National Training Aircraft Symposium (NTAS) 2022 - Bridging the Gap

A bidirectional deep LSTM machine learning method for flight delay modelling and analysis

Desmond B. Bisandu
Cranfield University, desmond.bisandu@cranfield.ac.uk

Irene Moulitsas
Cranfield University, i.moulitsas@cranfield.ac.uk

Authors Biography

Desmond Bala Bisandu is a PhD candidate and researcher at Cranfield University, in the Centre for Computational Engineering Sciences and the Digital Aviation Research and Technology Centre (DARTeC). He is working in the area of machine learning and passenger experience. He holds a BSc and MSc in Computer Science from University of Jos, Nigeria and the American University of Nigeria, Yola, respectively.

Irene Moulitsas holds a PhD in Scientific Computation from the University of Minnesota. She is currently a Senior Lecturer in Cranfield University in the UK, Course Director for the MSc in Computational & Software Techniques in Engineering and Team Leader for Machine Learning and Data Analytics at the Digital Aviation Research and Technology Centre (DARTeC). Her research has focused on the solution of complex problems by developing novel algorithms that enable the efficient execution of large scientific computations on parallel processing platforms. She has developed highly efficient serial and parallel algorithms and software, that are publicly available for use by numerous universities, research laboratories, and companies.

A Bidirectional Deep LSTM Machine Learning Method for Flight Delay Modelling and Simulation

[1,2]Desmond Bala Bisandu (desmond.bisandu@cranfield.ac.uk)
[1,2]Irene Moulitsas (i.Moulitsas@cranfield.ac.uk)

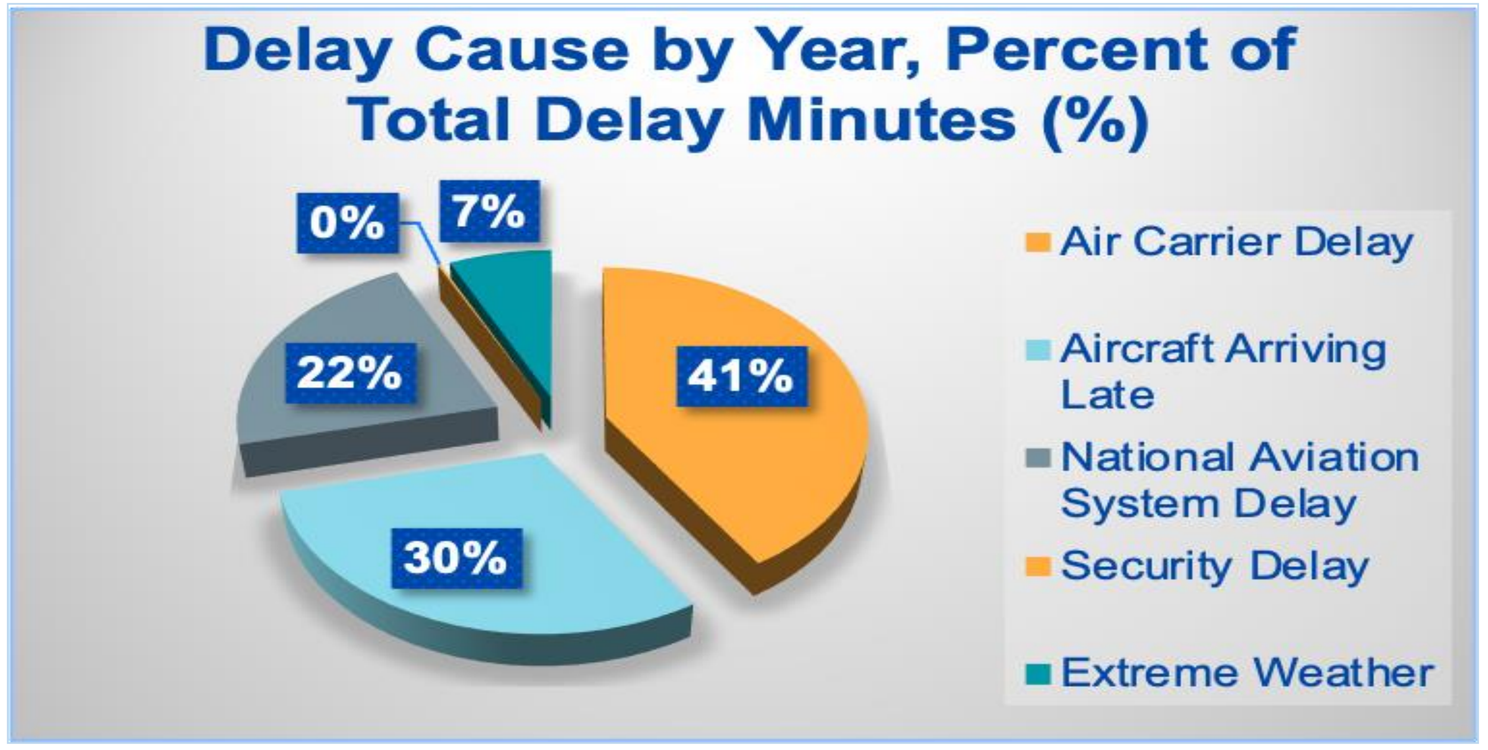
[1]Department of Computational Engineering Sciences
[2]Digital Aviation Research and Technology Centre: DARTeC (Machine Learning and Data Analytics Lab)

Overview

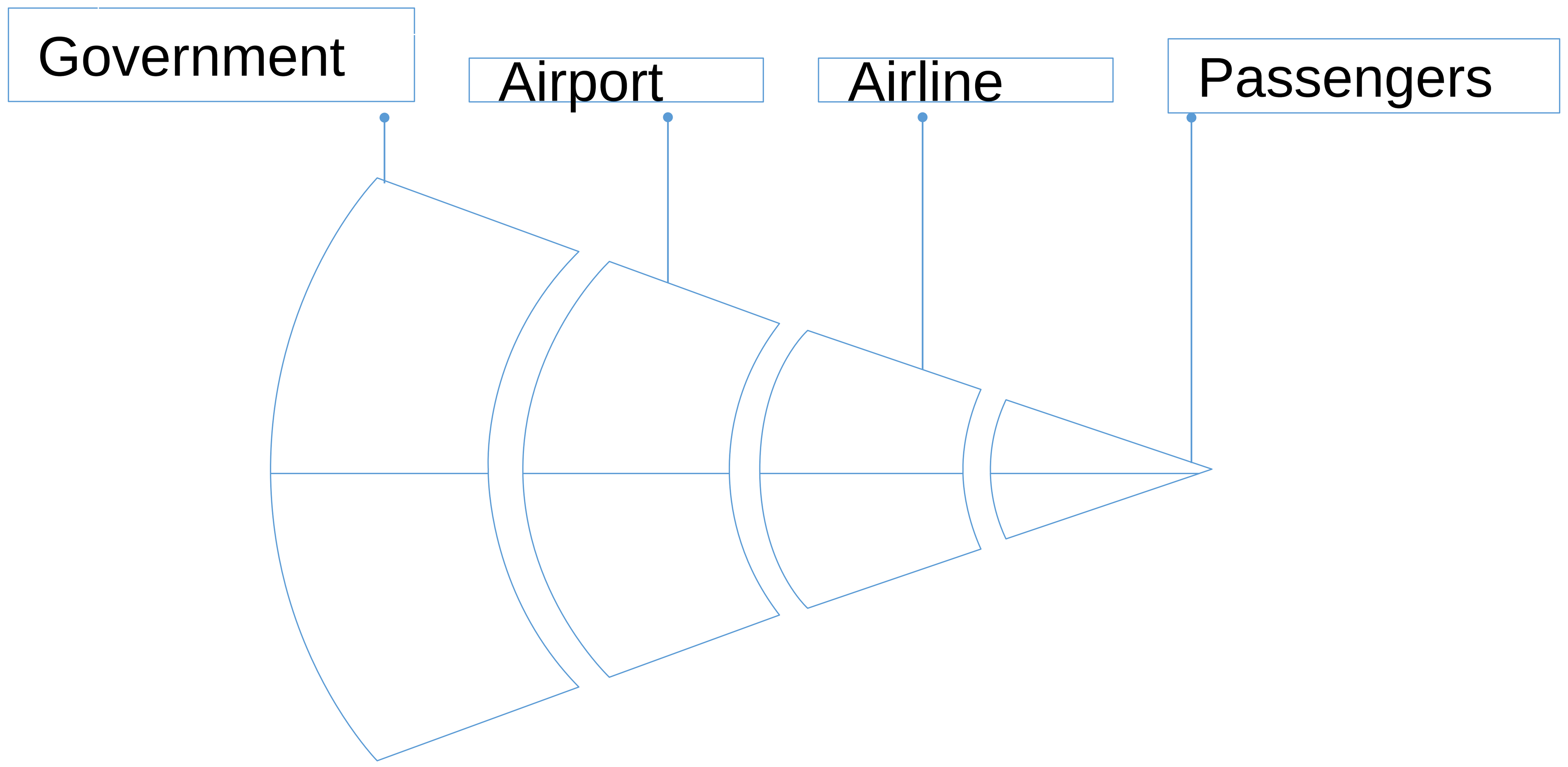
Introduction → Motivation & Problem Statement → Summary of Related Works → Our Contribution → Methodology → Results, Conclusion & Future Work

Introduction

Causes, effects & why flight delays prediction

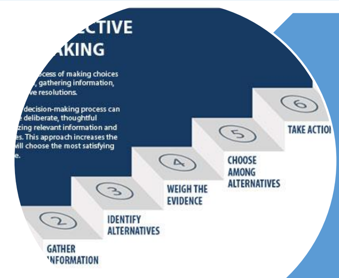


- Airline per hour - $1400 -$4500
- Passenger- $35 -$63



SOURCE: US Bureau of Transportation Statistics

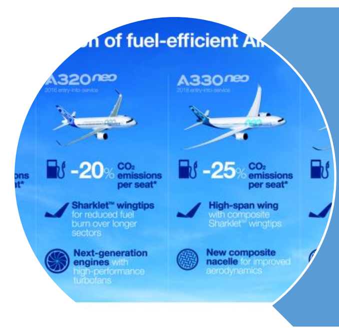
"Delayed flights remain a common problem on both sides of the Atlantic. Official figures show **11% of UK flights were delayed in the second quarter of 2021,** while **16% of US flights took off late in 2021 as a whole** (BBC, 7th Feb. 2022)."



Crucial in decision-making

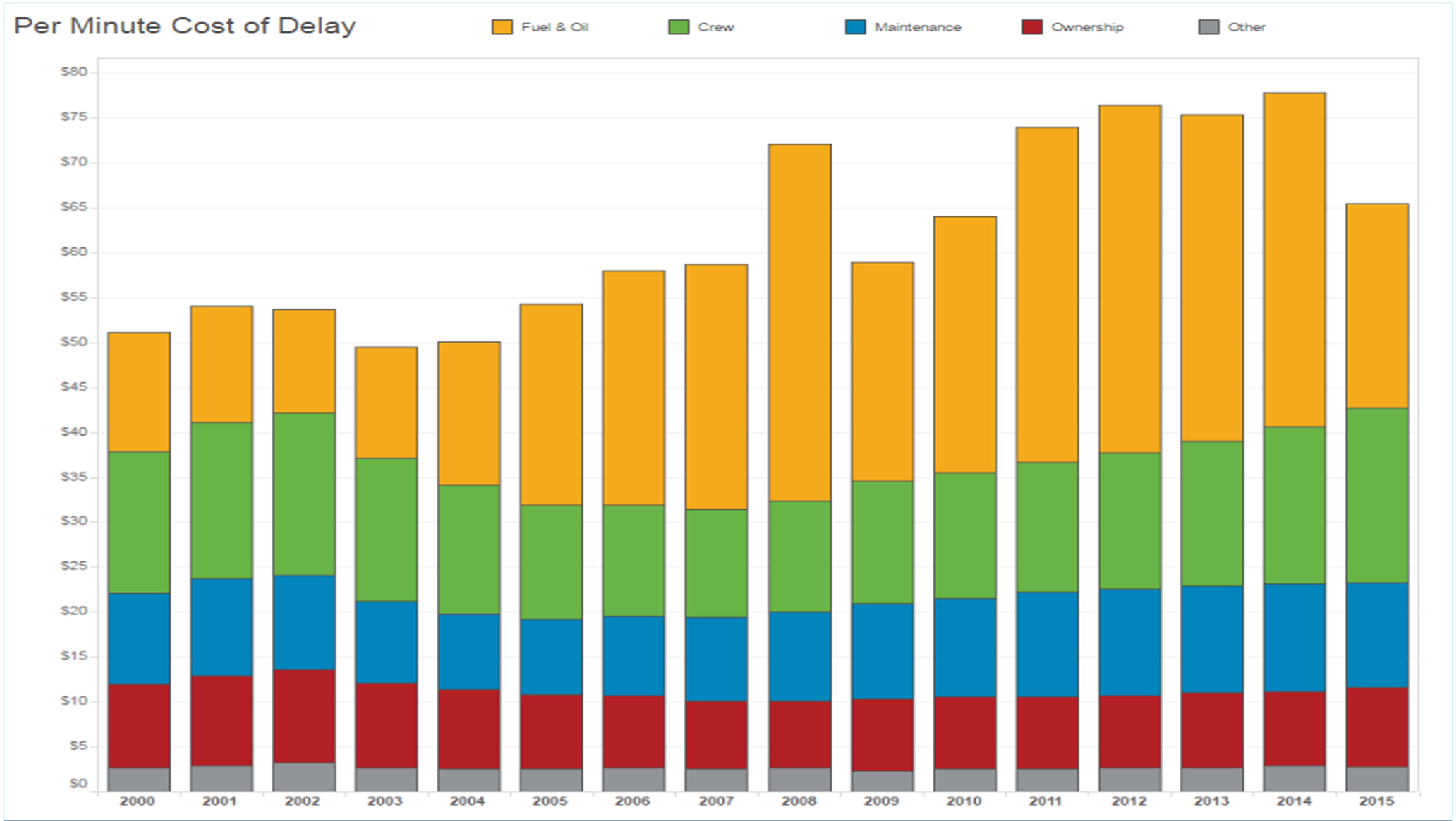
Negative economic impact on stakeholders'

Increase satisfaction & non-aviation revenue



Cut unnecessary expenditures. Thus, new & existing services can be improved

Motivation & Problem Statement



SOURCE: Flight Right and Bureau of Transportation Statistics (BTC)

Summary of Related Works

Bisandu et al., [1] utilise a special type of deep recurrent neural network (RNN) known as deep long short-term memory (LSTM) and social ski driver conditional autoregressive based deep learning to study non-weather impacted delays

In [2] and [3], the authors discuss the relevance of data in predicting flight delays and identify major methods such as machine learning, deep learning and statistical methods as the currently applied methods in the research of flight delays predictive tasks

In [4], the authors propose a method for predicting the flight departure delay in Nanjing Lukou International Airport by applying four different supervised machine learning algorithms

Kim et al., [5] proposed different architectural designs and implementation of LSTM and RNN in predicting flight delays using sequences of thresholds

Manna et al., [6] analyse air traffic data using a gradient boost decision tree, and the model produces higher accuracy based on their experiment

In [7], Chakrabarty used a grid search hyper-parameter tuning and gradient boosting classifier model for analysing and predicting the arrival delay of American Airlines using the top 5 must busiest airports with a binary classification technique.

Our Contribution

We proposed a deep BiLSTM architecture to perform flight delay analysis & prediction

We used real world dataset to train/test the deep BiLSTM model & then test the model in flight delays classification

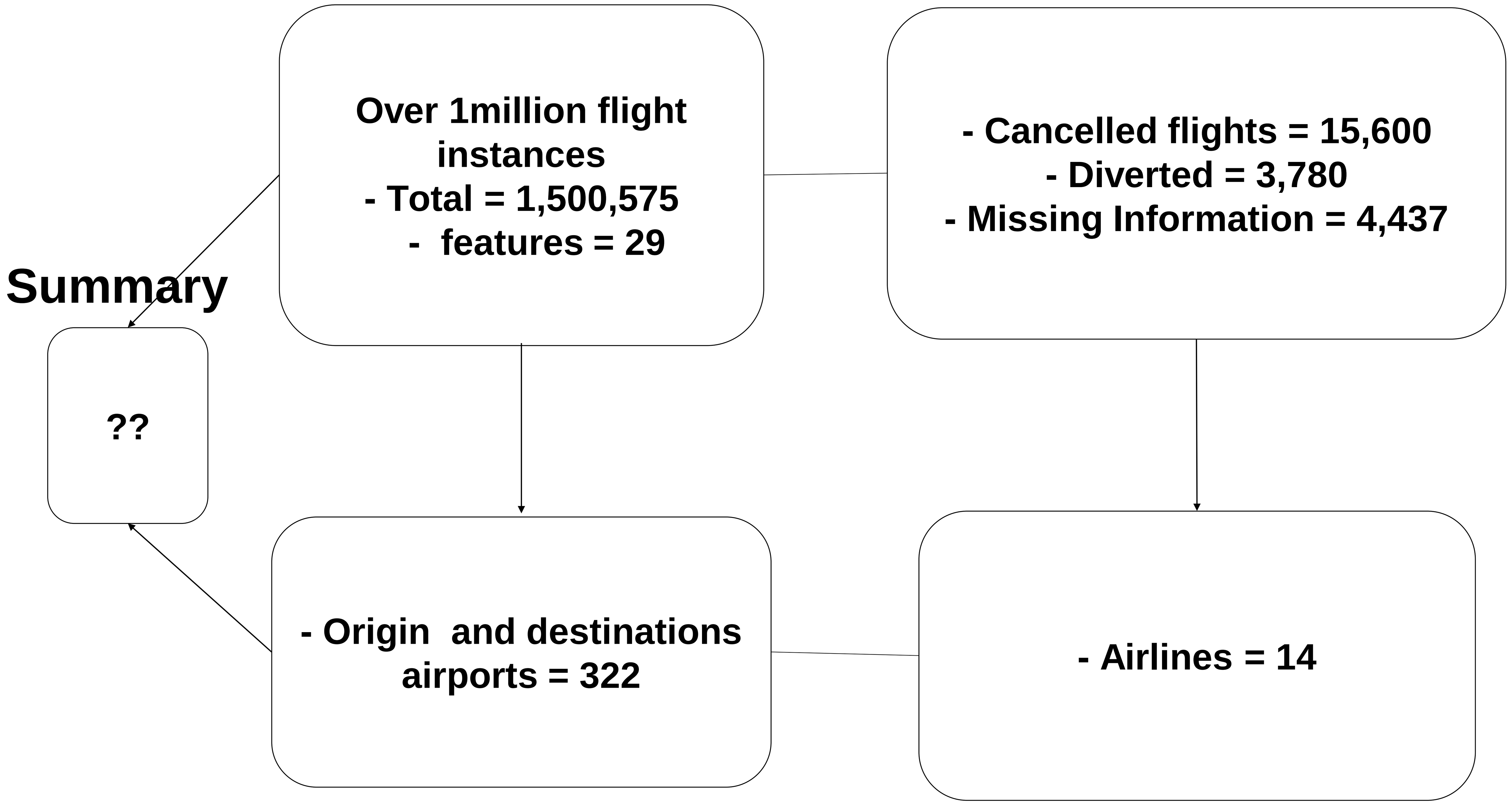
We compared the performances of our proposed deep BiLSTM with the LSTM on structure data

Dataset Description

Dataset Snippet

Year	Month	DayofMont	DayOfWee	DepTime	CRSDepT	ArrTime	CRSArrTi
2013	1	17	4	1038	1045	1451	1505
2013	1	18	5	1037	1045	1459	1505
2013	1	19	6	1035	1040	1515	1458
2013	1	20	7	1037	1045	1455	1505
2013	1	21	1	1044	1045	1446	1505
2013	1	22	2	1054	1045	1502	1505
2013	1	23	3	1036	1045	1504	1505
2013	1	24	4	1036	1045	1520	1505
2013	1	25	5	1042	1045	1525	1505
2013	1	26	6	1034	1040	1509	1458
2013	1	27	7	1042	1045	1506	1505
2013	1	28	1	1045	1045	1520	1505
2013	1	29	2	1116	1045	1607	1505
2013	1	30	3	1538	1045	1953	1505
2013	1	31	4	1059	1045	1520	1505
2013	1	1	2	1939	1840	29	2151
2013	1	2	3	1848	1840	2333	2151
2013	1	3	4	1617	1605	1925	1925

Period: October – December 2013



SOURCE: Flight Right and Bureau of Transportation Statistics (BTC: https://www.transtats.bts.gov/)

Feature Selection Criteria

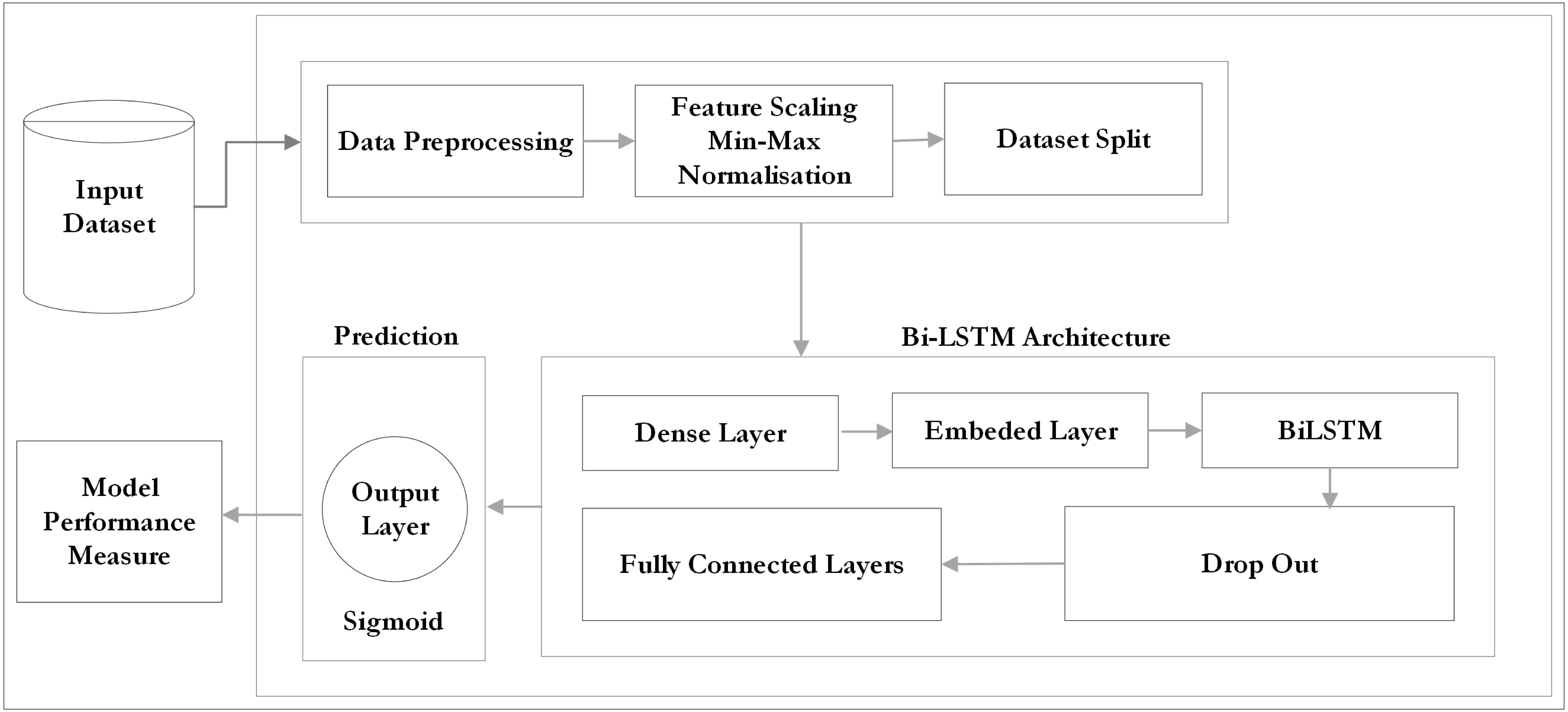
Feature should not contain information that indicates after a flight has taken off *e.g ArrTime & ActualEllapseTime etc.*

Features having no information about flight delays are removed *e.g TailNum & Year etc.*

All features with a high correlation with other features are removed *e.g DayofMonth & DayofWeek etc.*

S/No	Feature	Type	Description
1	Months	Date	Recorded flight months from October to December.
2	Distance	Date	Origin and destination distance in miles
3	ScheduleDepTime	Date	Departure schedule.
4	TaxiOut	Hours and Minutes	Taxi-out time in minutes.
5	DepDelay	Hours and Minutes	Time difference between a scheduled and actual departure.
6	ArrDelay	Hours and Minutes	Time difference between a scheduled and actual arrival.

Methodology



Results: Evaluation Metrics

All the computations were conducted on a Personal Computer (P.C.) with Intel(R) Core(T.M.) i7-9700 CPU with a processor speed of 3.00GHz and 32GHz RAM. We used libraries such as TensorFlow Core-2.4.1, TensorFlow GPU-2.4.1, Pytorch 1.9.1, NumPy-1.19.1, pandas-0.25.3, sci-kit learn-0.23.2, Scipy-1.5.2, PySimpleGUI-4.29.0 and Matplolib-3.3.1.

$$\text{Accuracy} = \frac{\sum_{a=0}^{z}(TP_a + TN_a)}{\sum_{a=0}^{z}(TP_a + TN_a + FP_a + FN_a)}$$

$$\text{Recall} = \frac{\sum_{a=0}^{z} TP_a}{\sum_{a=0}^{z}(TP_a + FN_a)}$$

$$\text{Precision} = \frac{\sum_{a=0}^{z} TP_a}{\sum_{a=0}^{z}(TP_a + FP_a)}$$

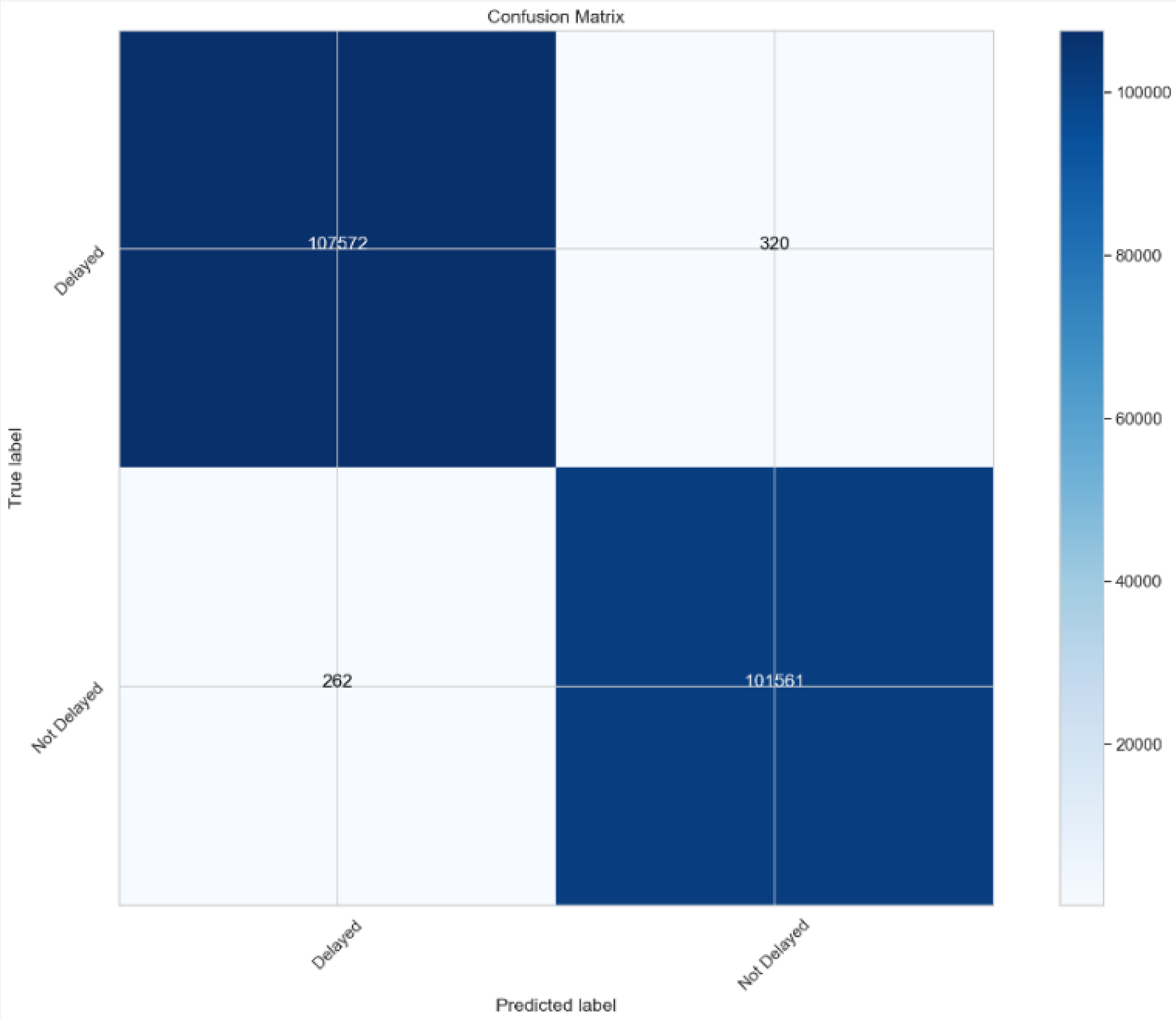
$$\text{F1} = 2 \text{ X } \frac{(Precision \text{ X } Recall)}{(Precision + Recall)}$$

$$\text{MCC} = \frac{\sum_{a=0}^{z}(TP_a \text{ X } TN_a) - (FN_a \text{ X } FP_a)}{\sum_{a=0}^{z}(TP_a + FP_a) + (TP_a + TN_a) + (TN_a + FP_a) + (TN_a + FN_a)}$$

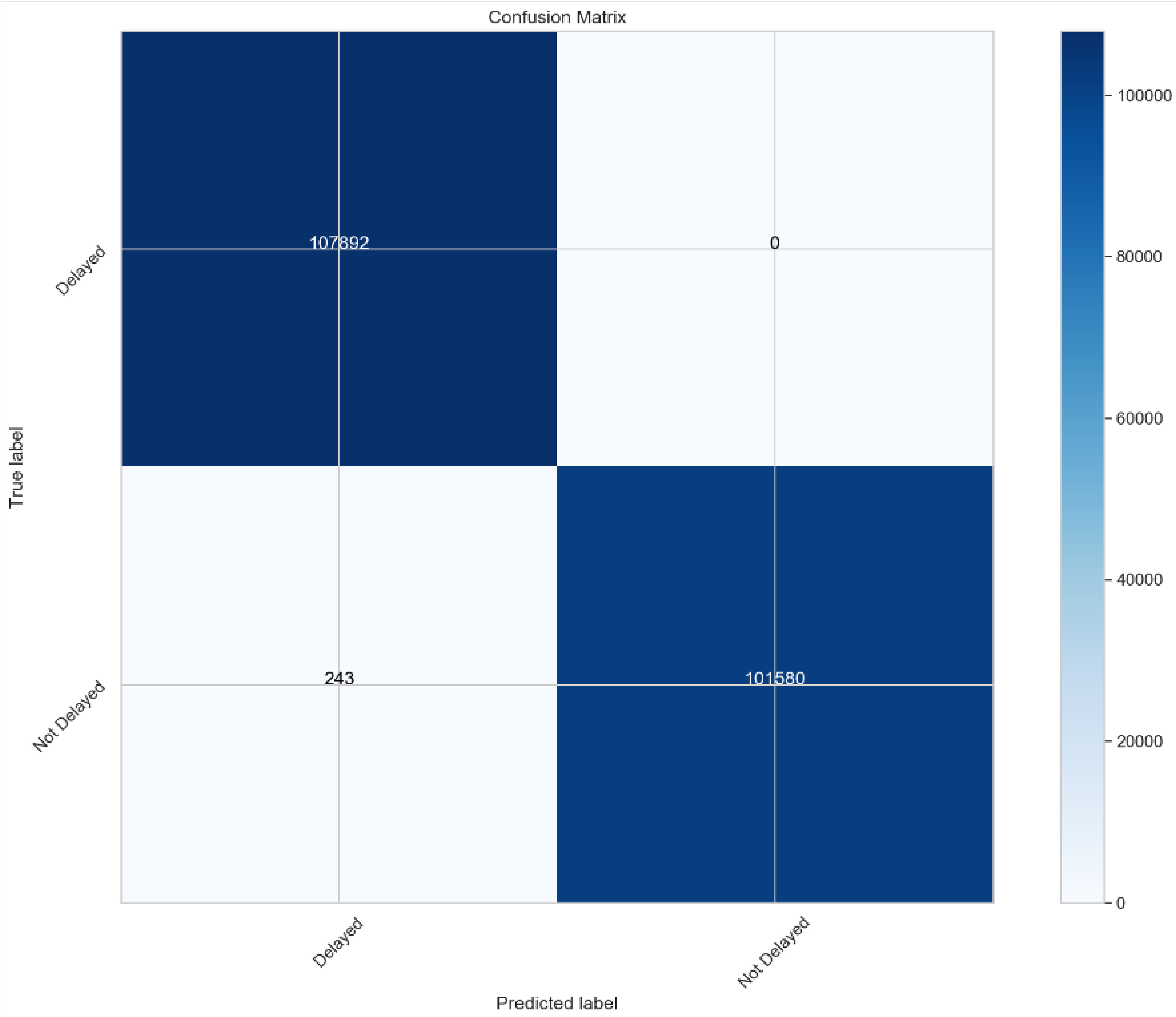
Results: Performance Comparison

S/No	Methods	Classes	Precision	Recall	F1-Score	Support	MCC
1	LSTM	Class 0	**0.8384**	0.9443	0.8756	107892	
		Class 1	0.4469	0.0743	0.0896	101823	
		Accuracy	-	-	0.7645	209715	0.4644
		Macro Average	0.4532	0.4532	0.4356	209715	
		Weighted Average	0.7362	0.7453	0.7453	209715	
2	BiLSTM	Class 0	0.8324	0.9898	0.8945	107892	
		Class 1	0.5643	0.0989	0.4988	101823	
		Accuracy	-	-	**0.9756**	209715	**0.9944**
		Macro Average	0.4202	0.4332	0.4122	209715	
		Weighted Average	0.70023	0.7234	0.7213	209715	

Results: LSTM Confusion Matrix



Results: BiLSTM Confusion Matrix



Conclusion & Future Work

Presented results were based on optimal results achieved by the LSTM & BiLSTM networks

Deep BiLSTM outperforms LSTM with higher average combine classification accuracy & Mathew's correlation coefficient value

Demonstrate an alternate approach for flight delay classification, and is expected to be among the recent contribution in the area

More cross-validation methods and larger sample sizes across different regions to further evaluate the models for a better generalisation

Modify neural network architectural design to achieve better tuning and higher accuracy of the models across different ratios of training/testing

LSTM low performance maybe improved by training the model with the weather, aircraft age, aircraft model & factors limiting airport infrastructure (i.e., available runways to flights)

The LSTM & BiLSTM models can be further investigated using weather-related delay since it is highly imbalanced

Thank you for listening

References

[1] D. B. Bisandu, I. Moulitsas, and S. Filippone, "Social ski driver conditional autoregressive-based deep learning classifier for flight delay prediction," Neural Comput. Appl., vol. 34, pp. 1–26, 2022, doi: 10.1007/s00521-022-06898-y.

[2] L. Carvalho et al., "On the relevance of data science for flight delay research: a systematic review," Transp. Rev., vol. 41, no. 4, pp. 499–528, 2021, doi: 10.1080/01441647.2020.1861123.

[3] E. R. Mueller and G. B. Chatterji, "Analysis of aircraft arrival and departure delay characteristics," AIAA's Aircr. Technol. Integr. Oper. 2002 Tech. Forum, no. October, pp. 1–14, 2002, doi: 10.2514/6.2002-5866.

[4] B. Ye, B. Liu, Y. Tian, and L. Wan, "A Methodology for Predicting Aggregate Flight Departure Delays in Airports Based on Supervised Learning," Sustainability, vol. 12, no. 7, p. 2749, 2020, doi: 10.3390/su12072749.

[5] Y. J. Kim, S. Choi, S. Briceno, and D. Mavris, "A deep learning approach to flight delay prediction," AIAA/IEEE Digit. Avion. Syst. Conf. - Proc., vol. 2016-Decem, pp. 1–6, 2016, doi: 10.1109/DASC.2016.7778092.

[6] 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," ICCIDS 2017 - Int. Conf. Comput. Intell. Data Sci. Proc., vol. 2018-Janua, pp. 1–5, 2018, doi: 10.1109/ICCIDS.2017.8272656.

[7] N. Chakrabarty, "A Data Mining Approach to Flight Arrival Delay Prediction for American Airlines," arXiv, pp. 4–9, 2019.

[8] X. Jinping, D. He, Liu, Xiangdong, D. Zhiyi, L. Erxue, C. Xiaofeng, and H. Zhenjiang, "Civil aviation administration of china," 2021 National Civil Aviation Work Conference and Safety Work Conference Held, 2021. [Online]. Available: http://www.gov.cn/xinwen/2021-01/12/content_5579282.htm.

[9] X. Fu, Z. Lei, S. Liu, K. Wang, and J. Yan, "On-time performance policy in the Chinese aviation market - An innovation or disruption?," Transp. Policy, vol. 95, no. June, pp. A14–A23, 2020, doi: 10.1016/j.tranpol.2020.06.008.

[10] Y. J. Liu, W. D. Cao, and S. Ma, "Estimation of arrival flight delay and delay propagation in a busy hub-airport," Proc. - 4th Int. Conf. Nat. Comput. ICNC 2008, vol. 4, pp. 500–505, 2008, doi: 10.1109/ICNC.2008.597.

[11] D. B. Bisandu, M. S. Homaid, I. Moulitsas, and S. Filippone, "A deep feedforward neural network and shallow architectures effectiveness comparison: Flight delays classification perspective," in The 5th International Conference on Advances in Artificial Intelligence (ICAAI 2021) in QAHE at Northumbria University London Campus, U.K., 2021, pp. 1–10, doi: https://doi.org/10.1145/3505711.3505712.

[12] R. W. Maxson, "Prediction of Airport Arrival Rates Using Data Mining Methods," 2018.