Tornado Classification

Jessica Warner  
 Master’s in Data Science University of Colorado-Boulder

ABSTRACT/INTRODUCTION

Tornadoes are among the most destructive natural phenomena, causing significant damage to property and posing severe risks to human life. This project proposes an approach to tornado prediction using machine learning techniques, specifically focusing on a classification model. The primary focus is to develop a model capable of accurately predicting tornado occurrences based on a comprehensive set of weather data.

The project will involve collecting and analyzing historical weather data from multiple regions across the United States, including a wide range of variables including temperature, humidity, pressure, wind speed, and precipitation, among others. Sophisticated feature engineering techniques and sampling will be used to capture subtle patterns in weather conditions that may be indicative of tornado formation.

The core of our approach is a machine learning classification model, utilizing algorithms such as Random Forests and cross fold validation. This model will be trained on a carefully curated dataset, balancing the rare occurrence of tornadoes with the more common non-tornado weather conditions. Sampling techniques and possible use of SMOTE will be employed to handle class imbalances.

Related Work

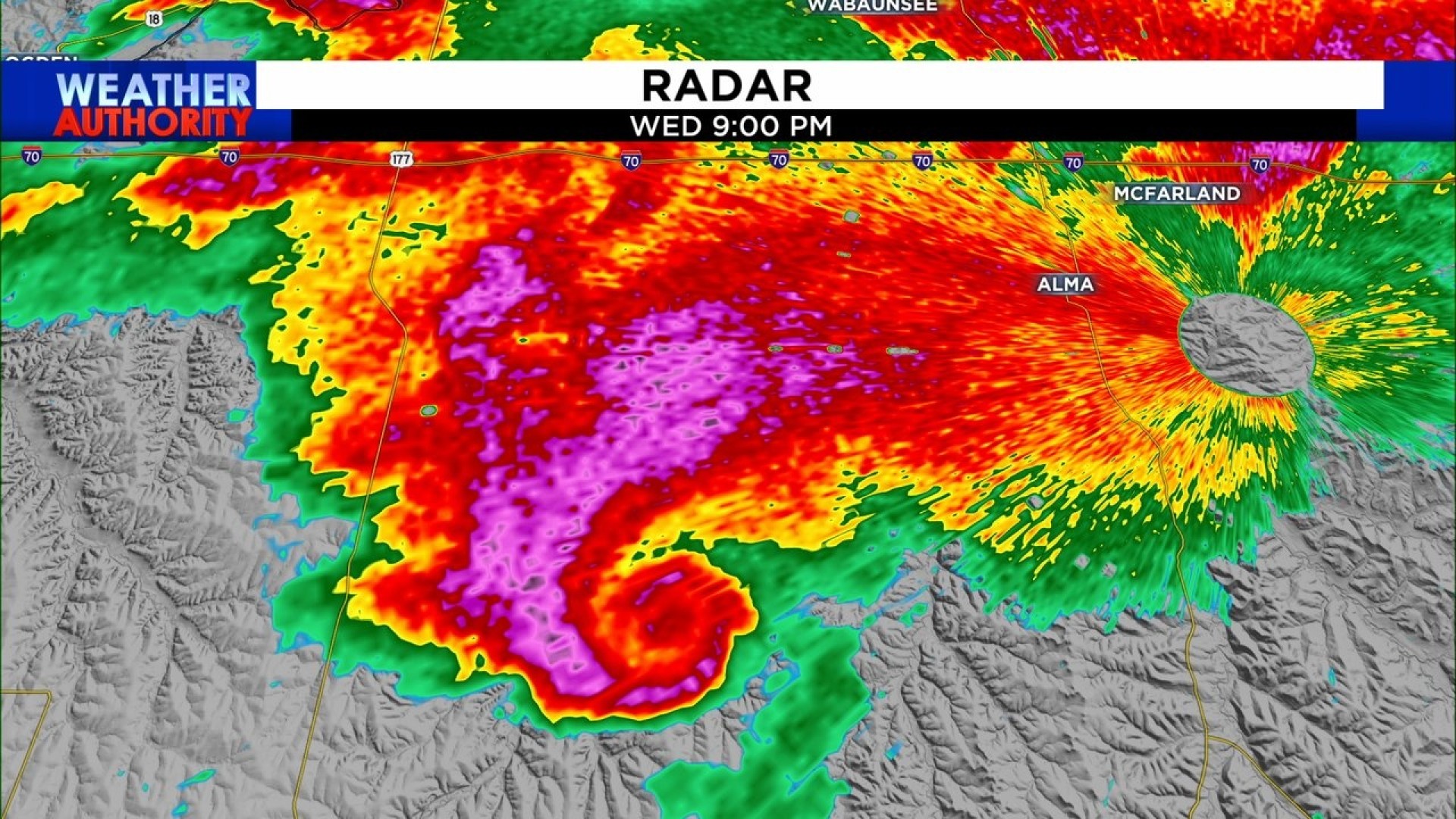
Other work has been done related to the topic of predicting tornado occurrences or classifying severe weather conditions that could create tornados. One of esteemed note, published on the American Meteorological Society, is titled “Machine Learning Classification of Significant Tornadoes and Hail in the United States Using ERA5 Proximity Soundings”. This journal also assesses a method to identify and predict an often-unpredictable weather phenomenon by analysis of meteorological data such as wind and hail.

Proposed Work

1. The incorporation of region-specific weather and current distribution of occurrences to account for geographical and seasonal variations in tornado formation conditions.

2. Feature selection and engineering to know what the most important variables in tornado formation in order are to help with dimensionality reduction because of excessive number of columns.

3. Potential integration of time-series analysis to capture the evolution of weather patterns leading up to tornado events.



Evaluation

The project aims to achieve high recall without significantly sacrificing precision, recognizing the critical importance of minimizing missed tornado predictions. Rigorous evaluation of the model using cross-validation techniques focusing on metrics such as ROC-AUC, Precision-Recall curves, and F1 score will be included.

Conclusion

The successful completion of this project has the potential to significantly enhance current tornado warning systems, providing more accurate and timely predictions. This could lead to improved public safety measures and more effective emergency response planning, ultimately reducing the impact of these devastating natural disasters.

Conference Name:ACM Woodstock conference

Conference Short Name:WOODSTOCK’18

Conference Location:El Paso, Texas USA

ISBN:978-1-4503-0000-0/18/06

Year:2018

Date:June

Copyright Year:2018

Copyright Statement:rightsretained

DOI:10.1145/1234567890

RRH: F. Surname et al.

Price:$15.00