

# YAP470 Project Proposal

**Project Name:** Multi Label Weather Prediction

**Project Description:** The subject we will be investigating is weather forecasting. Weather forecasts are made by collecting quantitative data about the current state of the atmosphere, land, and ocean and using meteorology to project how the atmosphere will change at a given place. This subject is interesting because human beings have been dealing with weather forecasting since before Christ. Although prediction methods are updated frequently, they do not work perfectly when it comes to predicting the future.

Supervised Learning has an important place in our world when it comes to predicting the future based on past data. This is why we decided to do the weather forecast with Supervised Learning. First, we are going to use classical supervised learning methods. These are nearest neighbor algorithms, decision trees, support vector machines, naive bayes and different linear regression algorithms. We estimate that machine learning algorithms will yield approximately the same results. After that, we are going to try to use the deep neural networks structure. And then, we are going to compare machine learning algorithms among themselves, and compare the performance of deep learning and machine learning.

We are going to analyze codes at Kaggle[3-4] and read academic papers[1-2] about weather prediction. Our main source is Kaggle but we will deeply research the weather prediction domain academically.

Our data consists of historical temperature, precipitation, humidity, and wind speed for Austin, Texas.

There are several implementations in Kaggle, but these studies were done with the data not updated. So, they had less data than we have now. We think we will get a better result with more data and different models. Other studies have an accuracy of around 80%. We think we can increase this to 90% accuracy. We are going to use all evaluation metrics to compare our results.

[1] <http://cs229.stanford.edu/proj2016/report/HolmstromLiuVo-MachineLearningAppliedToWeatherForecasting-report.pdf>

[2] <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7856509>

[3] <https://www.kaggle.com/code/costantinomarco/multiclass-svc-and-ml-pipeline>

[4] <https://www.kaggle.com/code/costantinomarco/meteorological-data-classification-decision-tree>