Victor Pontis

6.UAP

Supervisor: Prof. Dennis Freeman

**Predicting Hourly Rainfall Given Radar Measurements**

**Introduction**

It is extremely important for a farmer to have an accurate idea of the amount of rainfall on their farm. The average farm size is now around 441 acres making it extremely difficult for a farmer to have an understanding on how different parts of the farm are doing without the help of data that is automatically collected.[[1]](#footnote-1)

Farmers need to understand the rain distribution over their farm in order to maximize their yield. Certain crops require precise amounts of rainfall and the rainfall over an average farm can deviate meaningfully within the space of a few acres. With accurate rainfall measurements, farmers can ensure that a certain crop gets the optimal amount of rainfall, which increases their production and profits.

The traditional way to measure rain is with a rain gauge. A rain gauge is simply just a cylinder that catches rain and the height can be measured at regular intervals to determine rainfall over time. These simple rain gauges are especially sensitive to wind, cold weather and splashing, and do not provide any more information other than the cumulative amount of rain fall between observations.

Farmers have been moving towards polarimetric radar measurements to get more information about rainfall. Polarimetric radar measurements provide information on the size of rain droplets and provide resolution on the order of seconds rather than the hour resolution of rain gauges. In this project, I will be using polarimetric radar data to predict the hourly rainfall on different regions. If successful this project will help farmers understand rainfall on their farms and increase crop yields.

**Technical Implications**

This project relies on machine learning and data mining. I found this problem with associated data sets on a Kaggle.com.[[2]](#footnote-2) Kaggle is a site that hosts data mining competitions and provides the contestants with clean data and test cases. I will be utilizing their data and test cases as well as communicating with other participants on the online forum to improve my results.

The input data consists of radar measurements at different intervals. I will output the probabilistic distribution of rainfall at hourly intervals. This output will be tested against rain gauge data.

To predict the rainfall I will be using fundamental techniques from machine learning and data mining. I will explore the usage of support vector machines, random forests, decision trees, neural nets, etc. I will also be using cross-validation to make sure that I am not contaminating my model with input from the test set.

In addition to just providing a framework for probabilistically predicting rainfall, I will analyze and report on the relative success of different metrics. I will use this analysis with academic literature on polarimetric radar data to explain the difference in results among the metrics I use.

**Outcomes**

At the end of this term I will have produced the following:

* A comparative analysis of the relative strengths of polarimetric radar and rain gauge measurements
* A framework for probabilistically predicting rainfall given polarimetric radar measurements
* Suggestions for future work and improvements, including ideas for resolving problems present in my solution

**Conclusion**

Understanding rainfall across a farm is crucial to maximizing crop yield. Polarimetric radar data provides strong potential to improving upon the information provided by rain gauges but are not yet fully understood. This project will create a framework for interpreting polarimetric radar data and predicting hourly rainfall in different regions of a farm.

1. I got the average size of farm from http://www.agday.org/media/factsheet.php. [↑](#footnote-ref-1)
2. https://www.kaggle.com/c/how-much-did-it-rain [↑](#footnote-ref-2)