**APPLIED DATA SCIENCE**

**DEVELOPING A FLIGHT DELAY PREDICTION MODEL USING**

**MACHINE LEARNING**

TEAM LEADER : S.JAYSRI 142219205038

TEAM MEMBER 1 : M.BEEULAH MARRY 142219205013

TEAM MEMBER 2 : R.HEMALATHA 142219205033

TEAM MEMBER 3 : K.PRIYA DHARSHINI 142219205065

LITERATURE SURVEY

|  |  |  |  |
| --- | --- | --- | --- |
| **Literature** | **Methodology** | **Merits** | **Demerits** |
| Study of Flight  Departure Delay and  Causal Factor Using  Spatial Analysis | Assuming delay as a spatially dependent variable, finds delay distribution pattern to predict delay. | Considers spatial factors, people, day types and time ranges of a day to contribute to the prediction | Some of the attributes considered cannot be obtained on large scale in real time. |
| Flight delay forecasting and analysis of direct and indirect factors | LSTM network with attention mechanism to remember spatial dependencies | Direct and indirect causing factors are weighed differently. | Air interaction of flights not taken into account |
| Flight delay prediction based on aviation big data and machine learning | Comparison of LSTM and Random forest; Uses ADS-B data for improved accuracy. | Use of ADS-B can be seen promising.  Showed that LSTM  suffers from overfitting on test set. | Deployment of ADS-B is hectic. More data handing takes place. |
| Prediction of weather- induced airline delays based on machine learning algorithms | Experimented predicting delay using supervised machine learning algorithms. Uses SMOTE for weaker class sampling | Found weather causes to be amounting to a significant percent of delay. | Not taken spatial dependencies into account. Amount of delay could have been found. |
| A deep learning approach to flight delay prediction | A deep RNN and  LSTM approach to prediction; uses limited data attributes | Predicting two sections namely day prediction and flight prediction seems more reasonable and can give more  insights for the airport managers to make necessary arrangements. | Air traffic/flight interaction doesnt play great roles. Biased towards weather attributes. |
| Flight Delay Prediction System | Supervised Machine Learning algorithm using Naive Bayes | Considers independence among the predictors making the system scalable. Good for real time prediction | Does not take into account the impact of unprecedented reasons such as major calamities in flight delays |

|  |  |  |  |
| --- | --- | --- | --- |
| Predicting flight delay based on multiple linear regression | A multiple linear regression algorithm to predict delay | Both airline and weather features are taken into consideration.The methodology used in  this gives better results compared to Naive- Bayes and C4.5 approach | Predicts only the flights which are delayed above 30 minutes |
| Airline Flight Delay  Prediction Using  Machine Learning  Models | Comparison among 7 classification machine learning algorithms | Among the considered algos using 4 performance indicators decision tree was found to be the best in predicting flight delays | The data imbalance issue even though handled through weighted evaluation methods does had a significant effect on performance on the algorithms |
| Flight delay prediction from spatial and temporal perspective | ST-Random Forest for flight delay prediction using spatial features of aviation network and temporal correlation of weather condition and airport crowdedness on flight delays. | A real-time, highly accurate prediction system that guaranties the influence of the air traffic network in the prediction. | Overfitting might occur due to LSTM. |

**References**

1. Shaowu Cheng, Yaping Zhang, Siqi Hao, Ruiwei Liu, Xiao Luo, Qian Luo, "Study of Flight Departure Delay and Causal Factor Using Spatial Analysis", Journal of Advanced Transportation, vol. 2019, Article ID 3525912, 11 pages, 2019.<https://doi.org/10.1155/2019/3525912>

1. Wang, F., Bi, J., Xie, D., Zhao,X.: Flight delay forecasting and analysis of direct andindirect factors. IET Intell. Transp. Syst. 16, 890–907(2022)[.https://doi.org/10.1049/itr2.12183](https://doi.org/10.1049/itr2.12183)

1. Gui, G., Liu, F., Sun, J., Yang, J., Zhou, Z., & Zhao, D. (2019). Flight delay prediction based on aviation big data and machine learning. *IEEE Transactions on Vehicular Technology*, *69*(1), 140- 150.

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

1. Kim, Y. J., Choi, S., Briceno, S., & Mavris, D. (2016, September). A deep learning approach to flight delay prediction. In *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)* (pp. 1-6). IEEE.

1. Borse, Y., Jain, D., Sharma, S., Vora, V., & Zaveri, A. (2020). Flight Delay Prediction System. *Int. J. Eng. Res. Technol*, *9*(3), 88-92.

1. Ding, Y. (2017, August). Predicting flight delay based on multiple linear regression. In *IOP Conference Series: Earth and Environmental Science* (Vol. 81, No. 1, p. 012198). IOP Publishing.

1. Tang, Y. (2021, October). Airline Flight Delay Prediction Using Machine Learning Models. In *2021 5th International Conference on E-Business and Internet* (pp. 151-154).

1. Li, Q., & Jing, R. (2022). Flight Delay Prediction from Spatial and Temporal Perspective. *Expert Systems with Applications*, 117662.