Keene Particulate Matter Project - Roadmap Document

Index

1. [Data Science Lifecycle Review](#_Data_Science_Lifecycle)
2. [Problem Definition and Domain Knowledge](#_Problem_Definition_and)
   1. Problem Definition
   2. Domain Knowledge
      1. Literature Review
3. [Data Collection and Sourcing](#_Data_Collection_and)
4. [Data Cleaning and Processing](#_Data_Cleaning_and)
5. [Exploratory Data Analysis](#_Exploratory_Data_Analysis)
6. [Model Building and Evaluation](#_Model_Building_and)
7. [Model Results](#_Model_Results)
8. [Model Deployment](#_Model_Deployment)

## Data Science Lifecycle Research/Review

A diagram of a process

Description automatically generated

## Problem Definition and Domain Knowledge

### Problem Definition

1. To forecast, with the highest possible accuracy, when PM 2.5 levels will be elevated in the Connecticut River Valley, specifically over the City of Keene, given meteorological data gathered by KSC’s Nora Traviss.
2. To evaluate each model to optimize for a variety of factors: error, compute, number of features, and more to be determined.

## Data Collection and Sourcing

## Data Cleaning and Processing

## Exploratory Data Analysis

## Model Building and Evaluation

## Model Results

## Model Deployment