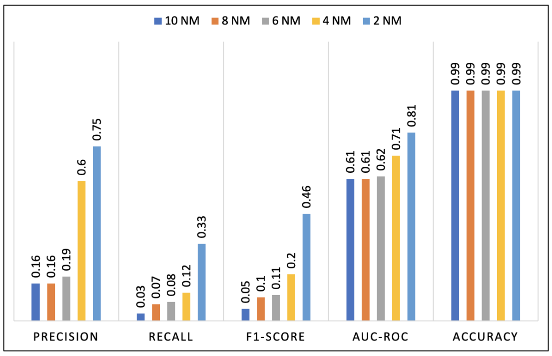

Figure 8. Histogram of the full data performance results in different radii. The results shows that the model performance increases while approaching the runway.

but only 15.5% of the detected go-arounds are actual go-arounds. Thus, we achieve a reduction in the false positive alerts by almost 60% while the detection rate decreases by approximately 8%. This means that an ATC can rely on the go-around prediction as being accurate the vast majority of the time, increasing trust in the system.

To sum up, the achieved results are more accurate when predicting a go-around compared to state-of-the-art models [7], [9], [11]. Furthermore, our findings confirm that predicting go-around occurrence is a very challenging problem. In fact, the number of detected go-around before 4 NM is very low (results in 7 and 8). The failure in accurately predicting a go-around can be justified through three main factors. First, the most important factor is that go-around events are subjective to the pilot decisions. Thus, the decision to go-around may be due to several factors that are hard to or even impossible be measured and/or collected, such as the pilot experience, personality, or his/her level of fatigue, etc. The second factor is the visibility. Even though this data is provided in METAR, it is not accurate and it does not provide the actual visibility level at the considered decision height. Finally, as we are only considering ADS-B data, our model lacks the airport surface movement data which we believe can affect a go-around decision. All these factors together mean that go-around and non go-around instances are closely aligned and similar.

V. Conclusion

The current work proposes a method to predict flight go-around events at the final approach phase for the purpose of increasing the situational awareness of tower controllers to safely perform their control tasks. It includes an innovative go-around trajectory labeling technique to detect go-around flights from historical data. Furthermore, it presents a data-driven model based on a binary classification prediction methods in order to predict flight go-around occurrence at different radii away from the runway threshold. In order to evaluate the performance of proposed method, computational experiments are conducted using ten months' of air traffic data for flights arriving at Philadelphia International Airport (PHL). Two types of experiments are conducted; the first includes down-sampling techniques and the second includes the full data set. The best prediction results are found at 2 NM away from the runway threshold. For the down-sampling data, the model is able to predict 56% of the go-arounds with only a 10% false positive alert rate, while for the full data set the model is able to detect 33% of the go-arounds with a 25% false alert rate. Our model outperforms state-of-the-art methods in terms of decreasing the false positive alerts in the system by 60%.

In future work we plan to address the go-around prediction problem by predicting the safety level of an approach profile. This will reduce the impact of the pilot subjectivity on the prediction, as well as alleviate the data imbalance problem.

Acknowledgement

This work was conducted under the Saab-NTU Joint Lab under its Machine Learning For Airport Management And Tower Control project with support from Saab AB (publ).

References

[1] Z. J. Lim *et al.*, "Causal effects of landing parameters on runway occupancy time using causal machine learning models," in *2020 SSCI*. IEEE.
[2] I. Dhief *et al.*, "Speed control strategies for e-aman using holding detection-delay prediction model," 2021.
[3] F. Dehais *et al.*, "Pilot flying and pilot monitoring's aircraft state awareness during go-around execution in aviation: A behavioral and eye tracking study," *The International Journal of Aerospace Psychology*, 2017.
[4] G. Adam and J. Condette, "Study on aeroplane state awareness during go-around," *Bureau d'Enquetes et d'Analyses pour la securite de l'aviation civile: Bourget, France*, 2013.
[5] T. Blajev and W. Curtis, "Go-around decision-making and execution project: Final report to flight safety foundation," *Flight Safety Foundation, March*, 2017.
[6] N. P. Singh *et al.*, "Real-time unstable approach detection using sparse variational gaussian process," in *2020 AIDA-AT*. IEEE.
[7] M. Gariel *et al.*, "On the statistics and predictability of go-arounds," *Computing Research Repository - CORR*, 2011.
[8] L. Dai *et al.*, "Modeling go-around occurrence," *13th USA/Europe ATM R& D Seminar, Vienna, Austria*, 2019.
[9] ——, "Predicting go-around occurrence with input-output hidden markov model," *ICRAT*, 2020.
[10] T. G. Puranik *et al.*, "Towards online prediction of safety-critical landing metrics in aviation using supervised machine learning," *Transportation Research Part C: Emerging Technologies*, 2020.
[11] B. Figuet *et al.*, "Predicting airplane go-arounds using machine learning and open-source data," *Proceedings*, 2020.
[12] FAA, "Air traffic by numbers," FAA report, Tech. Rep., 2019.
[13] M. Schfer *et al.*, "Bringing up opensky: A large-scale ads-b sensor network for research," in *Proceedings of ACM/IEEE IPSN*, 2014.
[14] I. E. Mesonet, "Asos-awos-metar data download."
[15] J. Hershberger and J. Snoeyink, "An o (n log n) implementation of the douglas-peucker algorithm for line simplification," in *Proceedings of the tenth annual symposium on Computational geometry*, 1994.
[16] T. Puranik *et al.*, "Energy-based metrics for safety analysis of general aviation operations," *Journal of Aircraft*, 2017.
[17] C. Wang *et al.*, "Imbalance-xgboost: leveraging weighted and focal losses for binary label-imbalanced classification with xgboost," *Pattern Recognition Letters*, 2020.
[18] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in 22^{nd} *International Conference on Knowledge Discovery and Data Mining*, 2016.
[19] D. Pham *et al.*, "An air traffic controller action extraction-prediction model using machine learning approach," *Complexity*, 2020.
[20] Y. Chang *et al.*, "Application of extreme gradient boosting trees in the construction of credit risk assessment models for financial institutions," *Applied Soft Computing*, 2018.
[21] Y. Sun *et al.*, "Classification of imbalanced data: A review," *International journal of pattern recognition and artificial intelligence*, 2009.