The paper "Predicting Time to Pushback of Flights in U.S. Airports" by Filienko et al. explores the development of a machine learning model to predict the pushback time of flights at U.S. airports. This prediction is crucial for efficient air traffic management as it affects subsequent estimates of taxiing and takeoff times. The authors propose a gradient boosting decision tree model that utilizes a comprehensive feature set, including weather conditions, airport activity, airline specifics, and aircraft characteristics. They trained and evaluated their model using a large dataset from 10 major U.S. airports, demonstrating that localized models for each airport are more memory efficient while maintaining prediction accuracy comparable to a global model.

The study highlights the importance of feature engineering in improving model accuracy. The features were derived from various data sources, such as weather forecasts and air traffic data processed by NASA's Fuser platform. Key features include estimated time of departure (ETD), aircraft type, and airport busyness indicators. The authors found that both global and local models could achieve similar mean absolute errors (MAE), but training local models per airport was computationally more efficient. This localized approach also allowed for parallel training, leading to better fault tolerance and memory efficiency.

The authors' model was one of the winners in the 2023 "Pushback to the Future" competition hosted by NASA. They achieved significant improvements in prediction accuracy over baseline models by integrating diverse data sources. The paper concludes with discussions on future work, including the potential for federated learning to incorporate sensitive airline data in a privacy-preserving manner. The results underscore the utility of their approach in enhancing the efficiency of air traffic management systems.