Abstract

Flight delays hurt airlines, airports, and passengers. Their prediction is crucial during the decisionmaking process for all players of commercial aviation. Moreover, the development of accurate prediction models for flight delays became cumbersome due to the complexity of air transportation system, the number of methods for prediction,and the deluge of flight data.In thiscontext, this paper presents a thorough literature review of approaches used to build flight delay prediction models from the Data Science perspective.We propose a taxonomy and summarize the initiatives used to address the flight delay prediction problem, according to scope, data, and computational methods,giving particular attention to an increased usage of machine learning methods.Besides,wealso present a timeline of significant works that depicts relationships between flight delay prediction problems and research trends to address them.