M3 Global Hackathon

PROJECT TITLE

Mixer Transformer with Context-Time Attention for Small Vessel Route Prediction

WHO WE ARE



Md Mahmuddun Nabi Murad
University of South Florida
mmurad@usf.edu



Yasin Yilmaz
University of South Florida
yasiny@usf.edu



Justin McMillen
University of South Florida
jmcmillen@usf.edu

Mixer Transformer with Context-Time Attention for Small Vessel Route Prediction

MAIN CHALLENGES

- **Filtering Out Smaller Routes**

  The dataset contains routes with lengths varying from 20 to 1000. The input sequence length is 60, and the prediction length is 120, though predictions may terminate before reaching 120 steps. A threshold is needed to remove smaller routes that don't contribute meaningful data for the model.

- **Representing the Endpoint for Learning**

  To predict the endpoint effectively, it is necessary to represent it in a way that machine learning models can learn efficiently. This includes differentiating the actual endpoint from waiting positions where the vessel remains stationary.

- **Feature engineering**

  The dataset contains profile ID, vessel ID, timestamp, latitude, and longitude. Challenges include identifying relevant features, engineering new features, and distinguishing between end positions and waiting states to improve model accuracy.

Mixer Transformer with Context-Time Attention for Small Vessel Route Prediction

METHODOLOGY

❑ **Data pre-processing**

- ➢ **Step 1:** Removing the duplicate values
- ➢ **Step 2:** Interpolating the missing values
- ➢ **Step 3:** Creating the additional features
- ➢ **Step 4:** Padding and creating observation and prediction sequences

$$M_{sin} = \sin\left(\frac{2\pi \times M}{60}\right) \qquad H_{sin} = \sin\left(\frac{2\pi \times H}{24}\right) \qquad D_{sin} = \sin\left(\frac{2\pi \times D}{7}\right)$$

$$M_{cos} = \cos\left(\frac{2\pi \times M}{60}\right) \qquad H_{cos} = \cos\left(\frac{2\pi \times H}{24}\right) \qquad D_{cos} = \cos\left(\frac{2\pi \times D}{7}\right)$$

Submission Train Dataset:

	Total Sequences	Input features	Output features
Model Train	728310	10	2
Model Test	242770	10	2
Model Validation	242770	10	2

Submission Test Dataset:

Total Sequences	Input features	Output features
1947	10	2

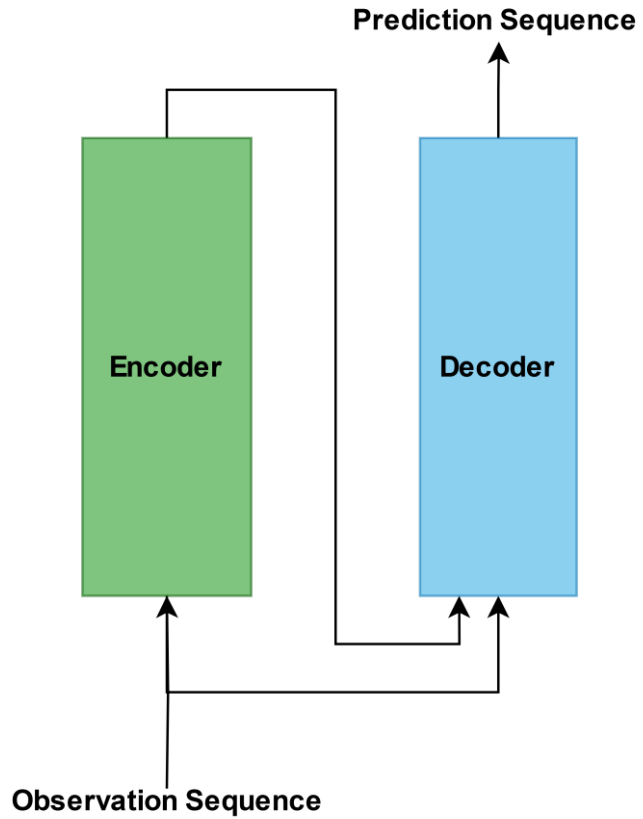
Input features: [Profile ID, M_sin, M_cos, H_sin, H_cos, D_sin, D_cos, Displacement, Latitude, Longitude]

Mixer Transformer with Context-Time Attention for Small Vessel Route Prediction

METHODOLOGY

❑ Route Forecasting Model

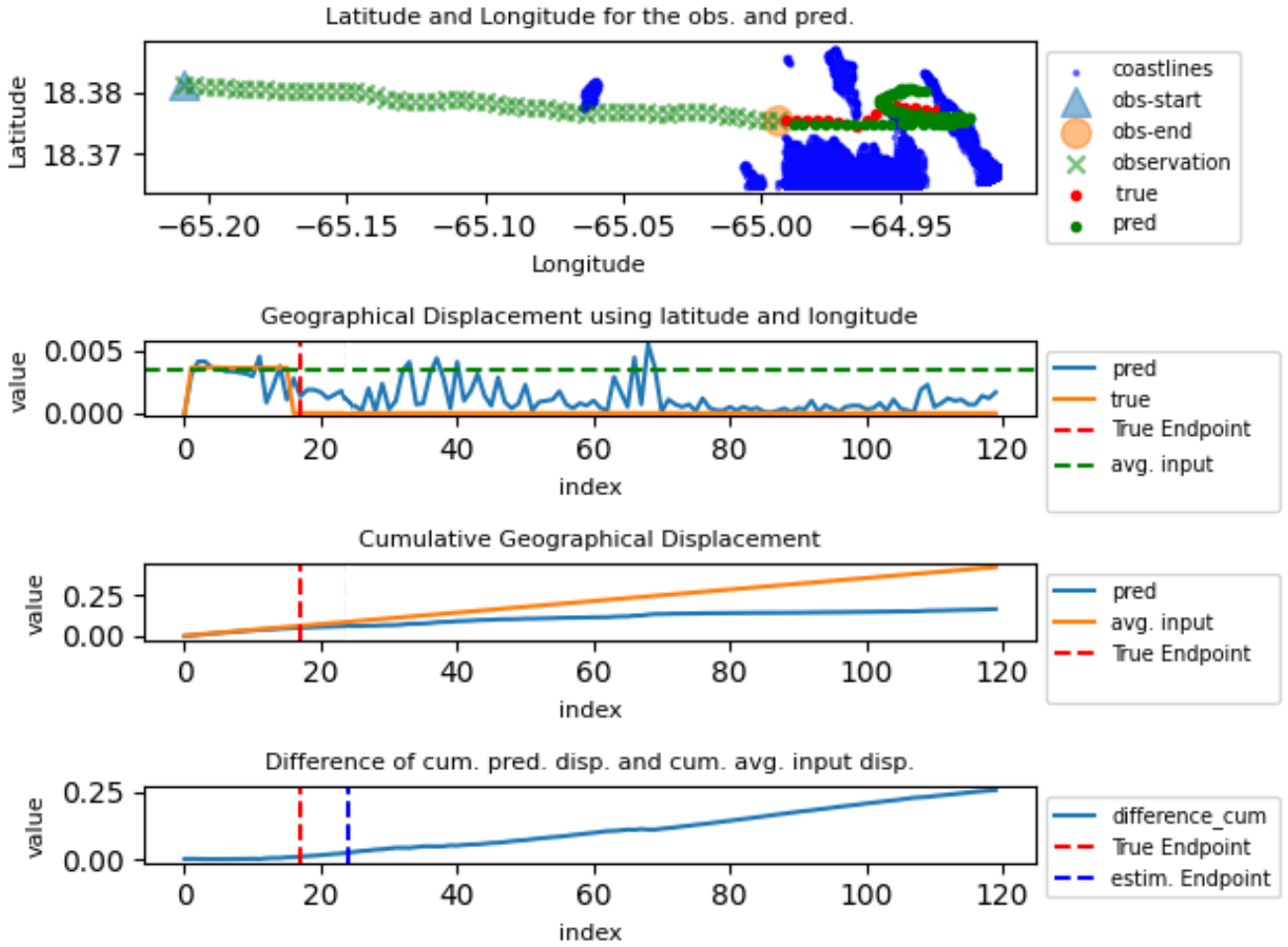
- Transformer model with encoder-decoder architecture.
- Input to the encoder and decoder is the observation sequence.
- Encoder includes Patching, Embedding, and Mixer modules.
- Decoder includes cross-attention and Mixer modules.
- Cross-attention to calculate the context-time attention.
- Mixer module includes a temporal mixer and an embedding mixer.



Mixer Transformer with Context-Time Attention for Small Vessel Route Prediction

METHODOLOGY

- **Endpoint Detection-1**
  - Derive the avg. input displacement from input sequence.
  - Derive predicted displacement from predicted latitude and longitude.
  - Accumulate the avg. input displacement and predicted displacement over the prediction window.
  - Applying a threshold on the difference between the avg. input displacement and predicted displacement.
  - Check the distance between the estimated endpoint and nearest land.



Mixer Transformer with Context-Time Attention for Small Vessel Route Prediction

METHODOLOGY

❑ Endpoint Detection-1

Step 1:

- Avg. Displacement of the input sequence: $d_{avg} \in \mathbb{R}$
- Predicted displacement: $\boldsymbol{d}_x \in \mathbb{R}^T$

Step 2:

- Cumulative avg. Displacement series: $\boldsymbol{C}_{avg} \in \mathbb{R}^T$
- Cumulative prediction displacement series: $\boldsymbol{C}_x \in \mathbb{R}^T$

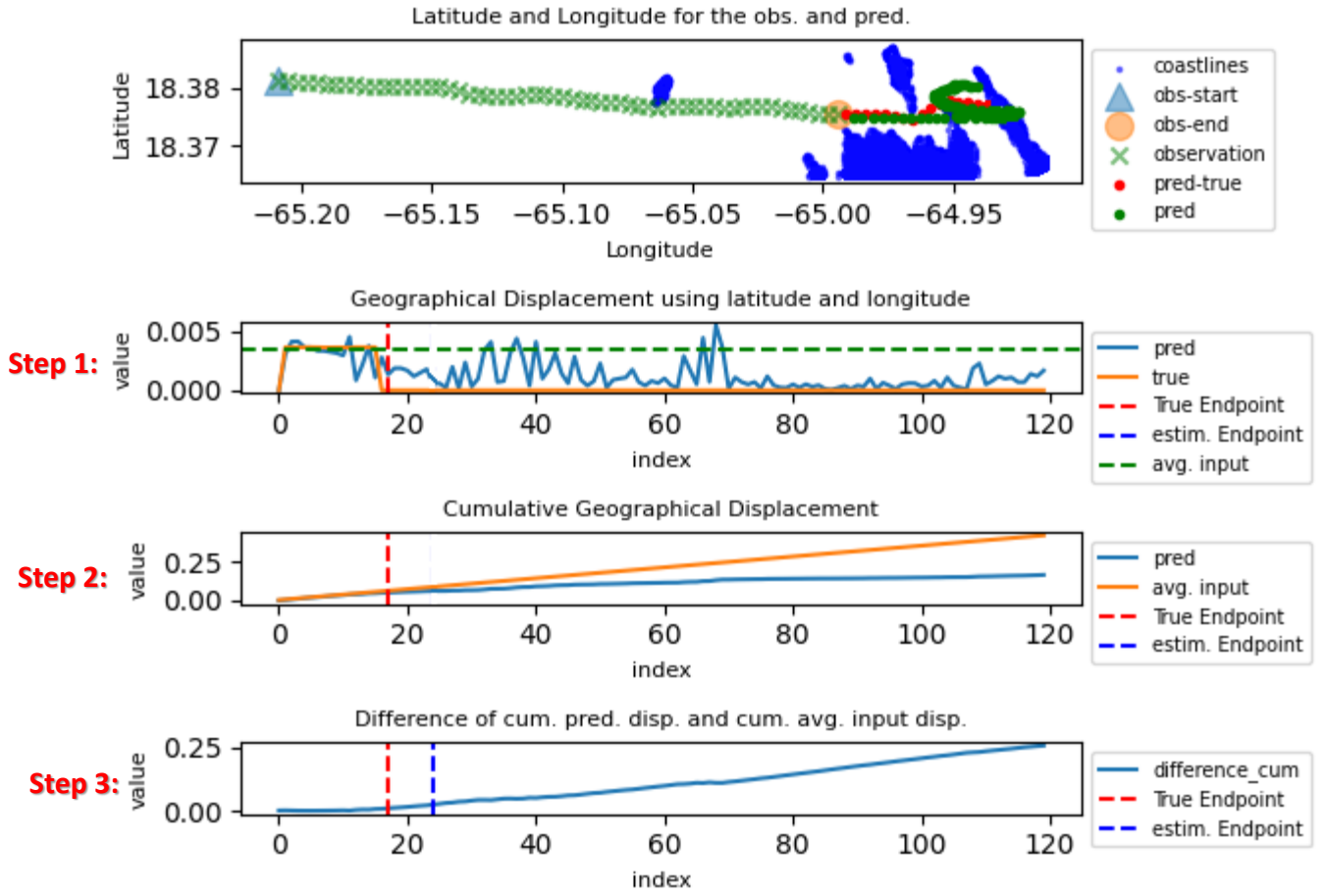
Step 3:

$$\boldsymbol{C}_\Delta = \max\{\boldsymbol{C}_{avg} - \boldsymbol{C}_x, 0\}$$

Initial endpoint: $$i^* = \arg \max_{i \in \{1,2,\ldots,T\}} \left\{ \boldsymbol{C}_\Delta^{(i)} < \delta \right\}$$

$$\delta = \boldsymbol{C}_\Delta^{(T)} \times h\%$$

Final endpoint: The endpoint should be closer to the land

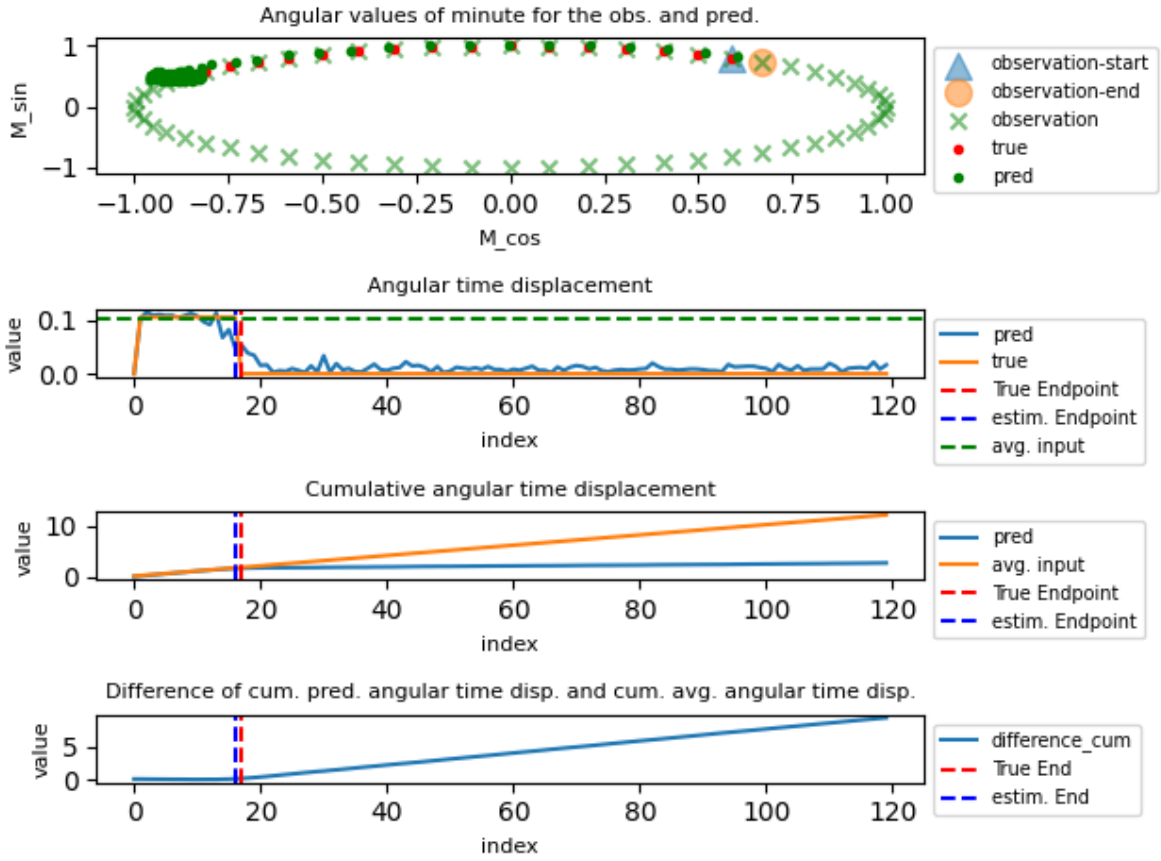


Mixer Transformer with Context-Time Attention for Small Vessel Route Prediction

METHODOLOGY

❑ Endpoint Detection-2

- Auxiliary model: output features *M_sin, M_cos*
- Followed the similar approaches as endpoint detection method-1.
- Benefit: No need to check the distance between the estimated endpoint and the nearest land.



Mixer Transformer with Context-Time Attention for Small Vessel Route Prediction

METHODOLOGY

❑ Endpoint Detection-2

Step 1:

- Avg. Displacement of the input sequence: $d_{\alpha.avg} \in \mathbb{R}$
- Predicted displacement: $\boldsymbol{d_{\alpha.x}} \in \mathbb{R}^T$

$$d_\alpha = \sqrt{(M_{sin}^{(i+1)} - M_{sin}^{(i)})^2 + (M_{cos}^{(i+1)} - M_{cos}^{(i)})^2}$$

Step 2:

- Cumulative avg. Displacement series: $\boldsymbol{C_{\alpha.avg}} \in \mathbb{R}^T$
- Cumulative prediction displacement series: $\boldsymbol{C_{\alpha.x}} \in \mathbb{R}^T$

Step 3:

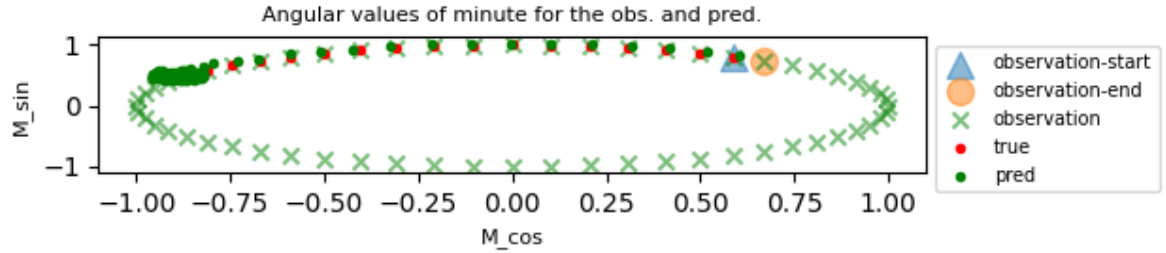
$$\boldsymbol{C}_\Delta = \max\{\boldsymbol{C}_{avg} - \boldsymbol{C}_x, 0\}$$

Final endpoint:

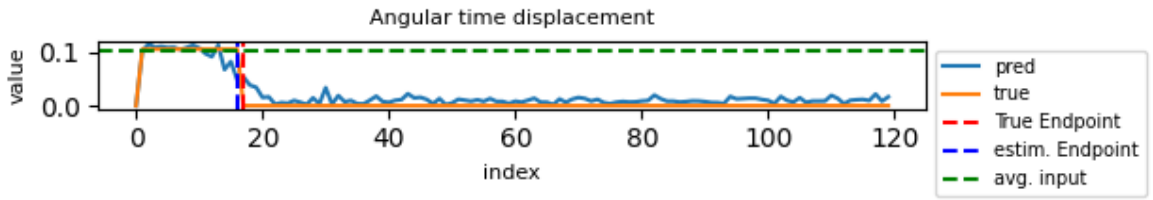
$$i^* = \arg \max_{i \in \{1,2,\ldots,T\}} \left\{ \boldsymbol{C}_\Delta^{(i)} < \delta \right\}$$

$$\delta = \boldsymbol{C}_\Delta^{(T)} \times h\%$$

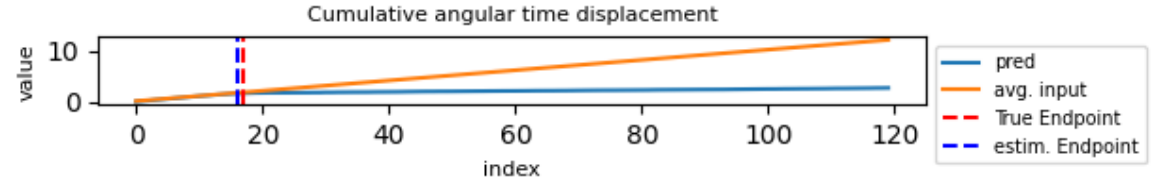
Step 1:



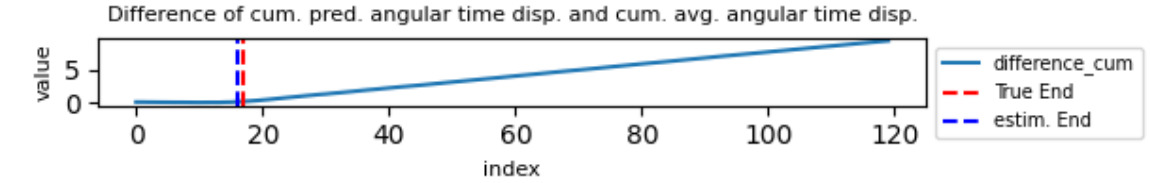
Step 1:



Step 2:



Step 3:



Mixer Transformer with Context-Time Attention for Small Vessel Route Prediction

RESULTS

When the predicted endpoint is similar to the true endpoint of a route:

Profile	Total Sequence	MSE	MAE	SMAPE	FDE	FD	AED	Termination Score(%)
1	26605	0.00007365	0.00704393	0.00021424	0.01290394	0.01541808	0.00576358	20.0
2	23127	0.00008860	0.00729869	0.00176160	0.01333685	0.01581776	0.00594290	20.0
3	27249	0.00009332	0.00747067	0.00068865	0.01398065	0.01662041	0.00609755	20.0
4	24103	0.00009145	0.00718019	0.00134623	0.01305474	0.01563586	0.00588066	20.0
5	25036	0.00009612	0.00772924	0.00071793	0.01430839	0.01710652	0.00631642	20.0
6	21494	0.00008497	0.00754831	0.00120433	0.01394757	0.01666818	0.00615726	20.0
7	26162	0.00007678	0.00719643	0.00103959	0.01323205	0.01570370	0.00587877	20.0
8	22485	0.00008748	0.00725741	0.00062658	0.01323773	0.01570538	0.00592770	20.0
9	24814	0.00009232	0.00761784	0.00113952	0.01425468	0.01696261	0.00624635	20.0
10	21695	0.00008462	0.00747250	0.00092197	0.01370754	0.01649944	0.00610740	20.0
Total:	242770							
Average:		0.00008693	0.00738152	0.00096606	0.01359641	0.01621379	0.00603186	20.0

- Assumption: Accurate detection of route endpoint.
- Full termination score with the expense of high MSE, MAE, SMAPE, FDE, FD, and AED.

Mixer Transformer with Context-Time Attention for Small Vessel Route Prediction

RESULTS

Performance Comparison:

- Termination score: Endpoint detection Method 2 performs well.
- Accuracy score: Endpoint detection method 1 performs well.
- Method 2 increases the termination score by 4.7%.
- Method 1 increases the route accuracy by 7.4%.

Overall performance with endpoint detection method 1:

Profile	Total Sequence	MSE	MAE	SMAPE	FDE	FD	AED	Termination Score(%)
1	26605	0.00006032	0.00666303	0.00020555	0.01154278	0.01399412	0.00543160	18.95
2	23127	0.00006462	0.00681019	0.00130099	0.01177562	0.01419493	0.00553162	18.83
3	27249	0.00007380	0.00702162	0.00061358	0.01249412	0.01501353	0.00571163	18.97
4	24103	0.00007045	0.00679522	0.00087912	0.01171879	0.01423883	0.00554204	18.95
5	25036	0.00007783	0.00733545	0.00061115	0.01288685	0.01559195	0.00597356	18.95
6	21494	0.00007551	0.00723493	0.00095458	0.01280026	0.01542004	0.00588557	18.95
7	26162	0.00006570	0.00684322	0.00079723	0.01195444	0.01432969	0.00556992	18.90
8	22485	0.00007272	0.00683353	0.00058915	0.01176749	0.01419179	0.00556033	18.80
9	24814	0.00007789	0.00722786	0.00090820	0.01278885	0.01543620	0.00590461	18.89
10	21695	0.00007436	0.00715759	0.00070165	0.01253834	0.01525856	0.00583489	18.89
Total:	242770							
Average		0.00007132	0.00699226	0.00075612	0.01222676	0.01476696	0.00569458	18.91

Overall performance with endpoint detection method 2:

Profile	Total Sequence	MSE	MAE	SMAPE	FDE	FD	AED	Termination Score(%)
1	26605	0.00007111	0.00694241	0.00021012	0.01270254	0.01516332	0.00568127	19.80
2	23127	0.00008079	0.00715744	0.00172702	0.01302962	0.01546201	0.00583281	19.78
3	27249	0.00008855	0.00735082	0.00064503	0.01372400	0.01630657	0.00599947	19.81
4	24103	0.00008113	0.00704901	0.00127190	0.01278550	0.01530791	0.00576814	19.81
5	25036	0.00009328	0.00762540	0.00070267	0.01409032	0.01683194	0.00623149	19.80
6	21494	0.00008203	0.00742843	0.00118573	0.01370405	0.01635827	0.00606004	19.80
7	26162	0.00007485	0.00710339	0.00103174	0.01305085	0.01546400	0.00580342	19.81
8	22485	0.00008227	0.00714711	0.00060953	0.01303930	0.01544676	0.00583823	19.79
9	24814	0.00008893	0.00749526	0.00112006	0.01398548	0.01662501	0.00614507	19.79
10	21695	0.00008046	0.00733547	0.00085017	0.01342062	0.01614145	0.00599654	19.79
Total:	242770							
Average:		0.00008234	0.00726347	0.00093540	0.01335323	0.01591073	0.00593565	19.80

OVERALL THOUGHTS & NEXT STEPS

- This dataset can be used as a baseline for route forecasting research works.
- In the future, we will integrate other information, i.e., coastline and port information.
- Right now, we are working to apply our model in the computer vision field to predict the route from the video information.
- We will publish our work with a citation of the M3 hackathon dataset.

For more information, please contact Md Mahmuddun Nabi Murad | mmurad@usf.edu