# Trail Maintenance Needs Prediction

Luke Rand and Isaac Olson

## Overview:

Using machine learning algorithms, we will generate a model that predicts trail maintenance needs.

## The Problem:

Predicting trail maintenance is a complicated problem involving many variables, and thus could prove difficult for humans to accurately predict. Common problems include trails widening due to overuse, developing trenches due to drainage issues, moving water causing erosion, and overgrowth of foliage (Trail Maintenance). Thus, trail maintenance needs depend on many factors, such as user traffic, rainfall, slope, soil composition, and surrounding environment. While large organizations like the United States Forest Service employ specific “standardized concepts and tools” (Trail Management), smaller organizations, like those in New Mexico, struggle to adequately maintain trails, a fact we have seen in our own personal experience on Santa Fe’s trail systems. At the same time, access to good trails has innumerable benefits to the population including physical exercise, mental health improvements, better relationships with others, and improved emotional, intellectual, and physical health in children (Benefits of Hiking, Why Trails Matter).

# Our Solution:

Given the complicated, interconnected, and patterned nature of trail maintenance needs, dependent on multiple factors which hold relation yet each influence the outcome, supervised machine learning models lend themselves to the problem. Supervised machine learning refers to providing an algorithm with multiple inputs and a desired output, from which it constructs a model to create these outputs from unfamiliar inputs (What Is Machine Learning). By feeding the model data that includes both input variables (slope, rainfall, soil, traffic) and desired outputs from a large trail maintenance organization (in the form of categorization or quantification) l, we can predict trail maintenance needs for areas serviced by smaller organizations, who may not have the resources to figure out how to best allocate time.

# Our Progress:

So far, we have had our “meet the scientist” meeting, which gave us important information on the “engineering” nature of our project, and how to approach it best. We recognized that developing a viable model early in the process and undergoing extensive testing will result in a stronger scientific leaning to our project. Furthermore, we have made contact with several trail maintenance organizations, the most notable being the Pacific Crest Trail Association, which has supplied us with work logs going years back, which we will scrape to generate the maintenance need variable of our training dataset. Additionally, we are currently planning a meeting with both the Pacific Crest Trail Association and United States Forest Service to better understand the current landscape of trail maintenance and determine whether we are missing anything, as well

as gain a better grasp of how to best proceed. Lastly, we have identified the following datasets to gather additional input variables to train the model:

- Traffic:

We are going to get data about trail traffic from Strava, a social media app for athletes to share their workouts, to compare the number of people that are using different trails. We will use the Strava API to get this data for different trail sections.

https://www.strava.com/maps/global-heatmap?sport=All&style=dark&terrain=false&labels=true&poi=true&cPhotos=true&gColor=blue&gOpacity=100#7.08/35.988/-118.154

https://developers.strava.com/docs/reference/#api-SegmentEfforts-getSegmentEffortById

- Slope:

There are multiple different data sets that we can use to find the slope of the trail. We are considering using the data from Strava to get the grade of the trail but it does have some limitations because of how it averages trail slope over longer sections.

https://developers.strava.com/docs/reference/#api-SegmentEfforts-getSegmentEffortById

- Rainfall:

The USDA Geospatial Data Gateway will allow us to download average rainfall data in different regions in the same data format as the pacific crest trail work logs witch should make the process much easier.

https://datagateway.nrcs.usda.gov/

- Soil:

We will get data about soil composition from the Soil Survey Geographic Database made by the USDA conservation of natural resources service. This data also exists in the same format as most of our other data.

https://coast.noaa.gov/digitalcoast/data/ssurgo.html

Works Cited

"Benefits of Hiking." *National Park Service*, www.nps.gov/subjects/trails/benefits-of-hiking.htm. Accessed 19 Dec. 2024.

"Trail Maintenance for Safer, Healthier Trails." *Productive Parks*, productiveparks.com/trail-maintenance-safer-healthier-trails#:~:text=Trails%20need%20to%20be%20maintained,resources%20in%20the%20long%20run. Accessed 19 Dec. 2024.

"Trail Management Tools." *Forest Service*, United States Department of Agriculture, www.fs.usda.gov/managing-land/trails/trail-management-tools. Accessed 19 Dec. 2024.

"What Is Machine Learning?" *MathWorks*, www.mathworks.com/discovery/machine-learning.html. Accessed 19 Dec. 2024.

"Why Trails Matter: Outdoor Learning." *American Trails*, www.americantrails.org/resources/why-trails-matter-outdoor-learning. Accessed 19 Dec. 2024.