FLIGHT PREDICTION

Introduction

Background:
From unreliable sensors to emergencies and bad weather, when ground control loses contact with airplanes, it spells disaster for ground control, airlines, and government agencies.

Motivation:
Predicting a flight's destination is an important challenge in motion analysis.

Methodology:
Picture historical flights with paint cans, filling buckets with its destination's color along its course. By comparing the colors of a flight of interest's buckets, we can predict its destination. (See 3)

Tools and Resources used:
Tracktable (trajectories), Anvil (Large Data Processing), Python

[1]: Problem:
If a flight disappears, our model should determine its destination from historical data. In the example above, the origin airport is known: LAX. If, halfway through, it loses connection to the ground. Our model should predict its destination, JFK, in a list of Top 20 airports.

[2]: Process Diagram:
Our process involves creating and filling a grid with historical data, determining what cells a flight of interest overlaps with, and adding weighted occurrences of the destination list in each cell to produce a final list of predicted destinations.

[3]: Overlapping Grid Cells Raster:
As a flight (paint cans) travels from LAX-JFK, it will overlap cells (buckets) in the national grid.

[4]: Content of Cells:
Using historical data, flights that pass through a cell have their destinations stored inside the cell; this is accessed and counted when the flight of interest overlaps that same cell.

[5]: Combine (Weighting):
After finding a cell that overlaps, weighing the destination count in it gives a more accurate prediction.

[6]: Unweighted Final Results:
Example: For flights with 5% remaining, 40% of them were in the top 20.

[7]: Final Weighted Results: Predictions for 500,000 flights at every percentage of trajectory remaining after truncation.

[8]: Colored US Map result visualization: The top predicted airport is in green, second and third prediction in yellow, and fourth and fifth prediction in blue.

[9]: Final Result: Increased Top 1 and 3 prediction accuracy while increasing the percent of flight remaining.

Conclusion

Conclusion:
Our forecasting accuracy for identifying the top airport stands at 31% and Top 20 Leading Airports it is 75%. Major US airports are so dominant in air traffic that randomly picking from the top airports often lands you within the Top 20 predictions.

Future Work:
What makes flights to smaller airports different from those to larger ones? How can we predict their destinations?

Acknowledgements:
Dr. Andy Wilson, Dr. Kat Ward, and The Data Mine staff!

References:
Tracktable and Python Documentation