

Kedari Chowtoori

KCC Technical Assignment

Use the NOAA Best Track Data (HURDAT2) to identify all hurricanes that have made landfall in Florida since 1900.

Using a programming language of your choice, build an application to parse the HURDAT2 data, identify the storms the hurricanes that made landfall in Florida, and output a report listing the name, date of landfall, and maximum wind speed for each event.

Your code should be well commented and organized so other developers can understand how it works. During your interview, be prepared to walk through the code, discuss your design choices, and respond to questions or feedback. After the interview, you will have the opportunity to take the same project further and incorporate feedback from the interview to expand your solution.

You will find the HURDAT2 data set (6.5 MB file) and the HURDAT2 format specification on the NOAA website here: <https://www.nhc.noaa.gov/data/>.

Implementation:

Programming language used: Python

React is used to render the obtained result.

Website: <https://hurricane-tracker.vercel.app/>

Codebase: <https://github.com/ChowtooriKedari/hurricane-tracker>

Python Scripts and Obtained CSVs: <https://github.com/ChowtooriKedari/hurricane-tracker/tree/master/PythonScripts>

I took the HURDAT2 dataset and parsed into a csv for better understanding and accessibility. Once the parsed data is obtained, I used this CSV to find out the landfalls in three different ways.

Approach-1:

Using indicator “L”:

After obtaining all the data, the first thing I did is to check all the land falls in the data. This was done by using the “L” indicator. Once all these were obtained, the next step was to figure out how to get only the ones that hit Florida since 1900. To do this I have used the shape file from [Natural Earth website](#). This helped in separating out the hurricanes that hit Florida.

Approach 2:

Without using “L”:

Since the requirement was to obtain the landfalls, without using the indicator “L”, other available data in the dataset like the latitude, longitude, wind speeds, pressure were utilized. When a hurricane hits the land, its speed usually decreases and the pressure would increase. Comparing it with other existing entries of the same hurricane, previous and next entries of each pattern, we can figure out to the most extent when a landfall would have occurred. Another metric is taking a look at the distance. The distance that the hurricane would have travelled from one location to the next. If it is on land, the distance should be less, since the speed gets reduced. The distance between the two points is checked using the Geopy package and it is checked against the rough estimates like [50,60,70,80,100] miles. If they fall in the range, it is also selected as a criterion for landfall. I was able to achieve almost 70% accuracy with this approach when comparing the obtained results with the earlier approach.

Approach 3:

Machine Learning:

Since we have almost 54749 entries in the dataset, I thought of using ML could help. I trained the model with 80 and 20% split. This is just to verify how accurate the results could be. I tested it with the other splits too, like 70/30 and 60/40. But this approach was based on labeling the data with the indicator, “L”. Since the requirement was not to use it, I tried restricting to approach-2. The CSVs obtained for these are also included in the [GitHub](#).