**Hurricane Modeling and Prediction**

**I. Summary of Problem Statement or Research Question**

Our goal is to apply machine learning techniques to model future hurricane trajectories based on historical data. The focus of our project is on broad, historical trends in hurricane trajectories, incorporating recorded data from storms as far back as 1850.

The research we do will be guided by the following questions, in order of priority:

* Are hurricane trajectories predictable by machine learning models?
* What are the noticeable correlations between trajectories and other variables, such as sea surface temperature?
* What is the projected impact of climate change (measured, for example, by sea surface temperature changes) on hurricane trajectories?

We hope to identify useful correlations that human-based models do not predict, and also learn a considerable amount about how to interface with and analyze vast amounts of data by applying machine learning techniques.

**II. Motivation**

Hurricanes are some of the most impressive weather phenomena on the planet and are capable of catastrophic amounts of destruction. Particularly in underdeveloped countries, lack of preparation can cripple economies and lead to thousands of deaths. We hope to contribute to the field of hurricane prediction, which can help make these deaths preventable.

**III. Plan of Action**

All of our group members have notifications for our Slack channel enabled, which will be our main channel of communication for the semester. We have at least one hour of availability overlap on Monday, Wednesday, and Friday every week, and we plan on weekly meetings on Wednesday afternoons. Will and Mattheus have prior experience with machine learning, and can help guide how we use our data to build models. In addition, Matthew and Mattheus are comfortable with Python and Jupyter notebooks, and can facilitate the technical aspects of using these tools. Write-ups on our progress can be done by all group members. Our proposed timeline is given below.

|  |  |
| --- | --- |
| Week 3 (2-3) | Select most viable data sources. Begin using standard scaler to standardize all data sets. Create a github repository and begin organization of all assets. |
| Week 4 (2-10) | Deadline: Final project proposal due. Research models and decide which ones are most viable. Establish an understanding of keras, sklearn, matplotlib, and pandas libraries. |
| Week 5 (2-17) | Begin construction of models.  Run through deep learning examples. |
| Week 6 (2-24) | Model training and optimization. |
| Week 7 (3-2) | Model training and optimization. |
| Week 8 (3-16) | Model training and optimization. |
| Week 9 (3-23) | Creation of figures and analysis. Find resources for which to compare the data. Comparison to online published results if applicable. |
| Week 10 (3-30) | Finalize results. Create the first draft of the report. |
| Week 11 (4-6) | Finalize Results. Editing and drafting of final report. Begin presentation creation |
| Week 12 (4-13) | Editing. Finalize presentation. |
| Week 13 (4-20) | Deadline: Papers ready for peer review |
| Week 14 (4-30) | Deadline: Submit final presentation and paper |

**IV. Research**

We will be getting our sources from established authorities on hurricanes, including (but not limited to): NOAA, the National Hurricane Center, and datasets from Risk Management Solutions. We hope to find data about the position of hurricanes’ eyes to track the movement of storms, and potentially see if this correlates with other factors like sea surface temperature.

[NHC Data Archive](https://www.nhc.noaa.gov/data/)

The NHC has comprehensive hurricane trajectory data, dating back to before 1850. We can use this as a way to study broad trends in hurricane paths over close to 200 years.

[NOAA Hurricane Data](https://www.aoml.noaa.gov/hrd/data_sub/hurr.html)

NOAA maintains a large database relating to hurricanes, broken up storm-by-storm. We can use this source not only to dive deeper into certain facets of the storms, but also to present particular storms as case studies of broader trends.

[Risk Management Solutions](https://www.rms.com/perils/hwind/legacy-archive/storms)

This company has released data about hurricane wind speeds from 1995 to 2013. We can use this as a proxy for storm intensity, and see if there are any correlations with hurricane trajectories.

**V. Potential Challenges / Limitations**

We may need help on interpreting our data sets and understanding the best possible way to present them to our model. Converting satellite images or color maps to a set readable by the computer will be important in providing training data for the model. Another area we may need help in is the construction of a model and choosing which one is best for the problem we are trying to solve.

The members of our group have limited outside experience with the tools necessary to achieve our goals (which will only take us so far), and no experience in gathering and parsing earth science data. We hope that the instruction in class will give us the tools we need to overcome these challenges. In addition, the weather systems we are trying to model are controlled by the laws of physics, so any model we develop will likely only re-discover these existing laws. We hope to overcome this by discovering phenomena or correlations that are not predicted by current, human-designed weather models.

**VI. Final Project Summary**

We aim to apply machine learning techniques to predict the future position of hurricanes based on historical data from NOAA, NHC, and other sources.