Controlling Air Traffic Congestion By Predicting Flight Departure Delays And Duration: Integrating Machine Learning Sampling Techniques And Deep Learning Approaches

Waqar Ahmed Khan[1*], Sai-Ho Chung[1], Hoi-Lam Ma[2]
[1]Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong
[2]Department of Supply Chain and Information Management, The Hang Seng University of Hong Kong, Hong Kong

*Corresponding Author: waqarahmed.khan@connect.polyu.hk

Abstract
Flight delays that may propagate through the entire aviation sector are becoming a cause of flight cancellations and resource reallocation which can result in huge economic losses to airlines. In this study, a novel hierarchical integrated machine learning model is proposed to accurately predict flight delays and durations in series rather than in parallel in the future forecast horizon, above a certain threshold, to avoid ambiguity in decision making. The historical data was provided by an international airline operating in Hong Kong and our findings provide insights for airlines in controlling future flight delays.

1. Introduction
Flight delays can propagate through the entire aviation network and have a great effect on passengers and cargos demand, which may cause huge economic losses to airlines in terms of paying high compensation. Consequently, a decrease in demand and an increase in block time may pressurize airlines to raise airfares to accommodate for an increase in operational expenses, forcing them to be less competitive. To facilitate airlines making better informed decisions about flight delays in the future forecast horizon is now an important research topic. Existing flight delay models are predicted above a certain threshold (Belcastro et al., 2016; Rebollo and Balakrishnan, 2014) from online available and/or domestically operated flights data (Belcastro et al., 2016; Du et al., 2018; Khanmohammadi, Tutun, and Kucuk, 2016; Rebollo and Balakrishnan, 2014; Tu, Ball, and Jank, 2008; Yazdi, Dutta, and Steven, 2017; Yu et al., 2019). The prediction models are limited in scope and further improvement can bring a breakthrough contribution. Implementing a single threshold delay model may not generate sufficient information about flight delay durations and implementing more than one threshold model in parallel may create ambiguities in decision making. The effect of international legislation cannot be studied from domestically operated flight data, and this may limit the scope of existing prediction models to domestic flights and make it inappropriate for international flight delay predictions. Furthermore, few efforts have been made to improve the existing machine learning algorithms for flight delay prediction. Too much hyperparameter adjustment with a lot of trial and error work to find the best network architecture along with high dimensional data may cause an algorithm to converge in a suboptimal solution.

The study addresses the above limitations by proposing a novel hierarchical integrated machine learning model to collectively predict flight delays and durations in the future forecast horizon. The historical data of 19,105 international flights operated over two years, from April 2015 to March 2017 was provided by one of the international airlines in Hong Kong. The major contributions of this study are: First, the study proposes to implement multiple prediction models above a certain threshold in series rather than parallel for predicting flight delays and durations to avoid ambiguity in decision making. Second, unlike existing works that apply simple random sampling technique which may cause overfitting, the current study applied eight different sampling techniques to balance data and make decision boundaries smoother. Third, historically internationally operated flights are used to capture the influence of international legislation on flight delays. Fourth, a new constructive neural network (CNN) is recommended for flight delay prediction. The algorithm advantages that it determines its own network architecture and analytically calculate connection weights by avoiding trial and error approaches.

Various experimental sets and comparison with the Backpropagation Neural Network (BPNN) and Support Vector Machine (SVM) demonstrated that the CNN with a combination of Synthetic Minority Over Sampling Technique and Tomek Links (SMOTETomek) sampling techniques is capable of handling flight delays problems more accurately.

2. Problem Statement

An international airline operating in Hong Kong plans to implement a prediction model in their system that is capable of forecasting future departure delays and possible durations in order to implement a contingency plan to avoid unnecessary delays. The flight dispatchers make a flight plan four hours prior to each international flight by gathering information about various operational parameters in order to define flight trajectories and fuel consumption for each flight. Late flight departures may affect the normal operations of the airline and such consideration during the preparation of the flight plan cannot be ignored. The current study proposes to predict flight departure delays and possible delay durations in a four hours future forecast horizon. The selected airline classifies flight delays into nine categories as communicated by international aviation transport authority (IATA). Figure 1 illustrates the percentage contribution of each category causing airline delays. The reactionary and miscellaneous category contributes highly to flight delays at 38.31% and the lowest contribution recorded is for damage to aircraft and automated equipment failure. In the literature, the top reason for flight delays is considered to be the weather, however, in the current study, it only contributes 1.43% of the total delay. Weather is an essential factor and may become an indirect cause of delay for the other categories. The analysis of the reasons for delays facilitates the selection of relevant operational parameters from historical high dimensional data that may be the cause of flight departure delays.

3. Solution Approaches

The popularity of machine learning techniques is gaining significant interest compared to traditional statistical techniques because of its improved ability in learning from historical events (Khan, Chung, Awan, et al., 2019a, 2019b; Tkáč and Verner, 2016). BPNN has attracted much attention in the aviation domain (Khanmohammadi, Tutun, and Kucuk, 2016; Lin and Vlachos, 2018) because of its universal approximation capability. The major limitation of fixed topology BPNN is its slow convergence because of the iterative tuning of the connection weights by gradient methods, complex hyperparameters adjustments, and the need for a lot of trial and error experimental work to determine the network topology (Huang, Zhu, and Siew 2006; Liew, Khalil-Hani, and Bakhteri 2016; Srivastava et al., 2014). The limitations together with lower user expertise may cause the network to converge at a suboptimal local minimal solution rather than the global minimal. To address the generalization performance and learning speed of BPNN, we plan to apply a novel self-organizing constructive neural network named the cascade principal component least squares neural network (CPCLS) (Khan, Chung, Ma, et al., 2019). The important characteristics of CPCLS are that it analytically calculates the connection weights on both sides of the network

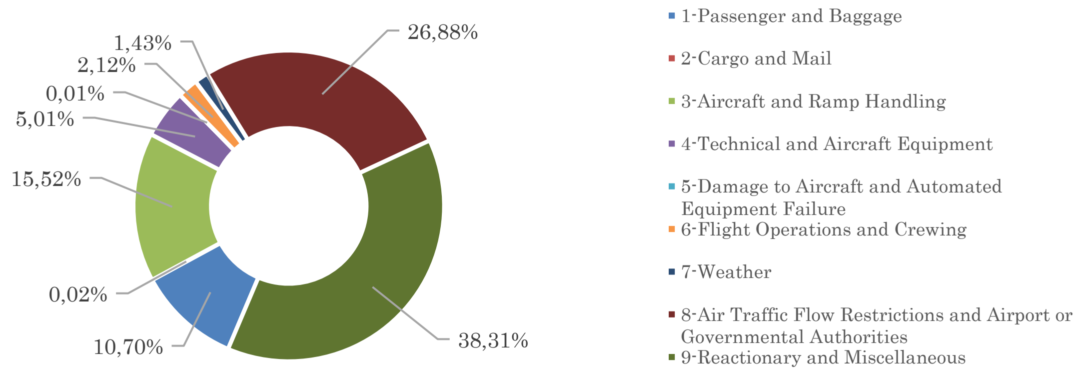

Figure 1. Reasons contributing to departure delay

rather than using the gradient method and determines its own network topology by adding multiple hidden units at each hidden layer.

CPCLS initializes by defining a number of hidden units h to be generated in the first layer H for given training data (x_i, y_i) with n samples, such that $x_i \epsilon R^n, i = 1,2, \dots \dots, l$ and $y \epsilon R^l$, such that $y_i \epsilon \{1,0\}$. For input weight connections w_{inp}, it orthogonally transforms the input operational parameters x and any preexisting h_i in to linearly independent h_{i+1} by eigendecomposition of covariance matrix C:

$$C = \frac{1}{n-1}(\boldsymbol{x} - \overline{\boldsymbol{x}})^T(\boldsymbol{x} - \overline{\boldsymbol{x}}) \tag{1}$$

The highest eigenvalues explaining maximum variance in operational parameters are selected and corresponding eigenvectors are determined:

$$|C - \lambda I| = 0 \tag{2}$$

$$(C - \lambda I)w_{inp} = 0 \tag{3}$$

The calculated eigenvectors are considered as w_{inp}. The h are determined by taking activation $\emptyset$ product of x and w_{inp} with added bias b:

$$h = \emptyset\left({w_{inp}}^T x + b\right) \tag{4}$$

For output weight connections w_{out}, the h and y are considered linear and calculated by ordinary least squares:

$$w_{out} = (h^T h)^{-1} h^T y \tag{5}$$

The flight delay is predicted by taking the product of h and w_{out}:

$$\hat{y} = {w_{out}}^T h \tag{6}$$

The error is calculated and if it is less then target error then stops, else add hidden units in the proceeding layers with respect to the previous layer and repeat Equations (1) to (6).

The orthogonal linear transformation in CPCLS ensures that h generated are linearly independent and capable of maximum error reduction. Besides, a small number of hyperparameters are generated not requiring trial and error experimental work or gradient derivation.

4. Experimental Results

The high dimensional historical flight data was provided by an international airline operating in Hong Kong for flight delays and possible duration prediction. The 19,105 international flights were undertaken over eight sectors from April 2015 to March 2017. In order to demonstrate the effectiveness of the CPCLS for delay prediction, a comparison study was performed with BPNN and SVM. The experimental work was carried out in Anaconda Python v3.2.6. The scikit-learn python tool was used for importing BPNN and SVM programming codes. The dataset was normalized in a range [0,1]. Rather than a simple accuracy metric, the classification and confusion matrix were also analyzed for a better conclusion of findings.

Table 1. Comparative study of machine learning algorithms in combination of sampling techniques for predicting departure delay status

Dataset type	Estimation Technique	Model Accuracy (%)		Label	Classification Report (%)		
		Train	Test		Precision	Recall	F1
SMOTETomek	BPNN	64.20	62.17	On-time	39	63	48
				delay	81	62	70
	SVM	63.40	63.35	On-time	39	58	47
				delay	80	65	72
	CPCLS	67.90	65.23	On-time	42	**66**	52
				delay	83	**65**	73

Table 2. Hierarchical integrated machine learning prediction by CPCLS with SMOTETomek

CPCL with SMOTETomek	Model Accuracy (%)		Label	Classification Report (%)		
	Train	Test		Precision	Recall	F1
Departure Delay (Level-1)	67.90	65.23	On-time	42	66	52
			delay	83	65	73
Threshold 60 min (Level-2)	68.17	62.95	1-60min	87	64	74
			>60min	28	59	38

Among different experimental work, BPNN with 20 hidden units and SVM with $C = 300$ and $\gamma = 1/i$ were selected as the best hyperparameters. The CPCLS was trained with 5 hidden units in the first layer with an addition of 150 hidden units in the proceeding layers.

An in-depth analysis of the data reveals that the classes are not balanced, and decision boundary overlapping exists. The majority of the data belongs to the delayed flights of 72% and a minority of data belongs to on-time flights of 28%. In order to improve class balancing and decision boundaries overlapping the study applied eight sampling techniques to select the one with a better prediction result. We employed undersampling techniques (such as Random Under-Sampling, Edited Nearest Neighbours (ENN), and Tomek Links), oversampling techniques (such as Random Over-Sampling, SMOTE, and Adaptive Synthetic), and combination techniques (such as SMOTETomek and SMOTEENN). Table 1 shows the comparative study among BPNN, SVM, and CPCLS for predicting the flight departure delay status. Among the various sampling techniques, SMOTETomek in combination with CPCLS shows a better recall of 66% for on-time flights and 65% for delayed flights.

The experimental work on the departure delay status (Level-1) prediction demonstrates that CPCLS with SMOTETomek has better performance, and the same techniques are used to predict the departure duration (Level-2). A novel hierarchical integrated machine learning model is proposed to predict the departure duration in series to avoid ambiguities in the decision making. The model is implemented in series rather than parallel for predicting various thresholds. For each threshold prediction, the previously labeled data not related to the threshold are eliminated during the training algorithms. For instance, for the threshold of 60min (Level-2), the algorithm defines two classes by considering flight delay data and eliminating the on-time data. This makes this model unique so it can be applied to any numbers of thresholds. Table 2 summarizes the results for the flight departure status and delay duration.

5. Conclusion and Future work

The hierarchical integrated model works by facilitating airlines to initially determine the flight status and if a delay is predicted in the future, then it helps to determine the possible delay duration. The Level-2 prediction can be extended to any number of thresholds to get insights on short delays. In the future, the CPCLS in combination with various sampling techniques can be applied to predict the delay reasons and integrate them into the existing hierarchical model. This can benefit airlines in predicting the delay status and duration, along with possible delay reasons, in order to recommend early remedial action for smooth operations.

References

Belcastro L, Marozzo F, Talia D, Trunfio P (2016) Using Scalable Data Mining for Predicting Flight Delays. *ACM Transactions on Intelligent Systems and Technology* 8(1):1–20.

Du WB, Zhang MY, Zhang Y, Cao XB, Zhang J (2018) Delay causality network in air transport systems. *Transportation Research*

Part E: Logistics and Transportation Review 118:466–476.
Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine: theory and applications. *Neurocomputing* 70(1–3):489–501.
Khanmohammadi S, Tutun S, Kucuk Y (2016) A New Multilevel Input Layer Artificial Neural Network for Predicting Flight Delays at JFK Airport. *Procedia Computer Science* 95:237–244.
Khan WA, Chung SH, Awan MU, Wen X (2019a) Machine learning facilitated business intelligence (Part I): Neural networks learning algorithms and applications. *Industrial Management & Data Systems* 120(1):164–195.
Khan WA, Chung SH, Awan MU, Wen X (2019b) Machine learning facilitated business intelligence (Part II): Neural networks optimization techniques and applications. *Industrial Management & Data Systems* 120(1):128–163.
Khan WA, Chung SH, Ma HL, Liu SQ, Chan CY (2019) A novel self-organizing constructive neural network for estimating aircraft trip fuel consumption. *Transportation Research Part E: Logistics and Transportation Review* 132:72–96.
Kim YJ, Choi S, Briceno S, Mavris D (2016) A deep learning approach to flight delay prediction. *Proc. 2016 IEEE/AIAA 35th Digital Avionics Systems Conf. (DASC)* (IEEE Computer Society, Washington, DC),1-6.
Liew SS, Khalil-Hani M, Bakhteri R (2016) An optimized second order stochastic learning algorithm for neural network training. *Neurocomputing* 186:74–89.
Lin Z, Vlachos I (2018) An advanced analytical framework for improving customer satisfaction: A case of air passengers. *Transportation Research Part E: Logistics and Transportation Review* 114:185–195.
Rebollo JJ, Balakrishnan H (2014) Characterization and prediction of air traffic delays. *Transportation Research Part C: Emerging Technologies* 44:231–241.
Srivastava N, Hinton G, Krizhevsky A, Sutskever I, and Salakhutdinov R (2014) Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Reserach* 15(1):1929–1958.
Tkáč M, Verner R (2016) Artificial neural networks in business: Two decades of research. *Applied Soft Computing* 38:788–804.
Tu Y, Ball MO, Jank WS (2008) Estimating Flight Departure Delay Distributions—A Statistical Approach With Long-Term Trend and Short-Term Pattern. *Journal of the American Statistical Association* 103(481):112–125.
Yazdi AA, Dutta P, Steven AB (2017) Airline baggage fees and flight delays: A floor wax and dessert topping? *Transportation Research Part E: Logistics and Transportation Review* 104:83–96.
Yu B, Guo Z, Asian S, Wang H, & Chen G (2019) Flight delay prediction for commercial air transport: A deep learning approach. *Transportation Research Part E: Logistics and Transportation Review* 125:203–221.