Lab Report

Title: (Prospectus) Air Traffic Analysis
Notice: Dr. Bryan Runck
Author: Megan Marsolek
Date: 9/28/21

Project Repository: NA
Google Drive Link: NA
Time Spent: 4 hours

Abstract

Air Travel has become an increasing and popular means of travel, with airlines expanding and offering more flights to more destinations, it becomes a challenge to monitor and safely manage airspace to avoid collisions around highly frequented areas like airports. There are many regulations in place to assure travel safely, but it has the potential to benefit from incorporating Machine Leaning (ML) to better estimate a plane's near future location based on constantly changing variables (weather) and fixed restrictions (airspace classifications).

Using a plane's location, the surrounding weather, and current airspace I would like to track and predict a plane's location in 3D space over time. Relying on three main sources for data, I will convert and prep the data to read into a ML model. I will then train and test the accuracy of by comparing it to recorded flight path. In the end I hope to visually display a sample of the ML predicted flight path with that of the real flight path and any surrounding weather and airspace classes for user interoperability.

Problem Statement

The intention is to predict a plane's flight path based on data that may influence a plane's course, such as Airspace classification or weather. Given we have access to its present and past locations as it flies (via ADS-B data) we hope to predict where it will likely be in 3D space for some specified time in the future.

Table 1. Relevant (attainable) variables for the problem

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Air craft locations	Point data (Lat., Long., Alt.) from ADS-B	Point data in csv (extracted from TAR files)	lat, lon, geoaltitude, heading, velocity, time	ADS-B Data	Possible web scraping, then subsetting of the data (may need to convert certain attribute units of measure e.g. feet to meters) Convert lat/lon to geometries
2	Weather	Measurements within each cell is 8 bit Base Reflectivity 0.5dbz (used to measure the severity of weather, the	Raster Sizes : 12000x5200 pixels OR	Pixel dbz measure in each cell	IEM NexRad Mosaic	Possible web scrapping (extracting archived raster images

		higher the dbz the more the severe the weather)	12200x5400 pixels (for the continental US)			over a specified time period) Matching CRS with rest of data, subsetting/cropping data to relevant spatial extant
3	Airspace Classification B	Regulated airspace across the US with flying permissions that dictat flight protocol (subset to focus only on class B airspace)	2D shapefile (vector) or access to a 3D Google kmz file	UPPER_VAL, LOWER_VAL, NAME, SECTOR, LEVEL	FAA Airspace Shapefiles	Matching CRS with all other data, Subsetting data to just class B

Input Data

All of the data proposed for this project are openly accessible and sourced data. Each is in a different format (csv, raster, and vector), with different temporal measures of time. The ADS-B data is relevant in locating where a plane in at any given time, based on initial glances planes constantly send out their location (roughly every few seconds), while the weather data updates new weather files every 5 mins. The only fixed data is the Airspace classification data, which contains information for B, C, D, and E airspaces.

The data will need to be converted to the same CRS system and PRS before loading into a ML model. I still need to explore whether I can store or locate more accurate weather data (for example 3D instead of 2D) data. I also need to spend more time on prepping (converting all the data to usable input formats for a ML) and subsetting (need to downscale the size of the project to the task at hand).

Table 2. Data Access

#	Title	Purpose in Analysis	Link to Source
1	ADS-B	ADS-B data records the location of plane at a given time in points (lat., long. Alt.). Its location can help predict where it might be at some time in the future. It also contains other attributes: make/model of aircraft, time, flight number, etc.	Open Sky Network
2	IEM NexRad Mosaic	This is Raster data for weather (radar). It will indicate where severe weather is located that may influence some flight paths. (smallest measure of time is 5 mins)	Iowa State University
3	Class B,C,D,E Airspace Shape Files	Define where Airspace classifications are in 3D space that may affect flight paths (represented as 2D but contains attributes that describe its min and max altitudes).	Federal Aviation Administration (FAA)

Methods

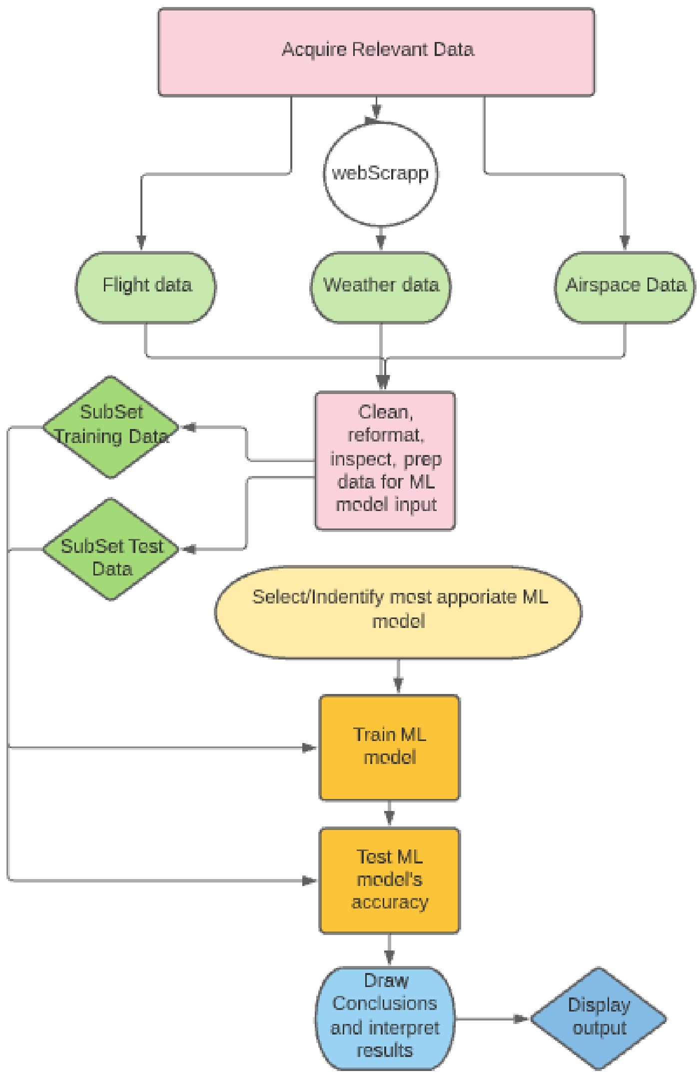

Figure 1. General Project Outline

At the moment I have yet to research and identify the most appropriate model for the task. I am uncertain if it is too ambitious as it is, given the large amount of data prepping, cleaning, and reformatting that needs to done before I can test the ML. Figure 1 demonstrates what I plan to do in order to accomplish the project. I believe large amounts of time will be required at most of these steps and may slow down the process since I cannot train a ML without first prepping the input data.

Results

I aim to map my results with the actual flight path so that you can see an example of the ML predictions next to the real flight path and with any relevant weather and airspace data.

Results Verification

I will measure my results accuracy by comparing the predictions with the recorded flight paths from the ADS-B dataset.

Discussion and Conclusion

I have yet to do the project, but I have a lot to learn. I have never web scrapped data before, and assume it will take a little time to code and attain the desired output. I have an introductory understanding of Machine Learning and still need to research what models are most affective with my given data and directive. Working in 3D with temporally sensitive data will be more difficult to map and format as input for a ML model. I am very used to 2D projections and am still trying to best figure out how to feed in 3D data to a ML. I have yet to do any of the steps mentioned above and anticipate there will be lots of issues in running and training the model (possible slow computing time or memory overloading).

References

Sakyi-Gyinae, M. K. (2019). *A Machine Learning Approach to Evaluating Aircraft Deviations from Planned Routes* (p. 73) [Thesis]. https://repository.tudelft.nl/islandora/object/uuid%3A274b4386-539a-4193-80e9-f120c8d4832e

Self-score

Category	Description	Points Possible	Score
Structural Elements	All elements of a lab report are included **(2 points each)**: Title, Notice: Dr. Bryan Runck, Author, Project Repository, Date, Abstract, Problem Statement, Input Data w/ tables, Methods w/ Data, Flow Diagrams, Results, Results Verification, Discussion and Conclusion, References in common format, Self-score	28	**28**
Clarity of Content	Each element above is executed at a professional level so that someone can understand the goal, data, methods, results, and their validity and implications in a 5 minute reading at a cursory-level, and in a 30 minute meeting at a deep level **(12 points)**. There is a clear connection from data to results to discussion and conclusion **(12 points)**.	24	**22**
Reproducibility	Results are completely reproducible by someone with basic GIS training. There is no ambiguity in data flow or rationale for data operations. Every step is documented and justified.	28	**26**
Verification	Results are correct in that they have been verified in comparison to some standard. The standard is clearly stated **(10 points)**, the method of comparison is clearly stated **(5 points)**, and the result of verification is clearly stated **(5 points)**.	20	**16**
		100	**92**