Weather Metrics and the Impact on Flight Delays

How strongly connected are weather metrics and flight delays (i.e. can we predict flight delays with weather metrics)?

The Data

Weather Metric Data

27 weather metrics, recorded daily for 50 years (18,250 entries) in Boston, MA

National Oceanic and Atmospheric Administration (NOAA)

Flight Delay Data

20 years of monthly departing flight delay data out of Boston Logan International

Bureau of Transportation Statistics (BTS)

Transforming the Data

Step 1

Converted Daily to Monthly Weather Data

Extracted 8 key priority metrics

Step 2

Step 3

Both datasets set to monthly timeline of **June 2003 - August 2023** to standardize dataset timelines

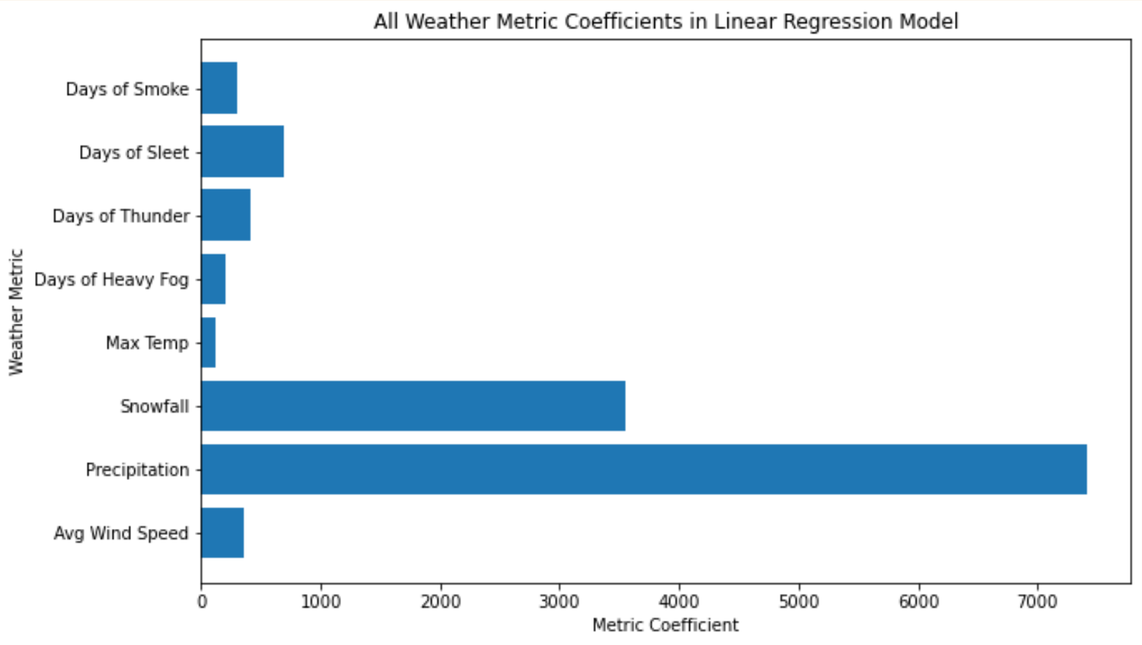
Continuous monthly flight delay data grouped into bins for categorical modeling

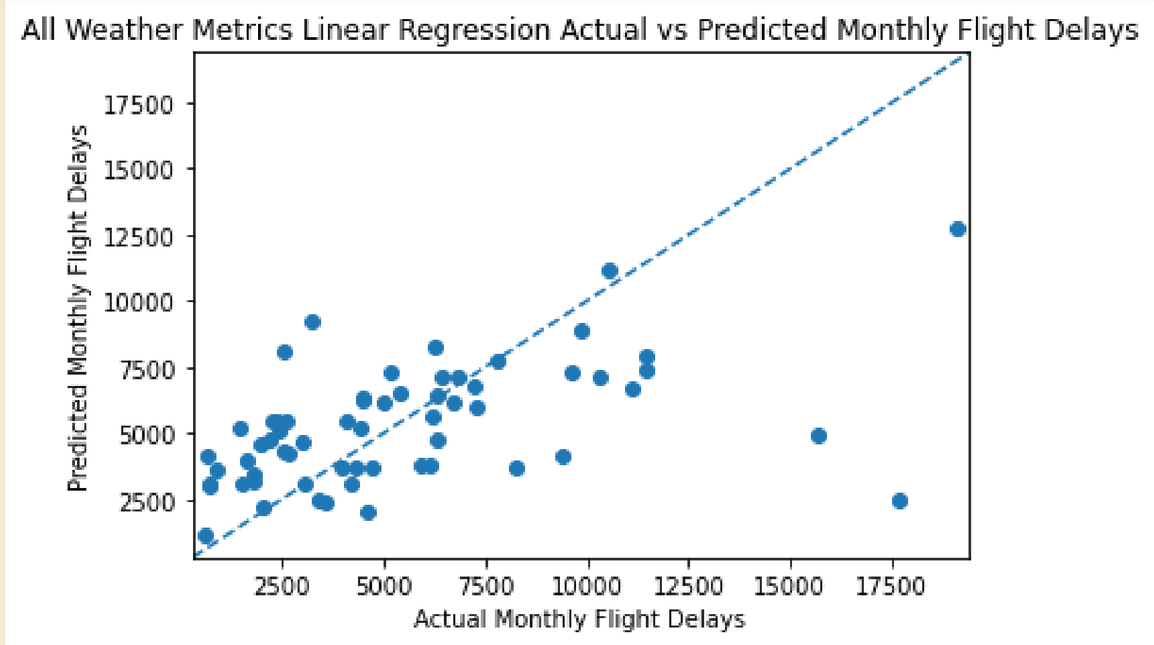
Step 4

Our Process

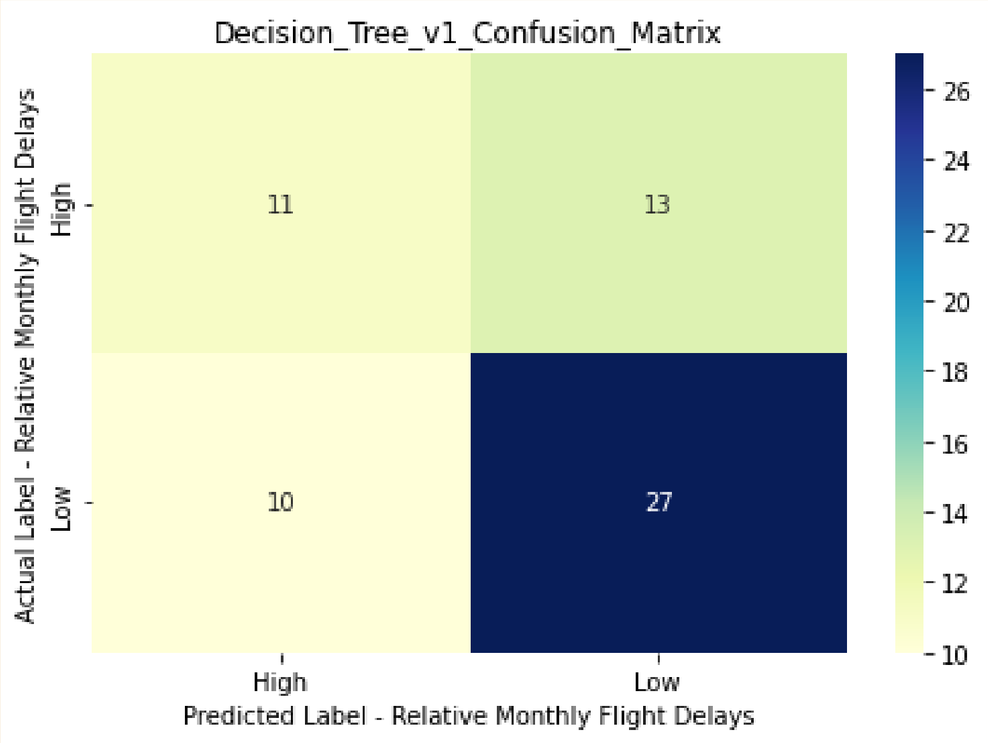
Multivariate Linear Regression

- Mean Squared Error (MSE): 3,448.69
- r-Squared Coefficient: 0.4006
- Continuous monthly flight delay data

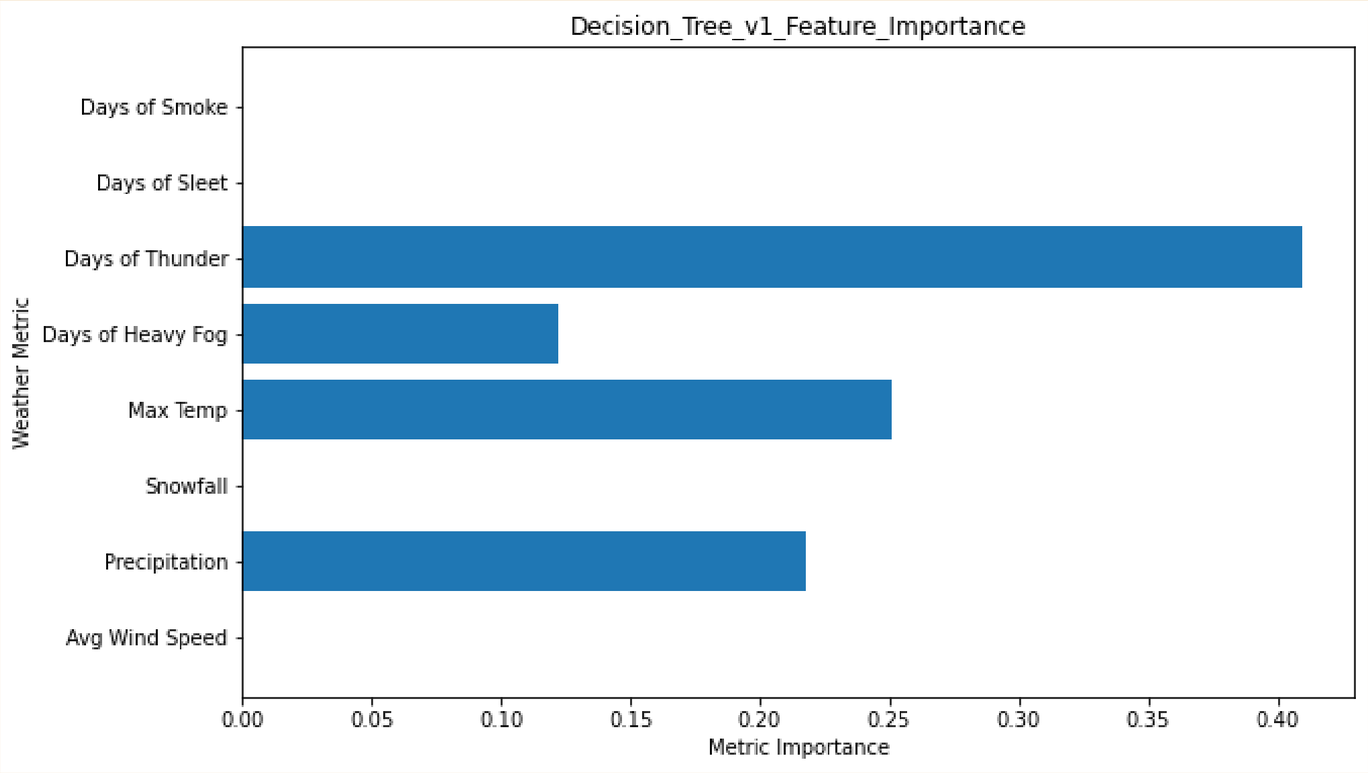




Decision Tree v1

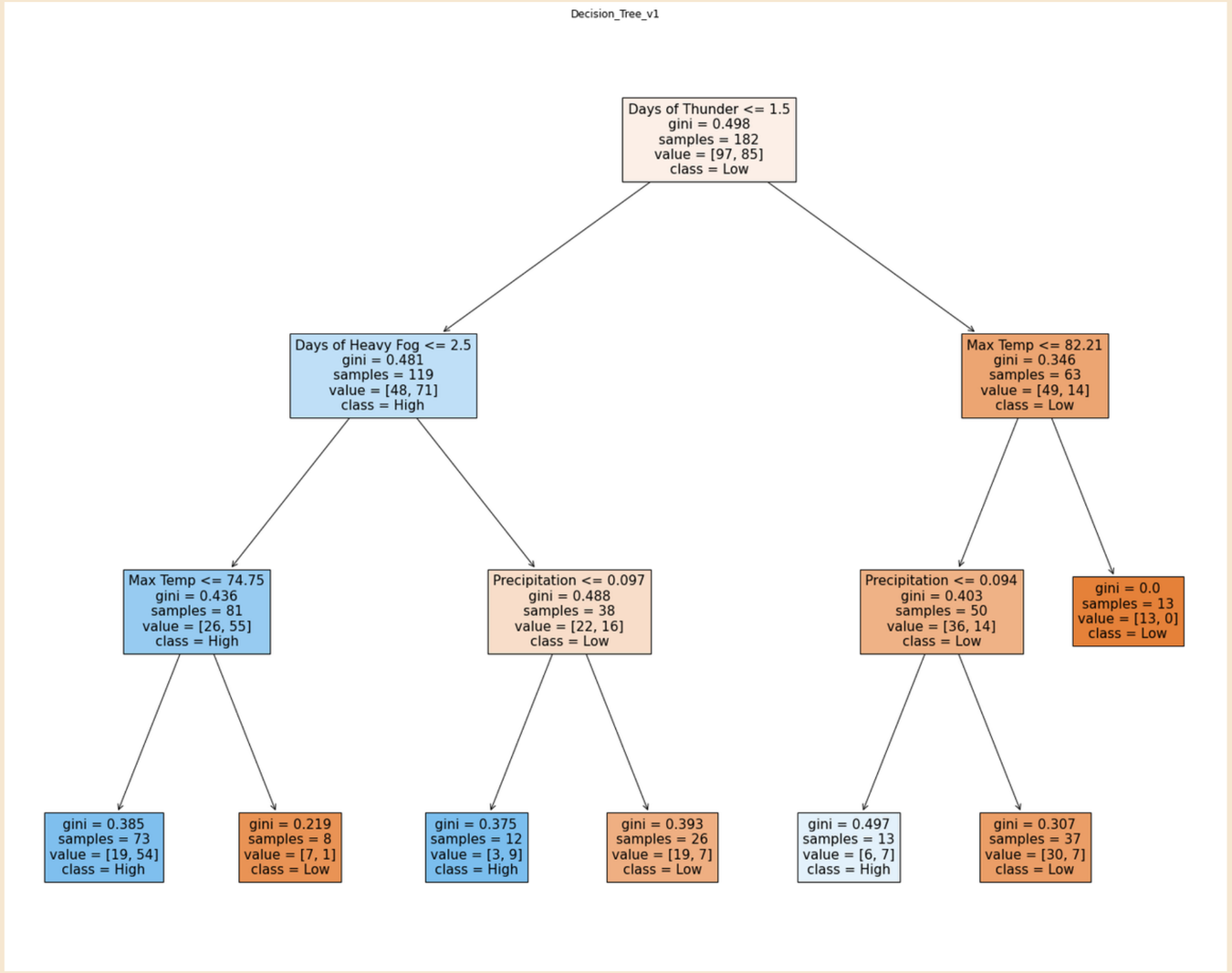


- max_length = 3
- Accuracy of ~62%
- Two label categories: high and low relative monthly flight delays

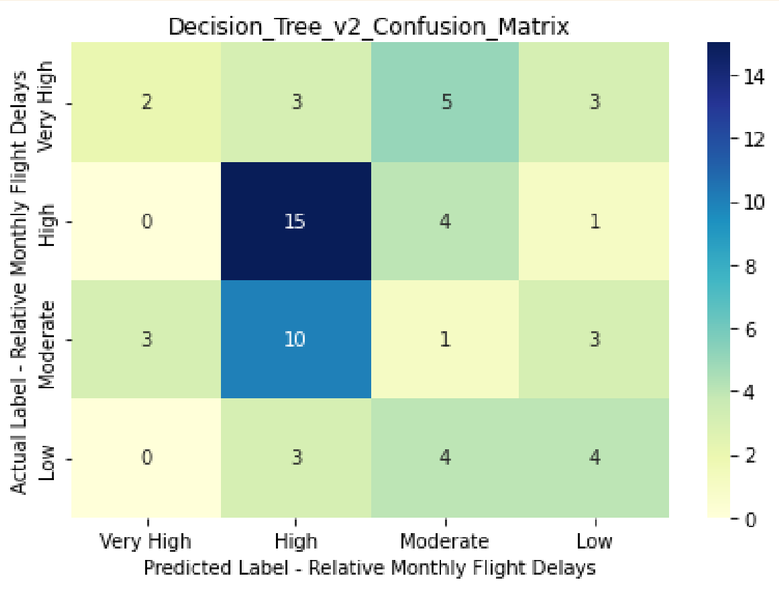


	Precision	Recall	f1 Score
Low	0.52	0.46	0.49
High	0.68	0.73	0.70
Accuracy	-	-	0.62

Decision Tree v1 - 2 Flight Delay Categories

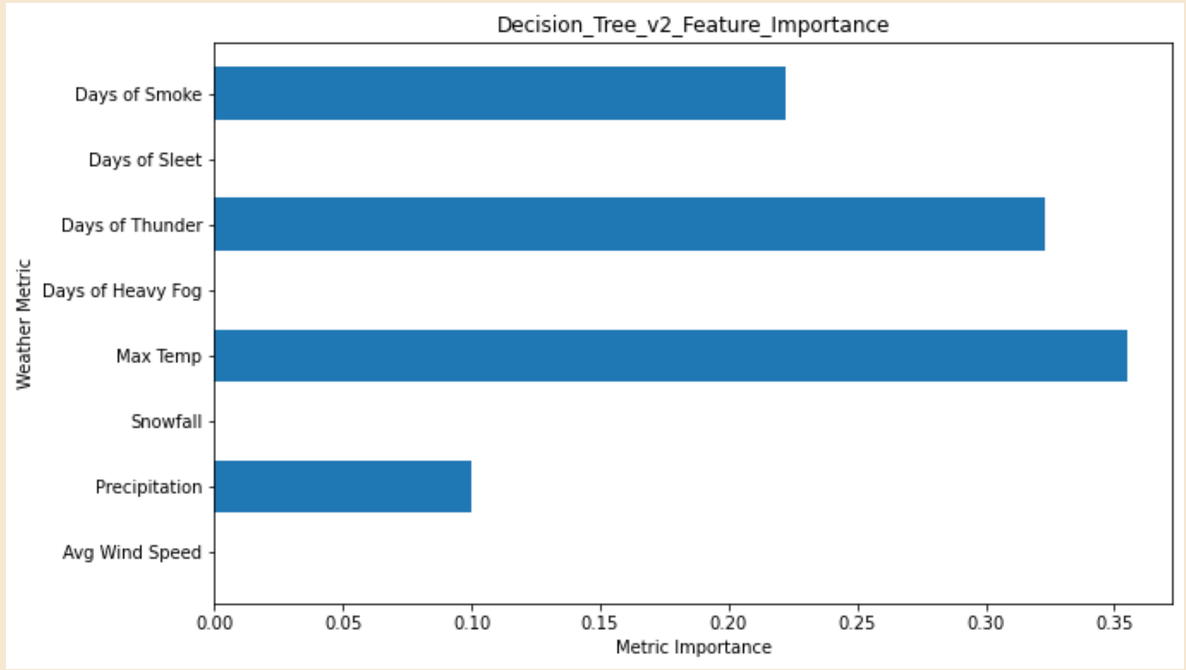


Decision Tree v2

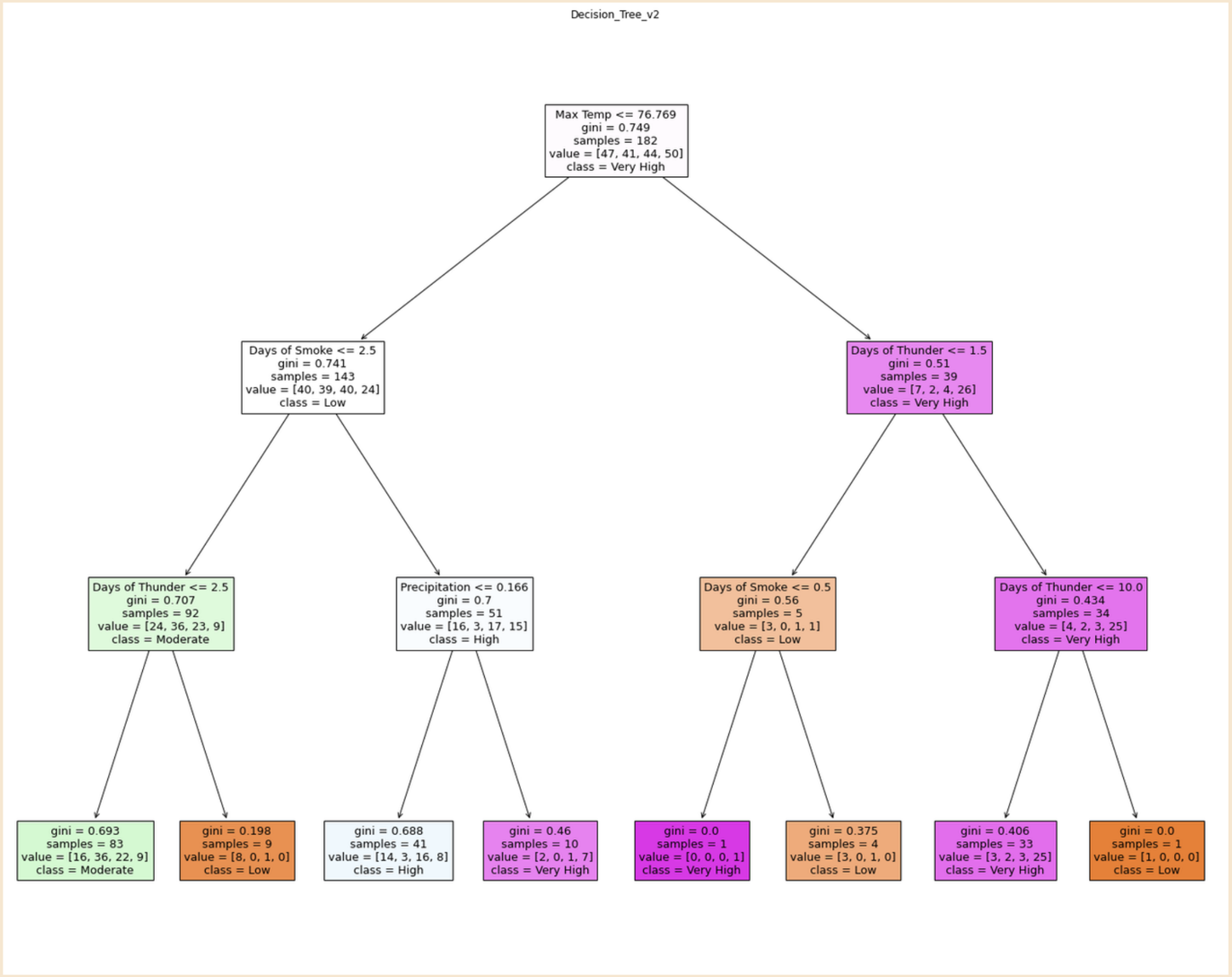


- max_length = 4
- Accuracy of ~36%
- Four label categories: low, moderate, high, and very high relative monthly flight delays

	Precision	Recall	f1 Score
Low	0.40	0.15	0.22
Moderate	0.48	0.75	0.59
High	0.07	0.06	0.06
Very High	0.36	0.36	0.36
Accuracy	-	-	0.36

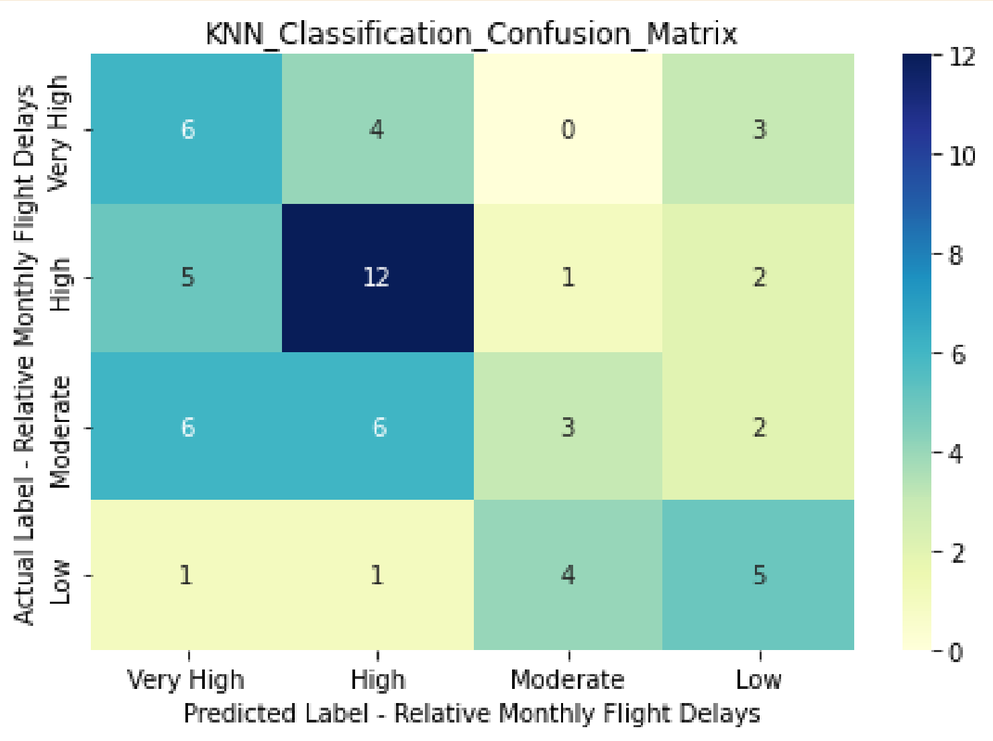


Decision Tree v2 - 4 Flight Delay Categories



KNN Classification

- Uses k = 12
- Accuracy of 43%
- Four label categories: low, moderate, high, and very high relative monthly flight delays



	Precision	Recall	f1 Score
Low	0.33	0.46	0.39
Moderate	0.52	0.60	0.56
High	0.38	0.18	0.24
Very High	0.42	0.45	0.43
Accuracy	-	-	0.43

Results and Takeaways

- Determining metrics for monthly flight delays
  - Days of thunder (storms)
  - Days of heavy fog (visibility)
  - Monthly avg precipitation
  - Monthly avg max temperature (seasonality)

- In general, weather metrics point to flight delay trends
  - i.e. more storms in a month would likely point to more delays that month

- Higher accuracy and specifics in monthly flight delay predictions are unattainable with monthly weather metrics

Future Work

- Obtain daily flight delay data rather than monthly flight delays
- Produce a tool for consumers & airlines to better predict flight delays using historical and forecasted weather
- Include data from weather stations and airports of other cities and hence, improve model accuracy