ONNX Implementation for Air Crash Probability Analysis with Flight Data EASA ATC public data.

This project focuses on the implementation of ONNX for predicting the likelihood of air crashes caused by unexpected flight path fluctuations. The dataset utilized originates from EASA Air Traffic control public data, encompassing information from various Air Traffic Control sources.

How to Run:

1. Create a virtual environment.

2. Activate the virtual environment.

3. Install all the requirements using the `requirements.txt` file:

   pip install -r requirements.txt

4. Run `ATC.ipynb` to perform exploratory data analysis (EDA) and data exploration. This notebook includes steps for data pre-processing and initial analysis.

5. Execute `onnx_1.ipynb` to develop the predictive model. This notebook focuses on the first level of accident probability prediction using critical data such as latitude, longitude, and airspeeds.

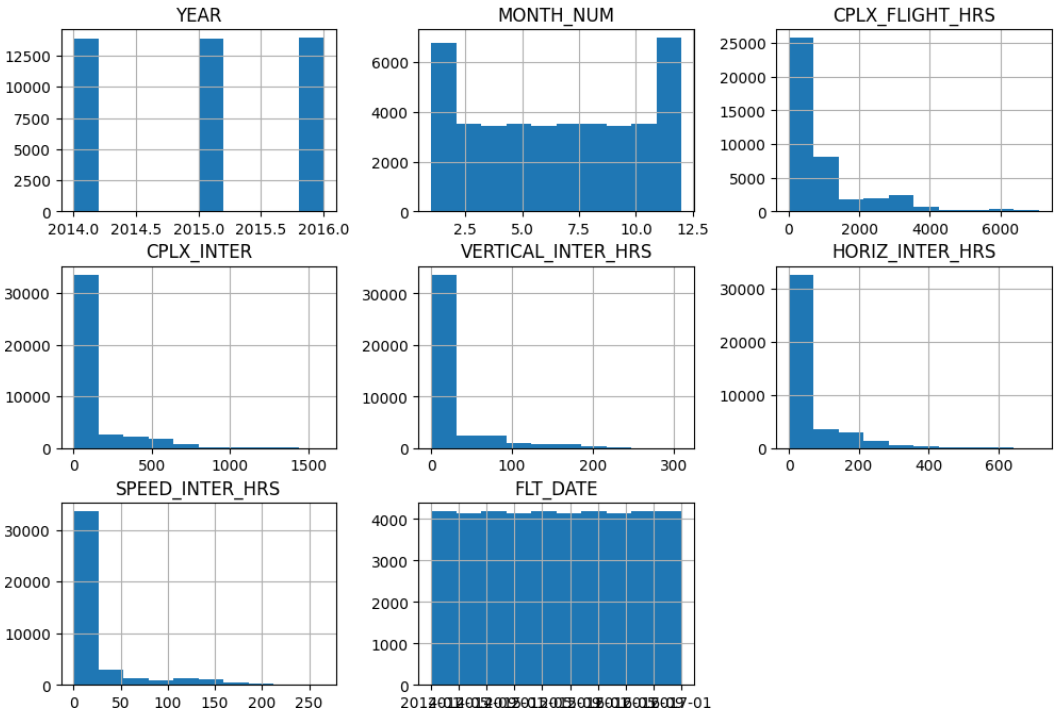
6. Launch `ONNX_2.ipynb` for a comprehensive analysis of flight path fluctuations and identification of high-risk geographic locations. This notebook also explores the potential for issuing "Mayday" messages in high-risk zones during flight.

Future Scope:
The current system can be developed as streaming system with live streaming data. For which stream processing and batch processing of data can be done using Apache spark streaming and Kafka which can process live data in our model for predictions.

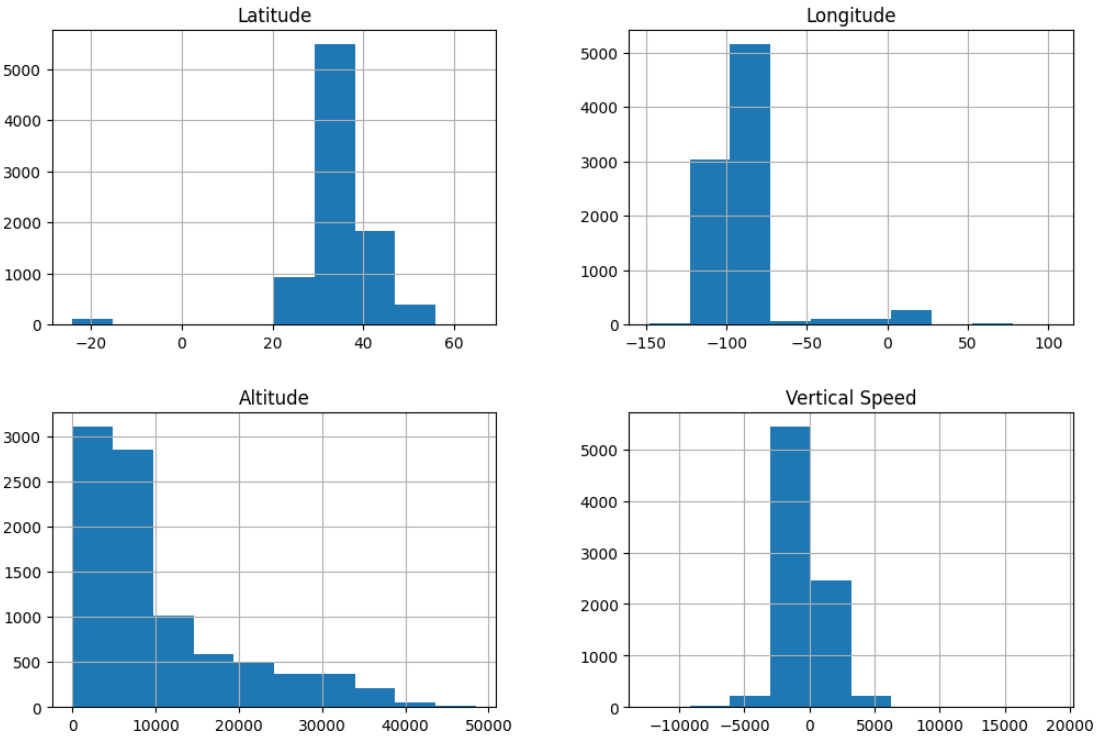
Runs and Plots from Project Analysis.

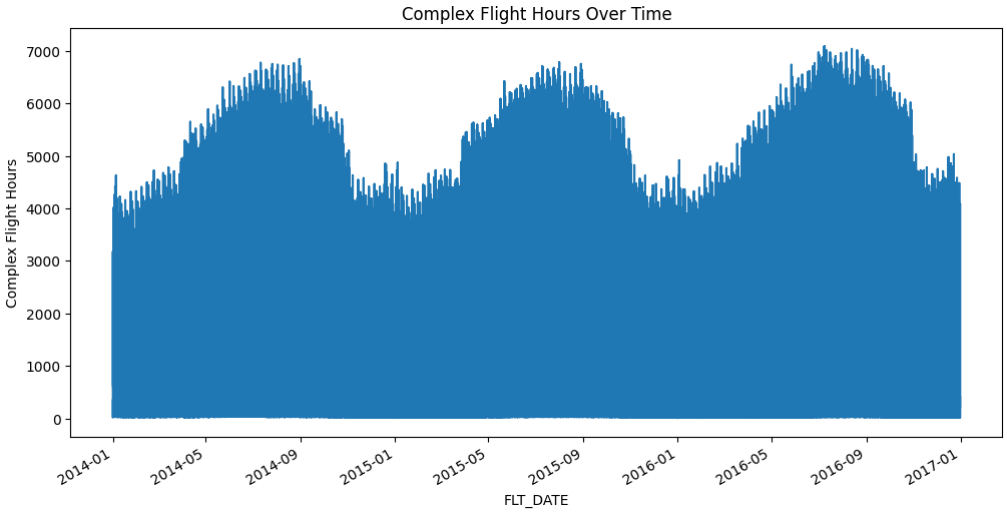
EDA Steps:

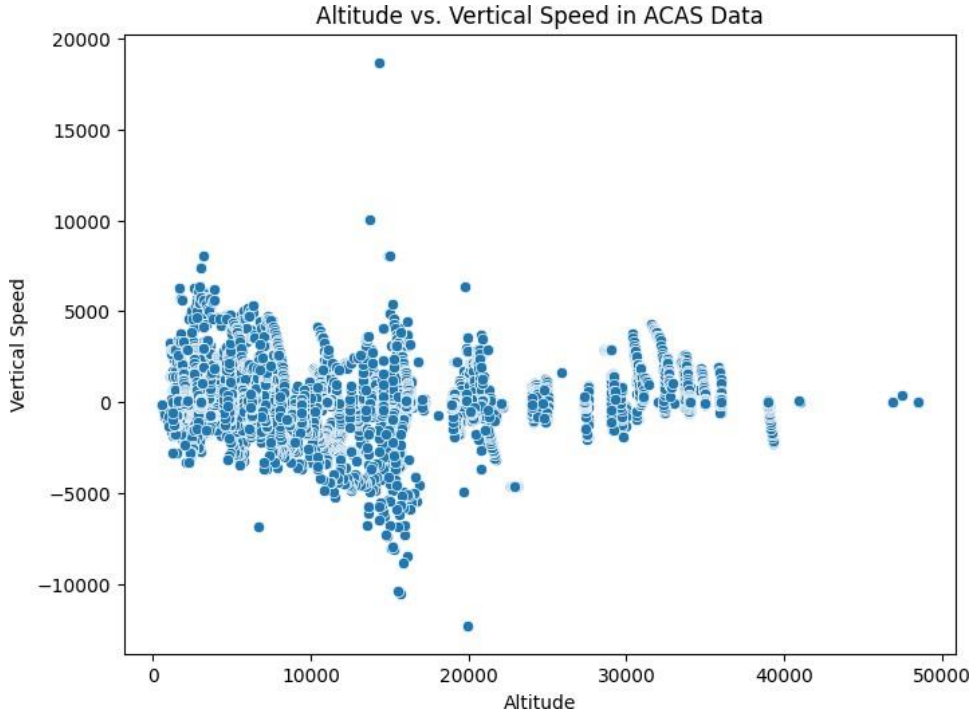
Distribution of TRF Complexity Features

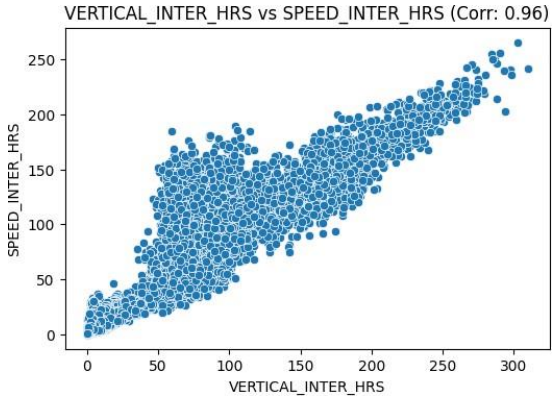


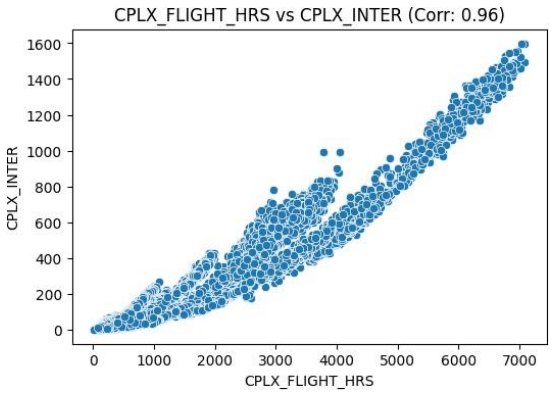
Distribution of ACAS Features

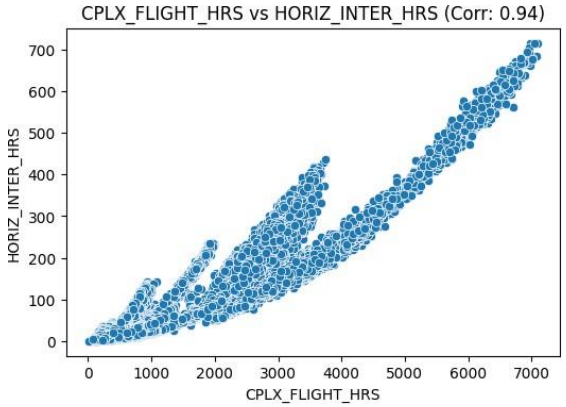


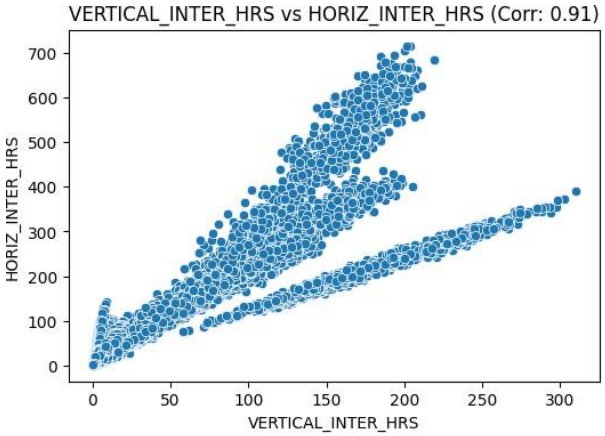

Complex Flight Hours Over Time
Complex Flight Hours
7000
6000
5000
4000
3000
2000
1000
0
2014-01
2014-05
2014-09
2015-01
2015-05
2015-09
2016-01
2016-05
2016-09
2017-01
FLT_DATE

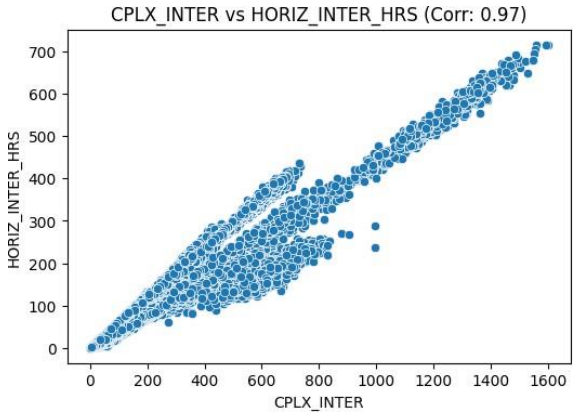

Altitude vs. Vertical Speed in ACAS Data
Vertical Speed
20000
15000
10000
5000
0
−5000
−10000
0
10000
20000
30000
40000
50000
Altitude

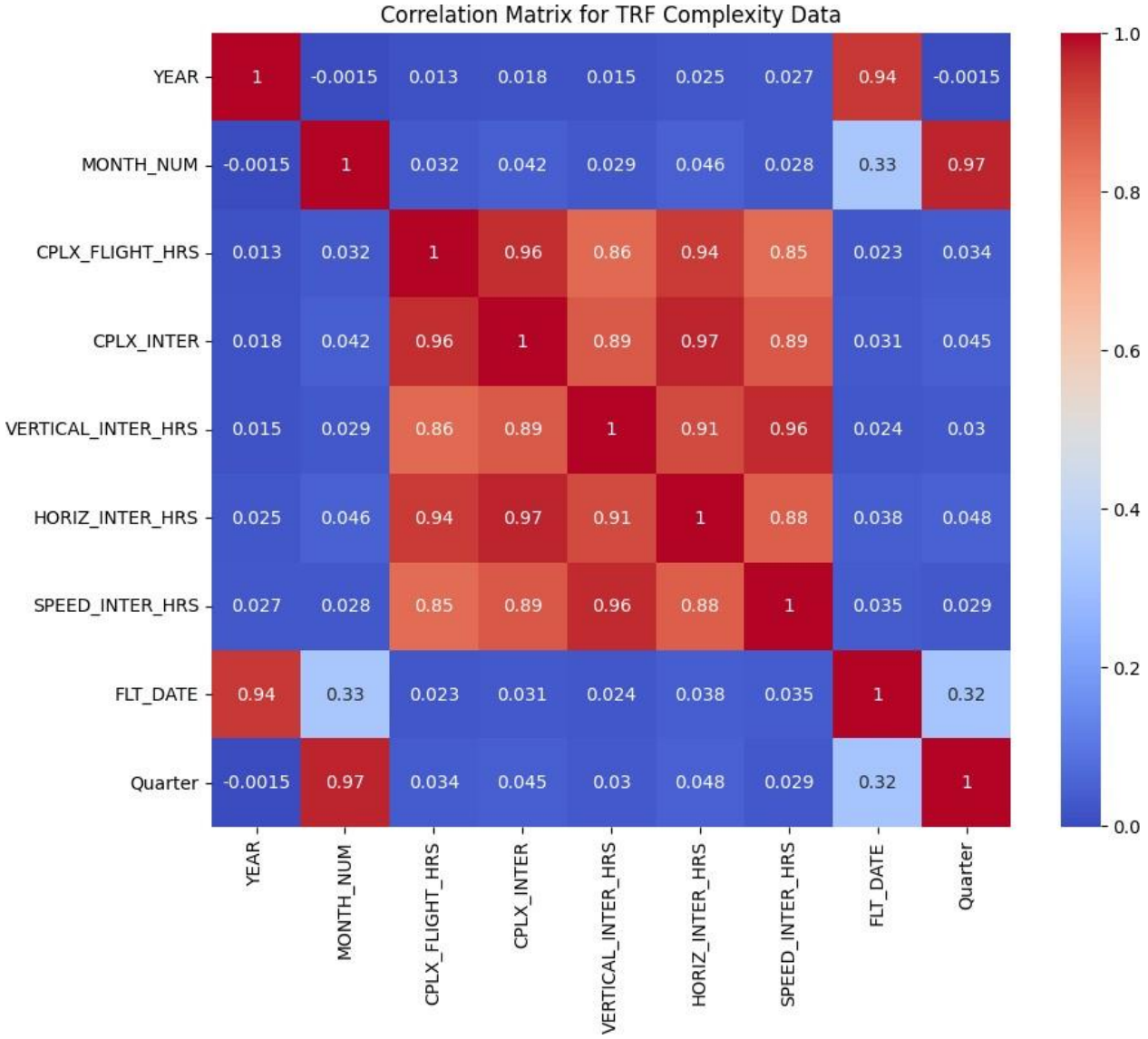

VERTICAL_INTER_HRS vs SPEED_INTER_HRS (Corr: 0.96)
SPEED_INTER_HRS
VERTICAL_INTER_HRS

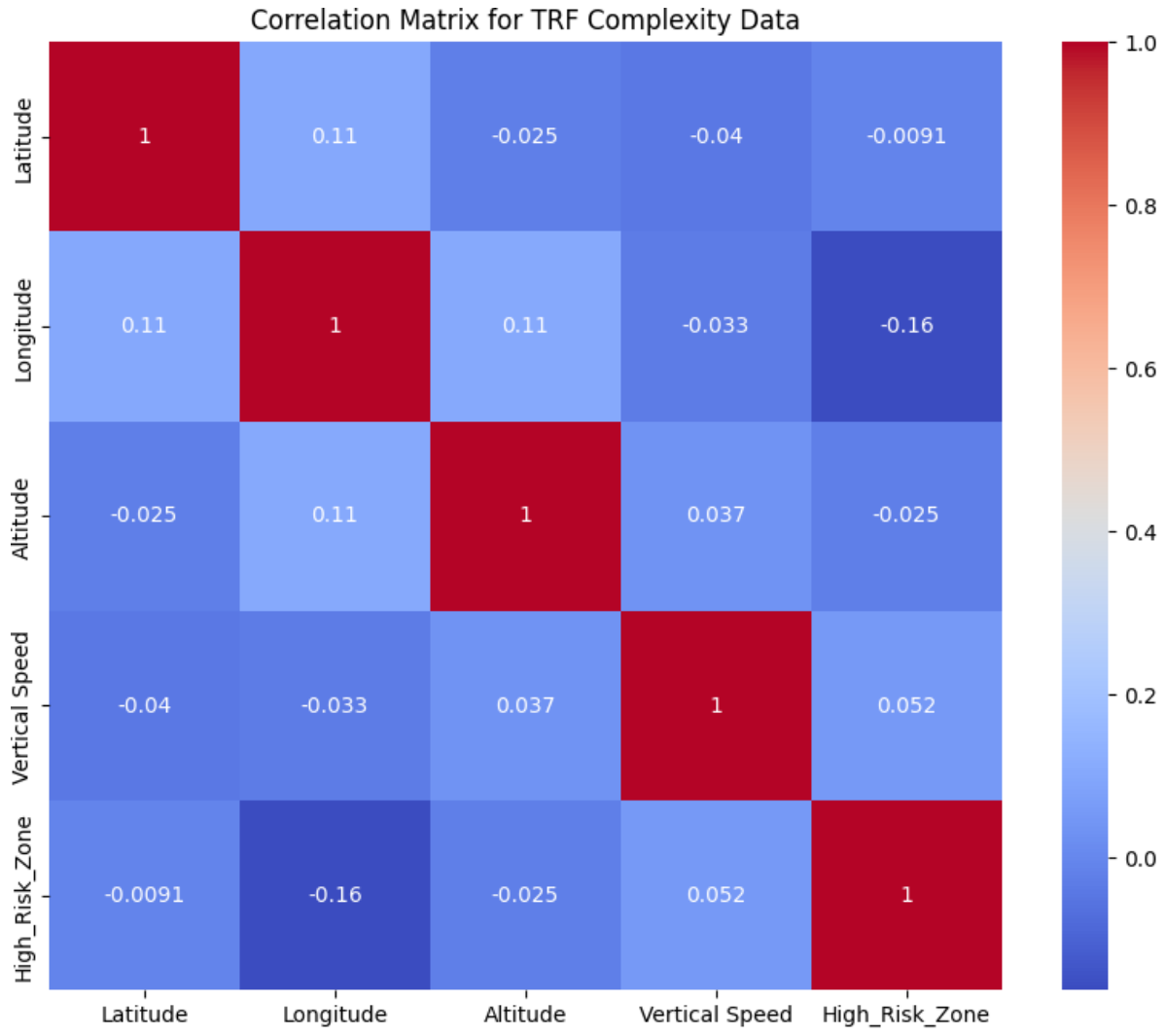

CPLX_FLIGHT_HRS vs CPLX_INTER (Corr: 0.96)
CPLX_INTER
CPLX_FLIGHT_HRS

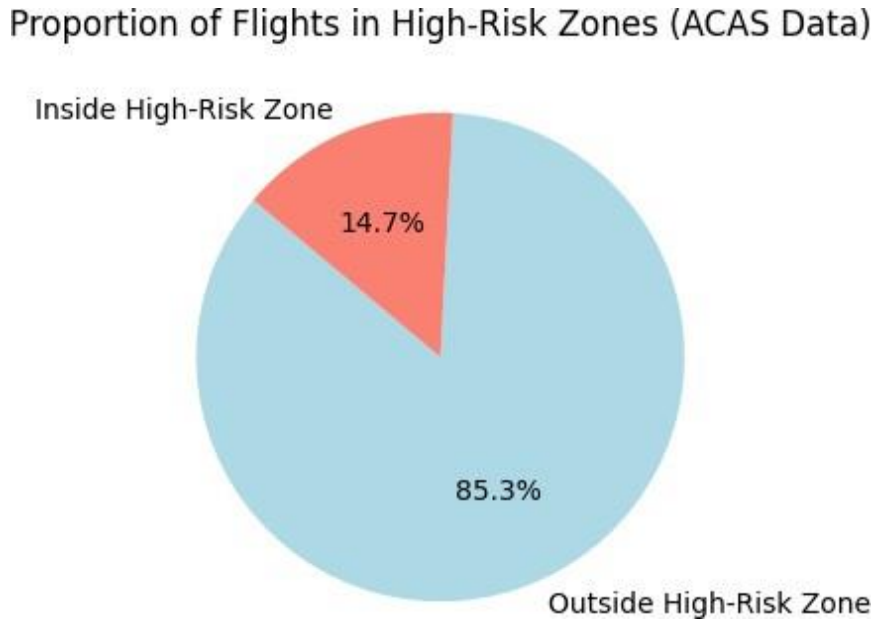

CPLX_FLIGHT_HRS vs HORIZ_INTER_HRS (Corr: 0.94)
HORIZ_INTER_HRS
CPLX_FLIGHT_HRS

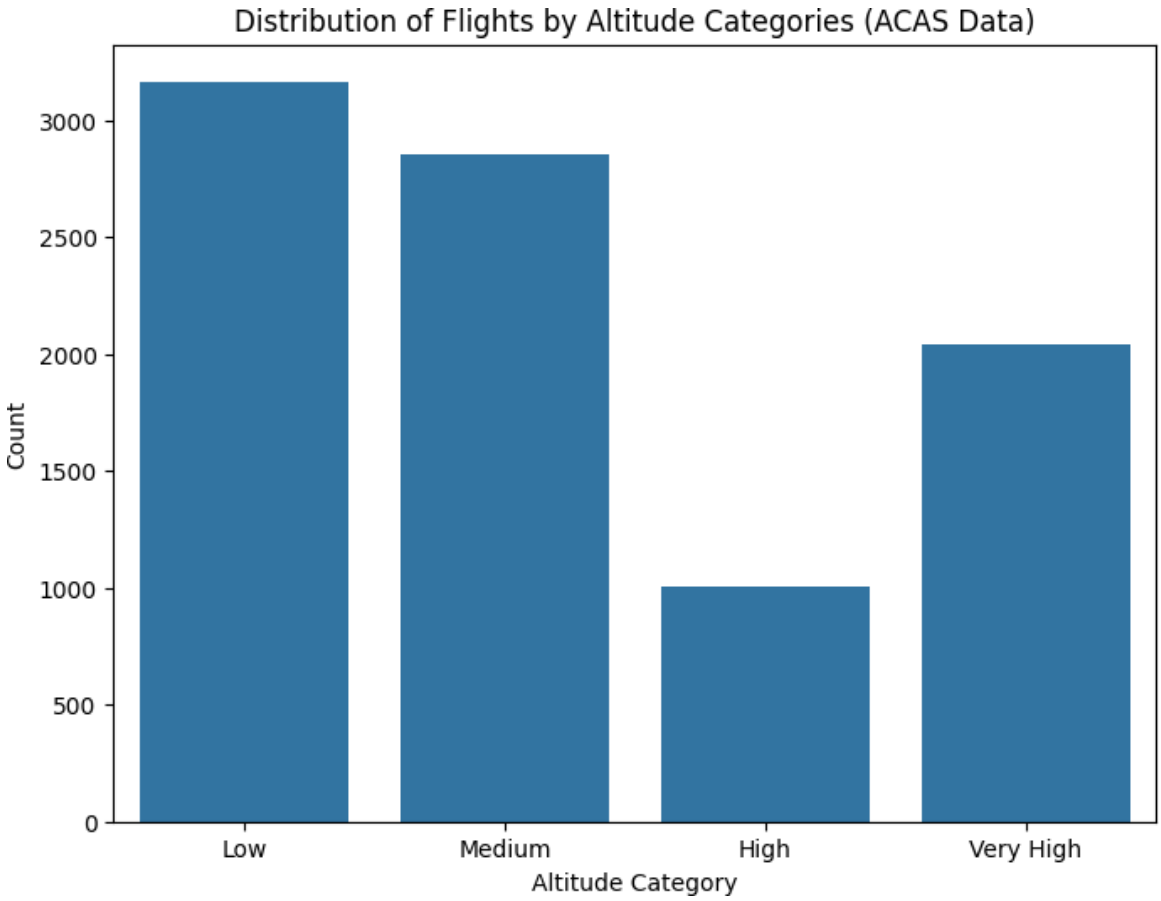

VERTICAL_INTER_HRS vs HORIZ_INTER_HRS (Corr: 0.91)
HORIZ_INTER_HRS
VERTICAL_INTER_HRS

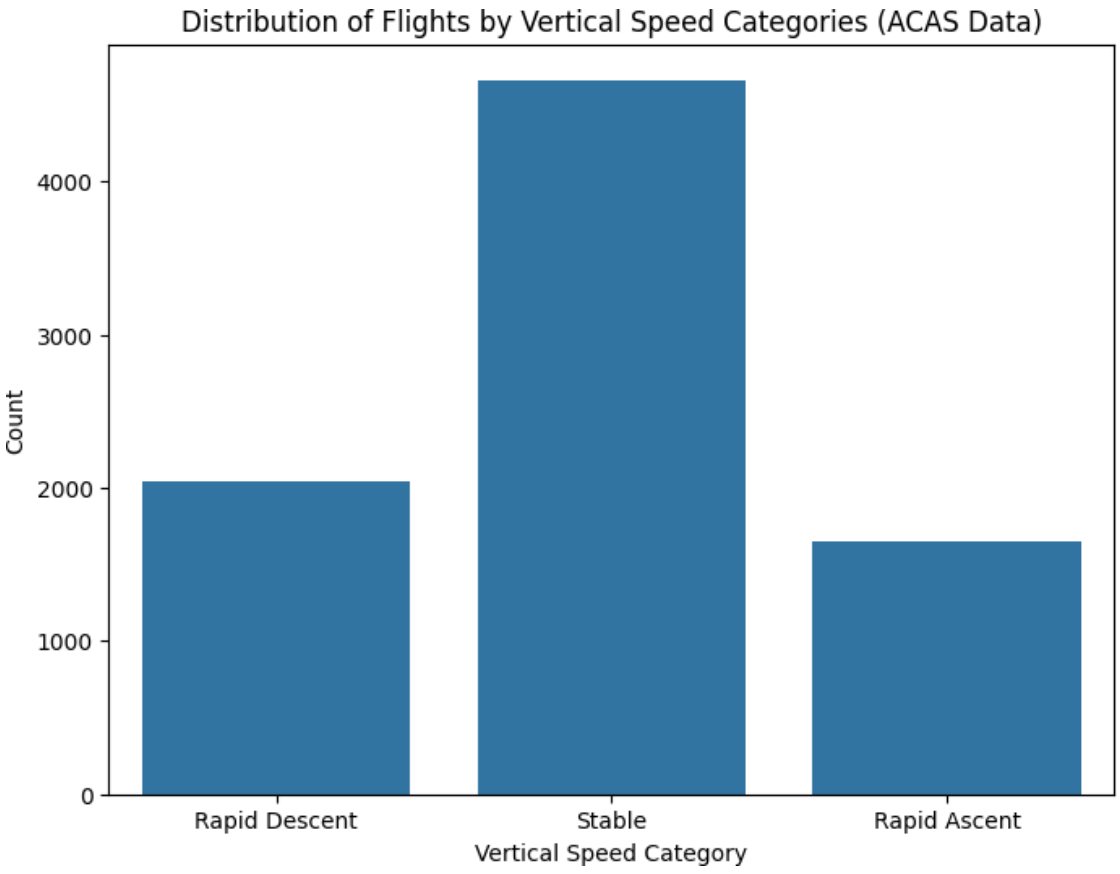

CPLX_INTER vs HORIZ_INTER_HRS (Corr: 0.97)
HORIZ_INTER_HRS
CPLX_INTER


Correlation Matrix for TRF Complexity Data
YEAR 1 -0.0015 0.013 0.018 0.015 0.025 0.027 0.94 -0.0015
MONTH_NUM -0.0015 1 0.032 0.042 0.029 0.046 0.028 0.33 0.97
CPLX_FLIGHT_HRS 0.013 0.032 1 0.96 0.86 0.94 0.85 0.023 0.034
CPLX_INTER 0.018 0.042 0.96 1 0.89 0.97 0.89 0.031 0.045
VERTICAL_INTER_HRS 0.015 0.029 0.86 0.89 1 0.91 0.96 0.024 0.03
HORIZ_INTER_HRS 0.025 0.046 0.94 0.97 0.91 1 0.88 0.038 0.048
SPEED_INTER_HRS 0.027 0.028 0.85 0.89 0.96 0.88 1 0.035 0.029
FLT_DATE 0.94 0.33 0.023 0.031 0.024 0.038 0.035 1 0.32
Quarter -0.0015 0.97 0.034 0.045 0.03 0.048 0.029 0.32 1
YEAR
MONTH_NUM
CPLX_FLIGHT_HRS
CPLX_INTER
VERTICAL_INTER_HRS
HORIZ_INTER_HRS
SPEED_INTER_HRS
FLT_DATE
Quarter
1.0
0.8
0.6
0.4
0.2
0.0

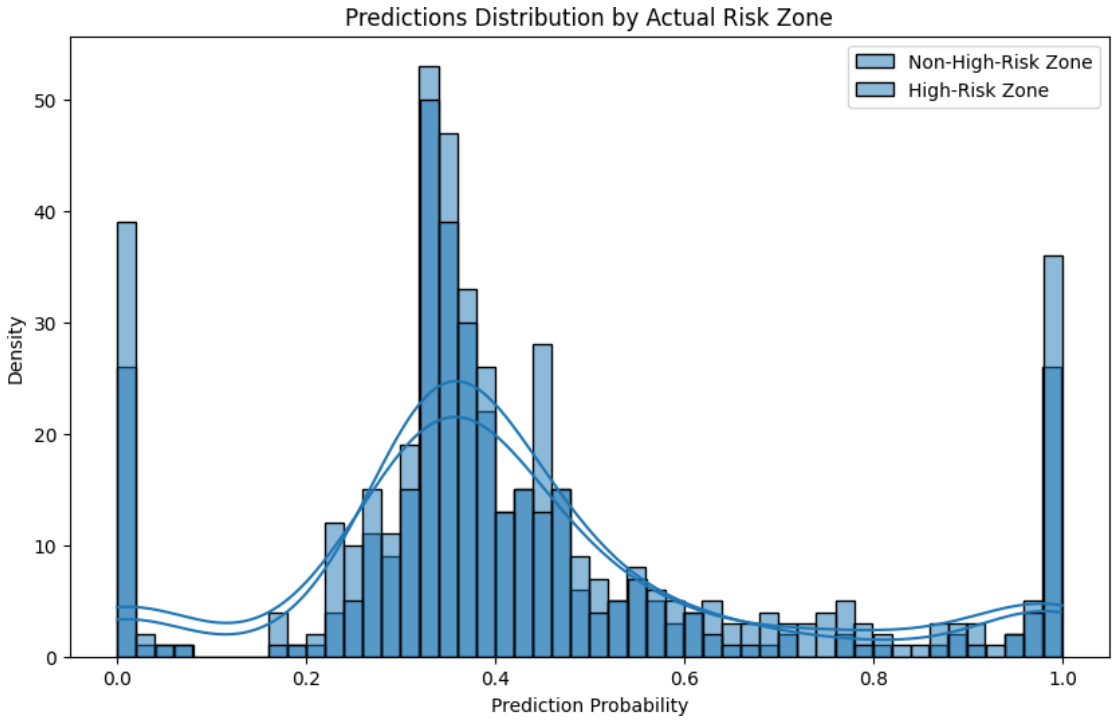

Correlation Matrix for TRF Complexity Data
Latitude
Longitude
Altitude
Vertical Speed
High_Risk_Zone
1
0.11
-0.025
-0.04
-0.0091
0.11
1
0.11
-0.033
-0.16
-0.025
0.11
1
0.037
-0.025
-0.04
-0.033
0.037
1
0.052
-0.0091
-0.16
-0.025
0.052
1
1.0
0.8
0.6
0.4
0.2
0.0

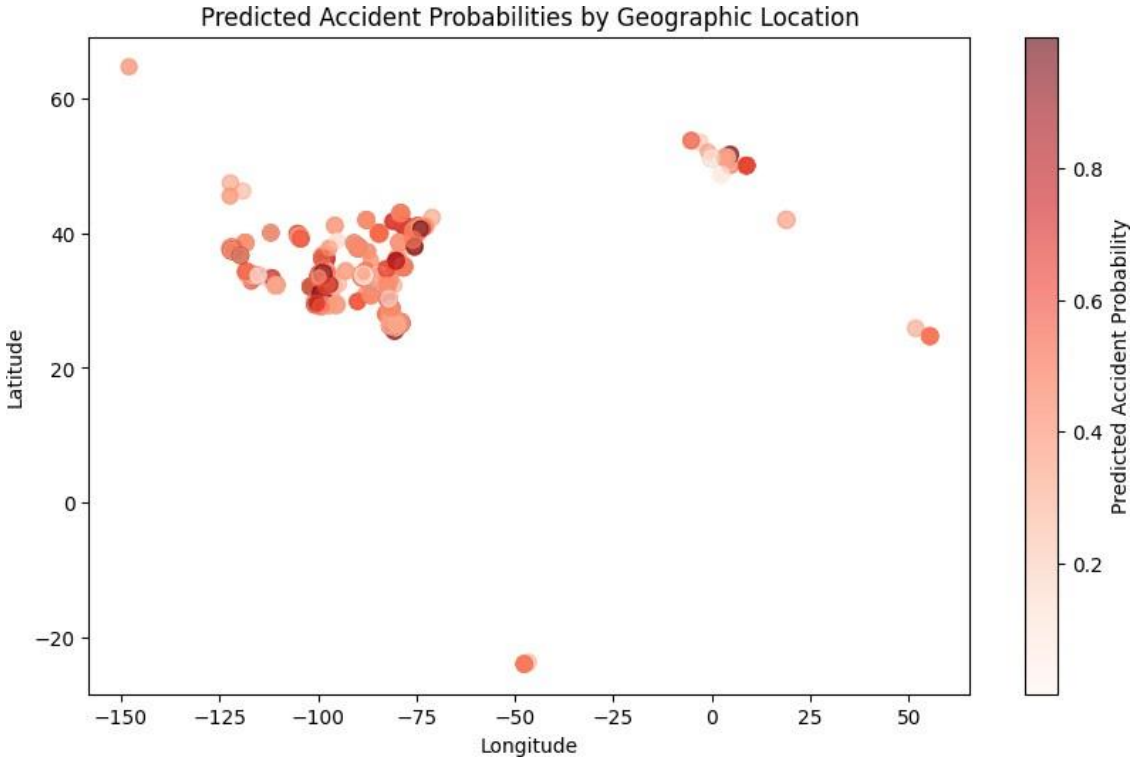

Proportion of Flights in High-Risk Zones (ACAS Data)
Inside High-Risk Zone
14.7%
85.3%
Outside High-Risk Zone

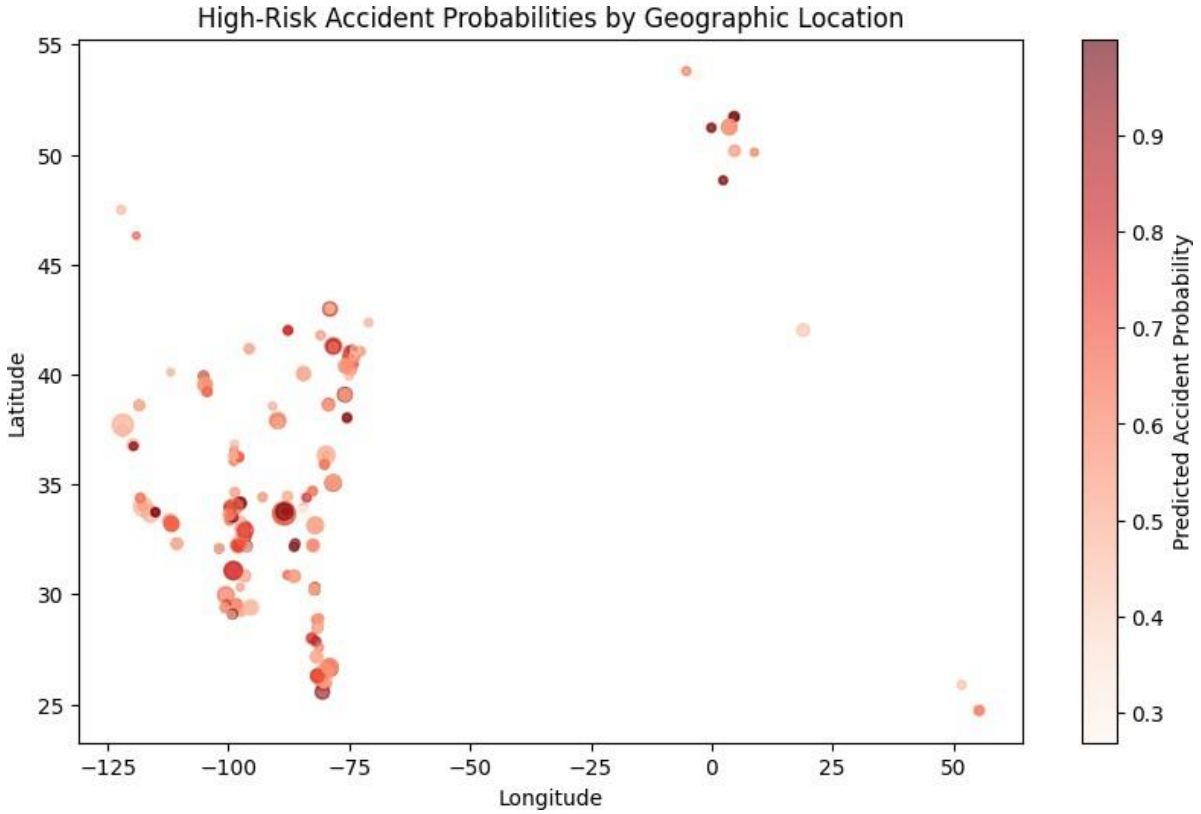

Distribution of Flights by Altitude Categories (ACAS Data)
Count
3000
2500
2000
1500
1000
500
0
Low
Medium
High
Very High
Altitude Category

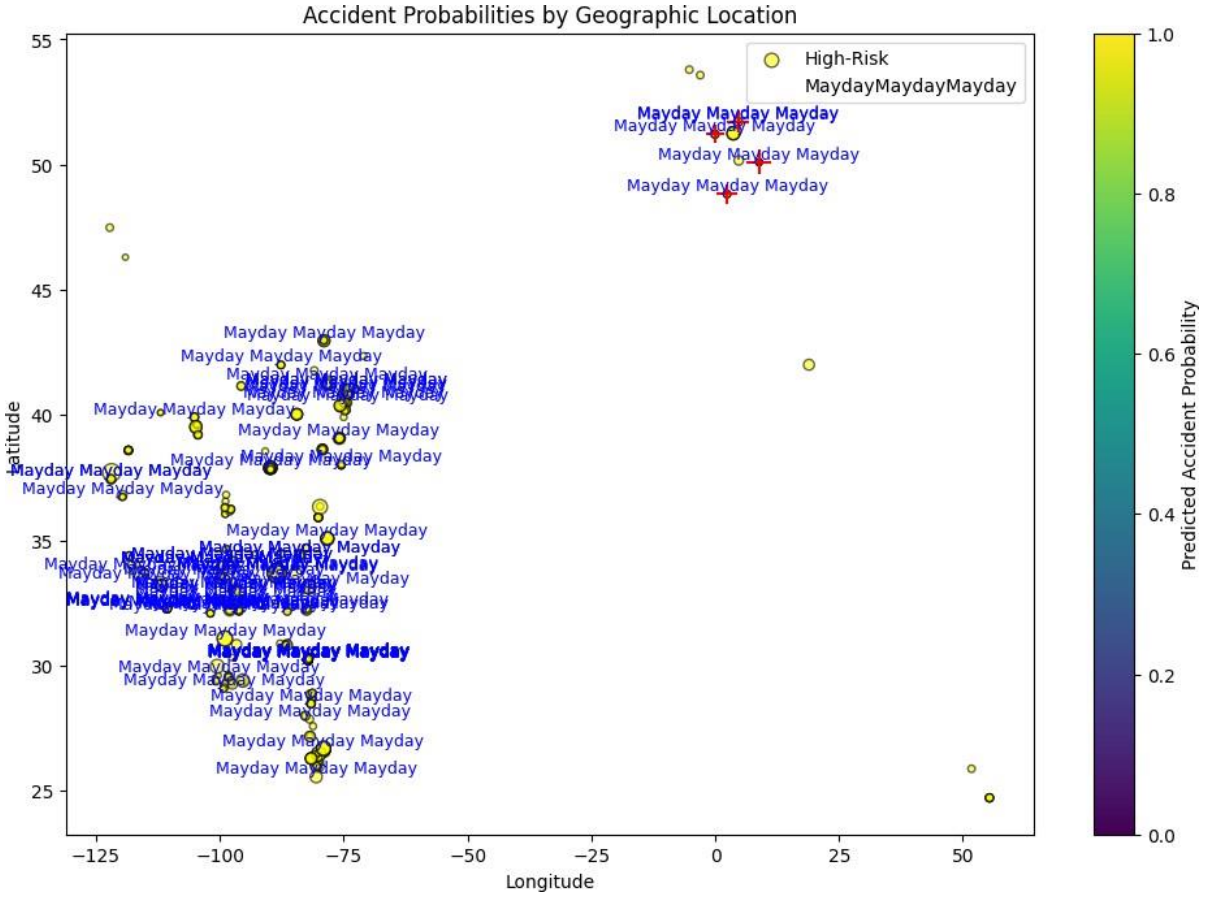

Distribution of Flights by Vertical Speed Categories (ACAS Data)
Count
4000
3000
2000
1000
0
Rapid Descent
Stable
Rapid Ascent
Vertical Speed Category

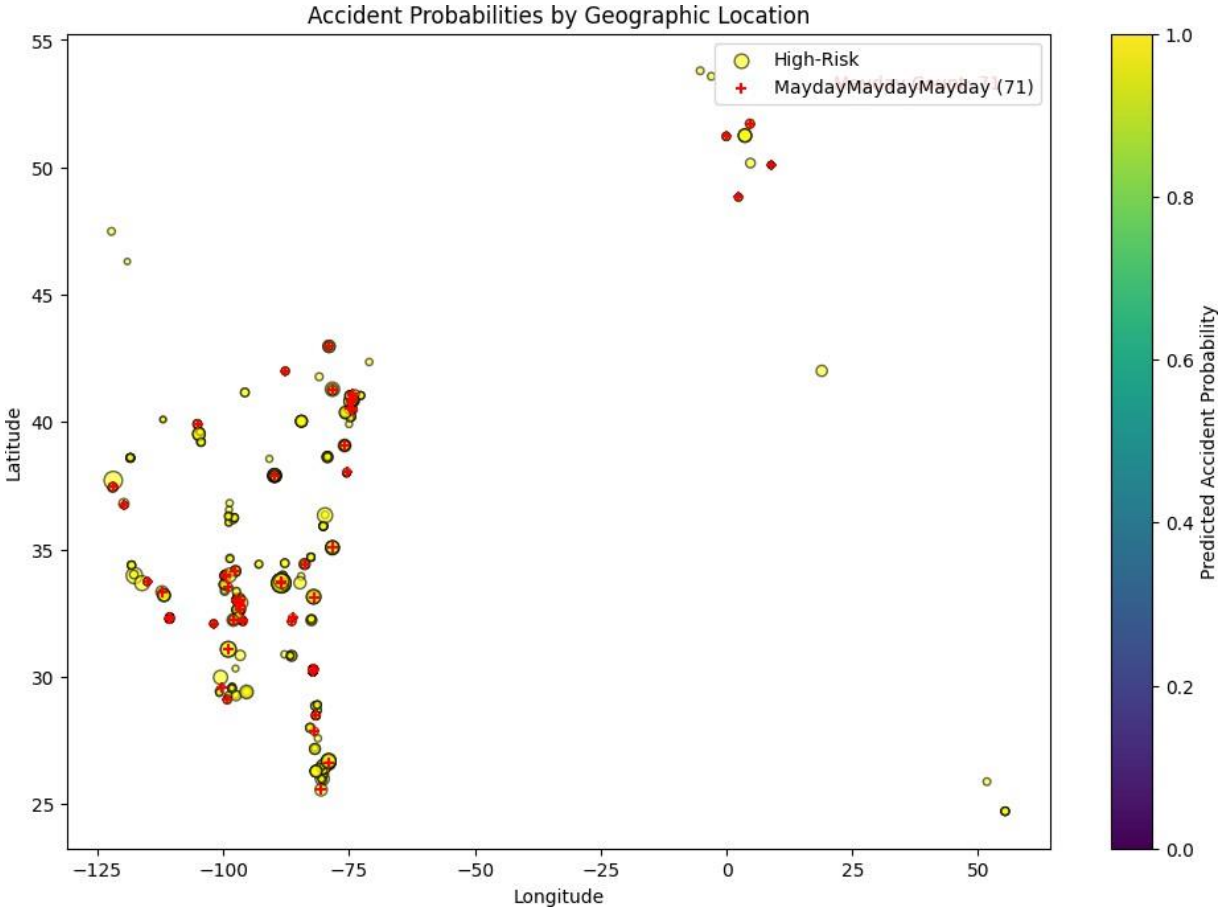
Visualizations' with ONNX implementation:






High-Risk Accident Probabilities by Geographic Location
Latitude
Longitude
Predicted Accident Probability


Accident Probabilities by Geographic Location
High-Risk
MaydayMaydayMayday
Latitude
Longitude
Predicted Accident Probability


Accident Probabilities by Geographic Location
High-Risk
MaydayMaydayMayday (71)
Latitude
Longitude
Predicted Accident Probability
55
50
45
40
35
30
25
−125
−100
−75
−50
−25
0
25
50
1.0
0.8
0.6
0.4
0.2
0.0