR which is a supervised learning method to predict delay in departure times of aircraft. They implement the algorithm on Microsoft Azure Learning Studio platform which is an Integrated Development Environment for utilizing machine learning for training and testing the model on the cloud. The datasets that were used are flight details dataset, weather dataset and airport dataset the period April-October 2013.They preprocessed and joined weather data such as temperature, humidity, precipitation, dew point along with the airport data to derive more accurate predictions as well as find out the effect. The model was trained on 70% of data and left 30% data to test the algorithm. The method was able to achieve about 80 percent accuracy in predicting whether a given aircraft would be delayed or not based on the training using past data.

S. Manna, S. Biswas, R. Kundu, S. Rakshit, P. Gupta and S. Barman [14] used a statistical approach that consider popular in machine learning to predict flight delay using gradient boosted decision tree. This method showed a great accuracy in modeling sequential data as compared to other methods because it can prove to be quite effective in handling regression tasks. With the help of this model, day-to-day sequences of the departure and arrival flight delays of an individual airport can be predicted efficiently. It has been implemented on the passenger flight on-time performance data taken from U.S. Department of Transportation to predict efficiently the arrival and departure delays in flights of an individual airport. The dataset used contained flight delay data for the period April-October 2013. It included all the incoming and outgoing flights from 70 busiest airports in the United States. They used R, SQL, and Python within Microsoft Azure Machine Learning Studio after pre-processing the data.

M. Baluch, T. Bergstra and M. El-Hajj [15] applied data mining approach on data from the Bureau of Transportation Statistics (BTS) .They used 4 years of data from July 2012 to July 2016 via data mining software. Their goal to predict the delays to aid consumers in the knowledge of the best and worst ways to travel the country and used the useful information to increase revenue, cuts, costs or both. Their results suggested that the airport size does not have an impact on flight delays. They found some carriers have much longer delays than others. In addition, that despite what the pilot might say, late departing flights typically do not make up for the lost time in the air. They discovered a strong correlation between month, travel day and departure delay as well as finding which day and month is best to travel on. Finally, they found that the longer the late arrival delay is, the greater amount of time between it and the departure delay.

V. Venkatesh, A. Arya, P. Agarwal ,S. Lakshmi and S. Balana [16] used the machine learning based classification models ANN and Deep Belief Networks (DBN) for delay prediction considering input parameters ranging from distance to their corresponding weather details to make a decision of whether certain flight is delayed or not. They implemented the work on real world flight big dataset from Kaggle.com. The accuracy results were 92% for neural networks and 77% of Deep Nets. That confirmed the effectiveness of using neural nets and similar deep architectures to categorize flight delay or no-delay. These models helped airline administrators besides the passengers who can rearrange their schedules and arrange accommodation.

Y. J. Kim, S. Choi, S. Briceno and D. Mavris [17] applied deep learning approach for air traffic delay prediction. The experiment applied on database from U.S. Department of Transportation for 10 airports through different parameter settings to each airport. Their model based on Recurrent Neural Networks (RNN) that has showed great accuracy in modeling sequential data of the departure and arrival flight delays.

L.Moreira, C. Dantas, L.Oliveira , J. Soares and E.Ogasawara [18] mentioned that prediction of delays is fundamental to mitigate their occurrence and optimize the decision-making process of an air transport system. They focused on different preprocessing method such as data integration & cleaning, data transformation, data reduction and data balancing. In data reductions they used different method like: absolute minimum shrinkage, information gain, attribute selection based on correlation, PCA and Sampling for the development of machine-learning flight delay classification models to be higher sensitivity to the occurrences of flight delays. They used supervised models for classification: neural networks, k-nearest neighbors (kNN), support vector machine (SVM), naive bayes classifier and random forest. Implementation is done on data provided by the Brazilian National Civil Aviation Agency (ANAC) between 2009 and 2015. Their results indicated the models that applied the balancing techniques performed much better in predicting the occurrence of delays, getting about 60% of hits.

N. Chakrabarty, T. Kundu, S. Dandapat, A. Sarkar and D. K. Kole [2] applied data mining and four supervised machine learning algorithms: random forest, SVM, gradient boosting classifier (GBC) and kNN to predict arrival delay of the flights. Data has been collected from BTS United States Department of Transportation and it included all the flights operated by American Airlines connecting the top five busiest airports of United States located in Atlanta, Los Angeles, Chicago, Dallas/Fort Worth, and New York, in the years 2015 and 2016. All the algorithms were used to build the predictive models and compared to each other to accurately find out whether a given flight will be delayed more than 15 min or not. The results showed that the gradient boosting classifier performs the best with a testing accuracy of 79.7% considering the four most important factors in deciding the best predictive model which are testing and training accuracy, nature of fit, number of false negatives present in the confusion matrix and area under the receiver operating characteristic (AUROC). This predictive model saved huge losses for the commercial airlines and offered on-time quality service to their passengers at industry competitive airfares.

In summary, the earlier methods are mostly based on machine learning, data mining or deep learning alone. In our proposed work therefore, PCA is used to minimize dimensions of different variables in the dataset[19]. PCA is used to handle high dimensionality and avoid issues like over-fitting in high dimensional space. It is used to reduce the number of variables in the dataset by extracting important one from a large pool. It reduces the dimensions with the aim of retaining as much information as possible. It combines highly correlated variables together to form a smaller number of an artificial set of variables (principal components) that account for most variance in the data[20].

The multilayer perceptron (MLP) is a special kind of ANN. MLP has been chosen because of its well-known learning and generalization abilities, which is necessary for dealing with imprecision in input patterns [21] and it is applied to deal with the prediction.

A. Our Contribution:

To the best of our knowledge, this is the first time that MLP and PCA have been used together for flight delay prediction.

B. Model Evaluation:

Receiver operating characteristic (ROC) curve is a plot of true positive rate (TPR) versus false positive rate (FPR) for a set of threshold τ, where TPR is known as the sensitivity performance metric. Then TPR and FPR are computed as follows as shown in Equations 1 and 2, respectively [2]:

$$\text{TPR} = \frac{TP}{(TP+FN)} \approx p(\hat{y} = 1|y = 1) \quad (1)$$

$$\text{FPR} = 1 - \text{sensitivity} = \frac{FP}{(TN+FP)} \approx p(\hat{y} = 1|y = 0) \quad (2)$$

The ideal point on the ROC curve is (0, 1) that includes all the positive examples which are classified correctly and no negative examples are misclassified as positive [2].

- **Accuracy**

Accuracy is one metric for evaluating classification models. Informally, accuracy is the fraction of the right predictions[22].Formally accuracy has the following definition:

$$\text{Accuracy}= \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

For binary classification accuracy can also be calculated in terms of positives and negatives as follows in Equ 3. :

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

Where TP = True Positives, TN = True Negatives, FP = False Positives and FN = False Negatives.

- **Confusion Matrix**

It is a performance measurement for classification problem where output can be two or more classes. It is a table with 4 different combinations of predicted and actual values as seen in Table 1 [23].

Table 1: Confusion Matrix

	Actual Values		
Predicted Values		positive(1)	negative(0)
	positive(1)	TP	FP
	negative(0)	FN	TN

- **Classification Report[precision, recall , f1-score, support]** [23] **:**

Recall:

Out of all the positive classes how much we predicted correctly. It should be as high as possible. Recall is the ratio of correctly predicted positive observations to the all observations in actual positive classes as in Equation 4.

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

Precision:

Precision is the ratio of correctly predicted positive observations to the total predicted positive observations as in Equation 5.

$$\text{Precision}=\frac{TP}{TP+FP} \quad (5)$$

F1-score:

F1 score is the weighted average of precision and recall. Therefore, this score takes both false positives and false

International Journal of Engineering Research and Technology. ISSN 0974-3154 Vol.13, No.4 (2020), pp. 814-822
© International Research Publication House. https://dx.doi.org/10.37624/IJERT/13.4.2020.814-822

negatives into account as in Equation 6.

$$\text{F1-score}=\frac{2*recall*precision}{recall+precision} \quad (6)$$

Support:

The support is the number of occurrences of each class in y_true.

III. METHODOLOGY

In the methodology section describes the flow of work for the algorithms that were applied .The main objective is to build a model to predict the delay of the flights that meets the state of art. neural network gave an amazing performance in terms of flight delay prediction [16],[18],[24],[25],[26],[27],[11], especially RNN,LTSM[17] and DBN [29].

The hybrid approach was compared by the most widely used models LR [9],[13] and DT [14],[30],[31],[11],[28] . In addition, it was compared with a benchmark from Kaggle.com to prove its authenticity. The process of predicting the arrival delay of flights operated in Saudi Arabia has been depicted in Fig. 1.

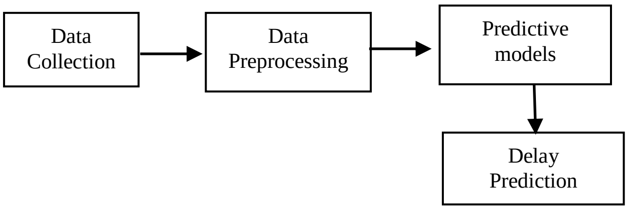

Figure 1: Methodology

3.1. Data Collection

The data used from two different sources GACA Saudi Arabia[32],and Kaggle.com[33] below description Table 2 for both datasets.

Table 2: Description of the Dataset

Name	Size kb	Number of columns	Number of rows	Year	Source
Flight Data	1989	14	15,668	2017	GACA
Airlines Delay	242,152	31	1936758	2008	Kaggle

3.2 Data Processing

The GACA dataset collected include 15,668 instances of flights connecting major airports through flights operated by different Airlines companies in Saudi Arabia. Of the total 15,668 instances, 187 instances had data missing in them, and hence, 15 instance were removed the rest were replaced with a value. The final dataset included 15,652 instances with specific selected attributes .The Kaggle benchmark contains 1936758 instances and 31 features, the unnecessary features were removed from the data.

3.3 Exploratory Data Analysis (EDA)

Analyzing the data sets lead to summarize their main characteristics. This section has two part: first for GACA dataset, the other for Kaggle dataset.

3.3.1 GACA dataset:

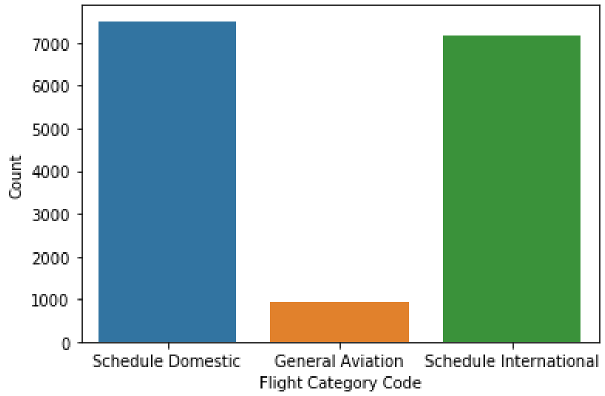

Figure 2: Flight Categories

- In Figure 2, flight categories distribution contains three types: scheduled domestic, general aviation and scheduled international.

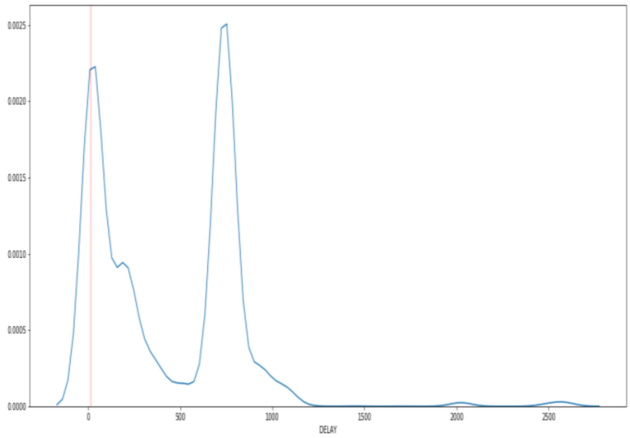

Figure 3: Highlighted Distribution of Delay Column

In Figure 3, highlighting the predefined threshold (15) for delay so we can visually identify the distribution. For further analysis of distributions of different columns, we introduce a temporary column named ‘DELAYED’ which will be ‘1’ for rows where delay value is above threshold (15) and ‘0’ otherwise. Now the plots would be more revealing for us as they would be highlighting different distributions individually for both delayed and not delayed flights.

International Journal of Engineering Research and Technology. ISSN 0974-3154 Vol.13, No.4 (2020), pp. 814-822
 https://dx.doi.org/10.37624/IJERT/13.4.2020.814-822

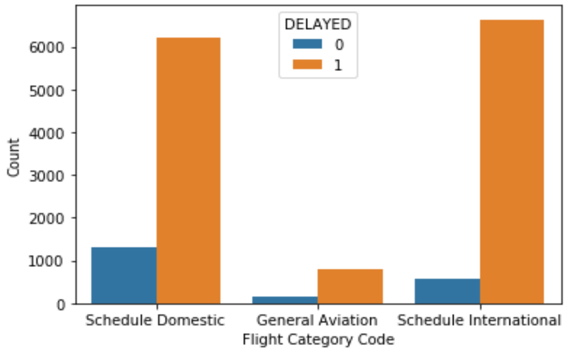

Figure 4: Distribution of Flight Categories for Delay

- In Figure 4, the distribution of flight categories in GACA dataset considering delay and the most delayed category is scheduled international.

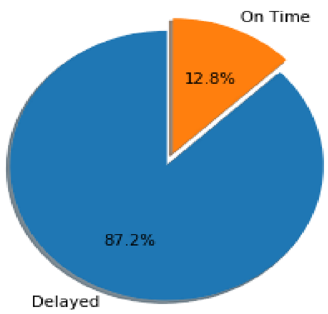

Figure 5: Distribution of Delay in the GACA Dataset

Figure 5, shows that our rows are divided into two categories (delayed and not delayed represented as 1 and 0 respectively). The delayed flights were 87.2% and 12.8% on time flights in GACA dataset.

3.3.2 Kaggle dataset:

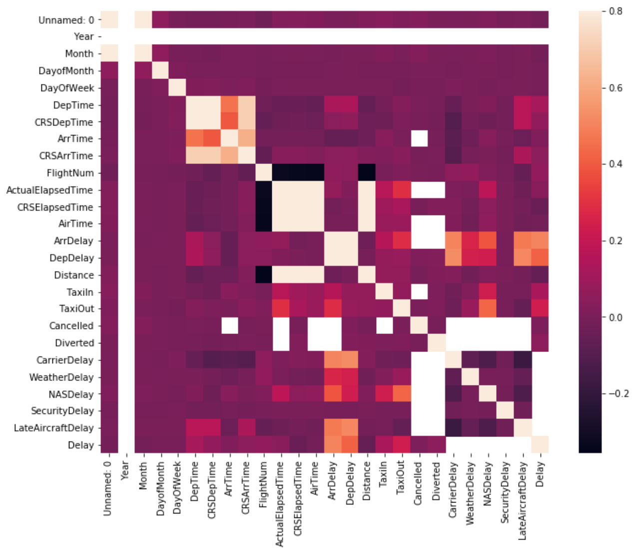

Figure 6: Correlation Between Features in Kaggle Dataset

Figure 6, showed the correlation between features and how they are related with each other in Kaggle dataset. Features with high correlation are more linearly dependent and hence have almost the same effect on the dependent variable. So when two features have high correlation we can drop one of the two features [34].

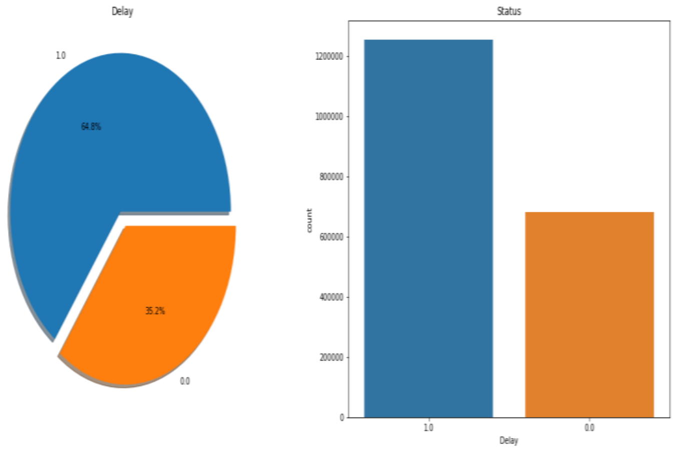

Figure7: Delay in Kaggle Dataset

In Figure 7, the delay distribution shown in Kaggle dataset showed the delay were 64.8% and the flights that weren't delayed were 35.2%precentage.

3.4 Classification Model

Here following the rules of the BTS flights arriving at the airport gate within 15 minutes beyond the scheduled arrival time are considered to be on-time and delayed otherwise [35].

Feature scaling was performed for algorithms LR, DT and the hybrid approach MLP with PCA using sklearn and keras libraries. For the analysis of the supervised machine learning algorithms used to classify the predictive delay of the Saudi Airlines' flights eight metrics have been primarily considered which are confusion matrix, accuracy, precision, recall, F1 score, ROC.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In this section a description for the results for both dataset used.

Predictive analytics was performed using several machine learning models (LR , DT and the hybrid proposed model MLP with PCA) in order to predict the delay/non-delay variable using other features on two datasets metrics training and testing accuracies, precision, recall, ROC, F1 Score are given in Table 3 for the GACA dataset results. As we see the best model considering both training and testing accuracies is MLP with PCA which have 0.8954 and 0.8957 respectively. While in the other hand the decision tree performed better on the training by having 0.909 but was less in the testing by 0.891.

International Journal of Engineering Research and Technology. ISSN 0974-3154 Vol.13, No.4 (2020), pp. 814-822
 https://dx.doi.org/10.37624/IJERT/13.4.2020.814-822

Table3: Results of GACA Dataset

	DT	LR	MLP with PCA (Sklearn)	MLP with PCA (Keras)
Training Accuracy	0.909	0.889	0.884	0.8954
Testing Accuracy	0.891	0.893	0.879	0.8957
MA Precision	0.76	0.79	0.74	0.77
WA Precision	0.88	0.88	0.86	0.89
MA Recall	0.72	0.66	0.62	0.73
WA Recall	0.89	0.89	0.88	0.90
MA F1 Score	0.74	0.70	0.65	0.75
WA F1 Score	0.89	0.88	0.86	0.89
ROC AUC	0.899	0.921	0.913	0.728
Avg Precision Score	0.979	0.987	0.985	0.925

MA = Macro Average
WA = Weighted Average
ROC AUC = Area Under Curve of ROC Curve

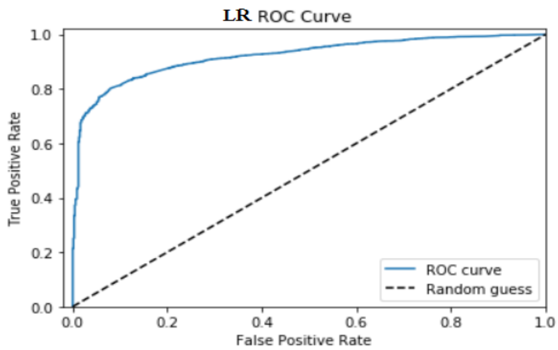

Figure 8: LR ROC Curve in GACA dataset

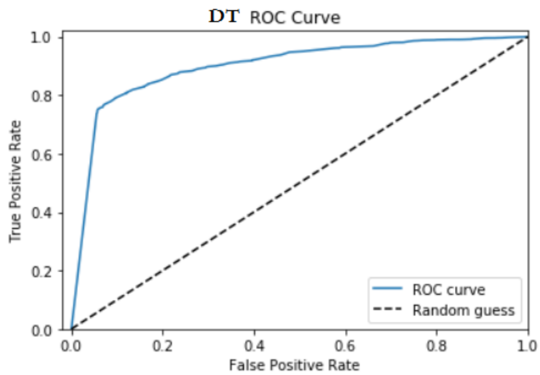

Figure 9: DT ROC Curve in GACA Dataset

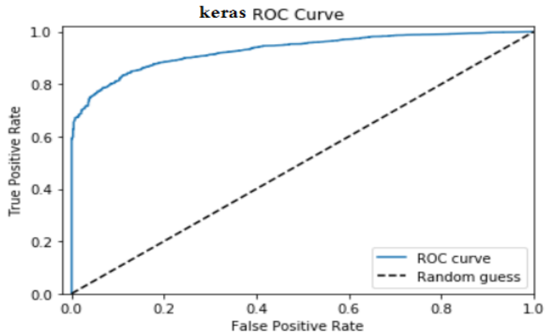

Figure 10:MLP with PCA ROC Curve(Keras) in GACA Dataset

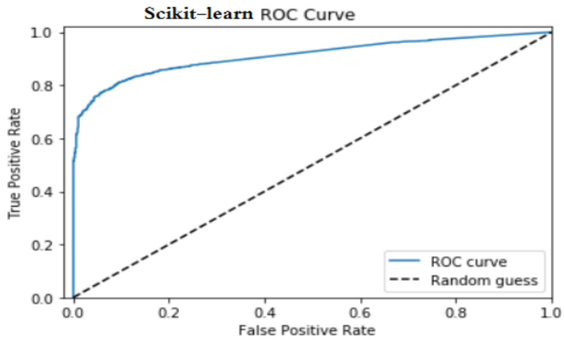

Figure 11: MLP with PCA (sklearn) in GACA Dataset

The ROC Curves gives out the predictive performance of every algorithm in Fig.[8-11], as seen all of the models give perfect ROC curve especially logistic regression and the proposed model because it reaches around score 1

Kaggle Benchmark Results:

The curves below describe the Roc, precision-recall curve and random guess for Kaggle dataset.

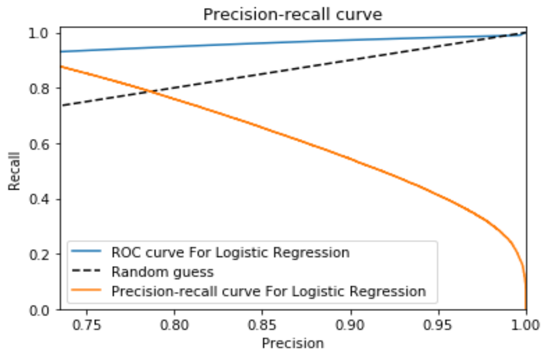

Figure12: Logistitc Regression Curve for Kaggle Dataset

In Figure 12 the curve shows the roc, precision-recall and random guess for the algorithm LR that applied on the Kaggle dataset.

International Journal of Engineering Research and Technology. ISSN 0974-3154 Vol.13, No.4 (2020), pp. 814-822
© International Research Publication House. https://dx.doi.org/10.37624/IJERT/13.4.2020.814-822

Table 4: Kaggle Dataset Results

Model	Train Accuracy	Test Accuracy	Average Precision score	Roc Accuracy
DT	0.92	0.86	0.944	0.923
LR	0.7114	0.7115	0.88	0.7917
MLP with PCA	0.9842	0.9843	0.9993828087319063	0.99882

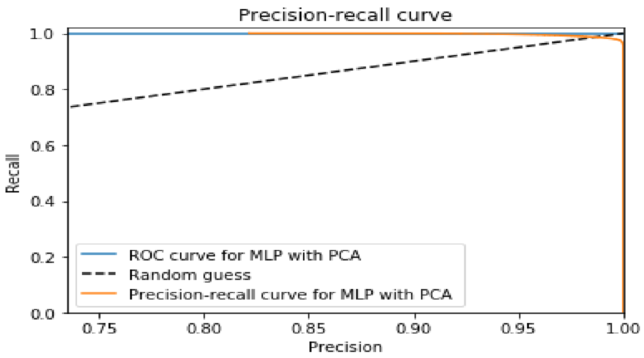

Figure 13:MLP with PCA Curve for Kaggle Dataset

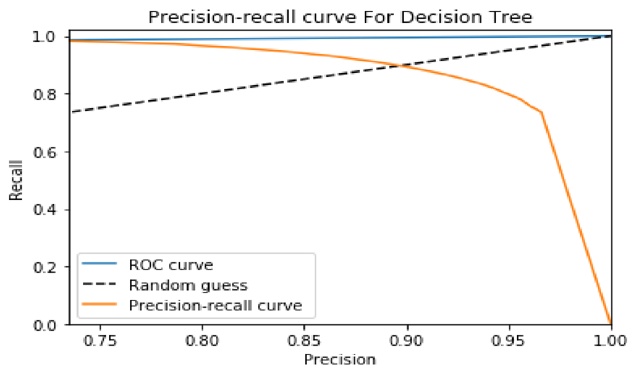

Figure 14: DT Curve for Kaggle Dataset

In Figure 13, the curves show the roc, precision-recall and random guess for the algorithm MLP with PCA that applied on the Kaggle dataset, it considers the best because it's get 1 degree.

In Figure 14, the curves show the roc, precision-recall and random guess for the algorithm DT that applied on the Kaggle dataset.

The predictive models were applied on a dataset from Kaggle there was a research on it that performed 92% as accuracy using neural networks. As seen in Table 4, the accuracy for the hybrid model was better by 0.98% for predicting the delay.

VII. CONCLUSION

With the continuously growing travel demand, limited capacity of airports and growing volume of aviation traffic flight delay has become a common phenomenon in aviation industry. Accurate delay prediction has thus become indispensable to alleviate airport congestion and improve the relatively low on-time performance of major

Commercial airports. The experimental results show that the proposed model performs the best with a testing accuracy of 89.57% for the local dataset GACA and for the other benchmark in Kaggle the results were 0.9843% using the suggested hybrid model. This is very good as compared with existing state-of-the-art. Predictive models with greater accuracy may be developed with the usage of larger datasets and more complex hybrid predictive models on appropriate amount of computer processing capabilities. Our research shows that supervised machine learning predictive models can be used to development intelligent systems of corporate use to multinational commercial passenger aircraft operator in their quest to offer on-time quality service to their passengers at industry competitive airfares. The motivation of this research was to apply a hybrid algorithm that contains PCA as a machine learning tool used to reduce the features in the dataset and neural network algorithm as a deep learning model for prediction. In addition, working on local Saudi airlines company's data is considered as first work in the area of kingdom Saudi Arabia.

The proposed method has proved to be highly capable of predicting accurate flight delay. The suggested path for further enhancement is to implement it as system in the airline companies, to announce early about the delay. In addition to apply it in the train data station related to delay issues.

REFERENCES

[1] G. Yablonsky, R. Steckel, D. Constales, J. Farnan, D. Lercel, and M. Patankar, "Flight delay performance at hartsfield-jackson atlanta international airport," *J. Airl. Airpt. Manag.*, vol. 4, no. 1, pp. 78–95, 2014.

[2] N. Chakrabarty, T. Kundu, S. Dandapat, A. Sarkar, and D. K. Kole, "Flight Arrival Delay Prediction Using Gradient Boosting Classifier," in *Emerging Technologies in Data Mining and Information Security*, Springer, 2019, pp. 651–659.

[3] H. Chen, J. Wang, and L. Feng, "Research on the Dynamic Data-driven Application System Architecture for Flight Delay Prediction.," *JSW*, vol. 7, no. 2, pp. 263–268, 2012.

[4] C. L. Monma and M. Stoer, "Handbook in Operations Research and Management Science," 1993.

[5] D. Long, D. Lee, J. Johnson, E. Gaier, and P. Kostiuk, "Modeling air traffic management technologies with a queuing network model of the national airspace system," 1999.

[6] S. T. Shorrock and B. Kirwan, "Development and application of a human error identification tool for air traffic control," *Appl. Ergon.*, vol. 33, no. 4, pp. 319–336, 2002.

[7] J. V Hansen, "Genetic search methods in air traffic control," *Comput. Oper. Res.*, vol. 31, no. 3, pp. 445–459, 2004.

[8] D. Bertsimas, G. Lulli, and A. Odoni, "An integer optimization approach to large-scale air traffic flow management," *Oper. Res.*, vol. 59, no. 1, pp. 211–227, 2011.

[9] L. Belcastro, F. Marozzo, D. Talia, and P. Trunfio, "Using scalable data mining for predicting flight delays," *ACM Trans. Intell. Syst. Technol.*, vol. 8, no. 1, p. 5, 2016.

[10] *Defining and Measuring Aircraft Delay and Airport Capacity Thresholds: ACRP report 104*. Transportation Research Board, 2014.

[11] N. Kuhn and N. Jamadagni, "Application of Machine Learning Algorithms to Predict Flight Arrival Delays." I:().

[12] L. Fausett, *Fundamentals of neural networks: architectures, algorithms, and applications*. Prentice-Hall, Inc., 1994.

[13] R. Nigam and K. Govinda, "Cloud based flight delay prediction using logistic regression," in *2017 International Conference on Intelligent Sustainable Systems (ICISS)*, 2017, pp. 662–667.

[14] S. Manna, S. Biswas, R. Kundu, S. Rakshit, P. Gupta, and S. Barman, "A statistical approach to predict flight delay using gradient boosted decision tree," in *2017 International Conference on Computational Intelligence in Data Science (ICCIDS)*, 2017, pp. 1–5.

[15] M. Baluch, T. Bergstra, and M. El-Hajj, "Complex analysis of united states flight data using a data mining approach," in *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*, 2017, pp. 1–6.

[16] V. Venkatesh, A. Arya, P. Agarwal, S. Lakshmi, and S. Balana, "Iterative machine and deep learning approach for aviation delay prediction," in *2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON)*, 2017, pp. 562–567.

[17] Y. J. Kim, S. Choi, S. Briceno, and D. Mavris, "A deep learning approach to flight delay prediction," in *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*, 2016, pp. 1–6.

[18] L. Moreira, C. Dantas, L. Oliveira, J. Soares, and E. Ogasawara, "On Evaluating Data Preprocessing Methods for Machine Learning Models for Flight Delays," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, pp. 1–8.

[19] M. Waqar, H. Dawood, P. Guo, M. B. Shahnawaz, and M. A. Ghazanfar, "Prediction of Stock Market by Principal Component Analysis," in *2017 13th International Conference on Computational Intelligence and Security (CIS)*, 2017, pp. 599–602.

[20] A. Nagpal, "Principal Component Analysis- Intro," 2017. [Online]. Available: https://towardsdatascience.com/principal-component-analysis-intro-61f236064b38.

[21] N. Das, B. Das, R. Sarkar, S. Basu, M. Kundu, and M. Nasipuri, "Handwritten Bangla basic and compound character recognition using MLP and SVM classifier," *arXiv Prepr. arXiv1002.4040*, 2010.

[22] "Classification: Accuracy." [Online]. Available: https://developers.google.com/machine-learning/crash-course/classification/accuracy.

[23] S. Narkhede, "Understanding Confusion Matrix," *towardsdatascience*, 2018. [Online]. Available: https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62.

[24] B. Thiagarajan, L. Srinivasan, A. V. Sharma, D. Sreekanthan, and V. Vijayaraghavan, "A machine learning approach for prediction of on-time performance of flights," in *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*, 2017, pp. 1–6.

[25] D. A. Pamplona, L. Weigang, A. G. de Barros, E. H. Shiguemori, and C. J. P. Alves, "Supervised Neural Network with multilevel input layers for predicting of air traffic delays," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, pp. 1–6.

[26] S. Khanmohammadi, S. Tutun, and Y. Kucuk, "A new multilevel input layer artificial neural network for predicting flight delays at JFK airport," *Procedia Comput. Sci.*, vol. 95, pp. 237–244, 2016.

[27] B. Sridhar, Y. Wang, A. Klein, and R. Jehlen, "Modeling flight delays and cancellations at the national, regional and airport levels in the United States," in *8th USA/Europe ATM R&D Seminar, Napa, California (USA)*, 2009.

[28] H. Alonso and A. Loureiro, "Predicting flight departure delay at Porto Airport: A preliminary study," in *2015 7th International Joint Conference on Computational Intelligence (IJCCI)*, 2015, vol. 3, pp. 93–98.

[29] B. Yu, Z. Guo, S. Asian, H. Wang, and G. Chen, "Flight delay prediction for commercial air transport: A deep learning approach," *Transp. Res. Part E Logist. Transp. Rev.*, vol. 125, pp. 203–221, 2019.

[30] H. Khaksara and A. Sheikholeslamib, "Airline delay prediction by machine learning algorithms."

[31] M. Bloem and N. Bambos, "Stochastic models of Ground Delay Program implementation for prediction, simulation, and insight," *J. Aerosp. Oper.*, no. Preprint, pp. 1–33, 2017.

[32] "GACA." [Online]. Available: https://gaca.gov.sa/web/en-gb/page/home.

[33] "Kaggle." [Online]. Available: https://www.kaggle.com/giovamata/airlinedelaycauses/data#DelayedFlights.csv.

[34] V. R, "Feature selection — Correlation and P-value." [Online]. Available:

International Journal of Engineering Research and Technology. ISSN 0974-3154 Vol.13, No.4 (2020), pp. 814-822
 https://dx.doi.org/10.37624/IJERT/13.4.2020.814-822

https://towardsdatascience.com/feature-selection-correlation-and-p-value-da8921bfb3cf.

[35] S. Choi, Y. J. Kim, S. Briceno, and D. Mavris, "Prediction of weather-induced airline delays based on machine learning algorithms," in *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*, 2016, pp. 1–6.

ABOUT THE AUTHORS:

Bashayer Alharbi received a B.S degree in computer science from Hail University, Saudi Arabia in 2015. Student at Qassim university for Master degree in Computer Science.

Master Prince received a B.S degree in computer science from Patna University, India in 1996. He received an M.S degree in computer science from Indira Gandhi National Open University, New Delhi, India in 2004 and the Ph.D. degree in computer science from Pune University, India 2008.

From 2009 to date, he has been working as Assistant Professor in the Department of Computer Science, Qassim University, KSA. His research interests include computer vision and machine learning.

Dr. Prince received the Best Ph.D. Thesis Dissertation of the Year 2009 Award of the Pune University, India.