



---

# Hike Finder

Through the National Park Service

---

Mary Velasquez

June 3, 2025

# Overview

**Goal:** Build an app that recommends National Park hiking trails based on your mood

## Data

National Park Service API for trail names, trail descriptions, duration, park location, etc.

## Model

Rules-based mood model based on duration + trained classifier to further refine using descriptive trail text

## App

Users select mood and park to get trail recommendations

# Data Collection & Cleaning

- **Collected 1,500+ activities from the NPS "Things to Do" API**  
Queried paginated results and filtered for those tagged with "Hiking"
- **Flattened and extracted key attributes for each trail**  
Parsed park name, duration, season, state, and description
- **Standardized trail durations into minutes**  
Converted varied duration inputs into numeric durations
- **Created "Energy Level" categories based on trail duration**  
Classified trails into Relaxed, Chill, Adventurous, and Challenging

```
1-15 Hours      1
0-90 Minutes    1
30-300 Minutes  1
30-75 Minutes   1
1-3 Days        1
5-20 Minutes    1
40-120 Minutes  1
30-50 Minutes   1
40 Minutes      1
5 Minutes       1
9 Hours         1
7 Hours         1
0-30 Minutes    1
0-3 Hours       1
30-80 Minutes   1
2-10 Hours      1
10-180 Minutes  1
40-240 Minutes  1
6-9 Hours       1
10-240 Minutes  1
45-180 Minutes  1
12 Hours        1
4-12 Hours      1
Name: count, dtype: int64
```

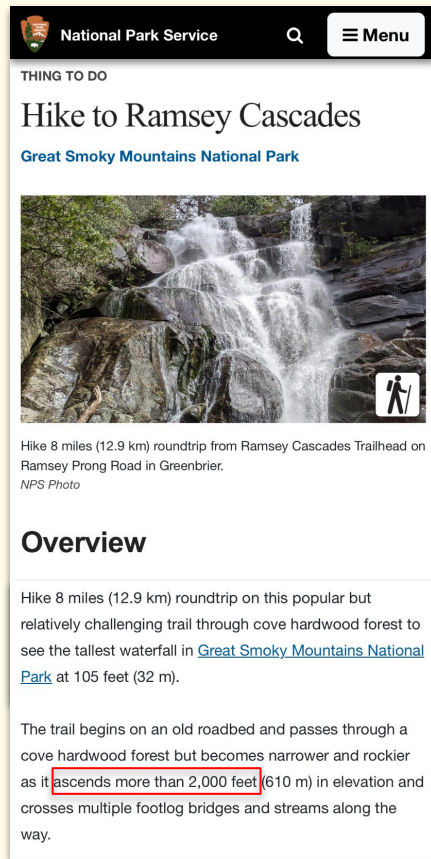
*Snippet of the wide variation in trail durations*



# The Model

**Goal:** Build on prior rule-based tagging by learning to predict energy level from trail descriptions and elevation cues.

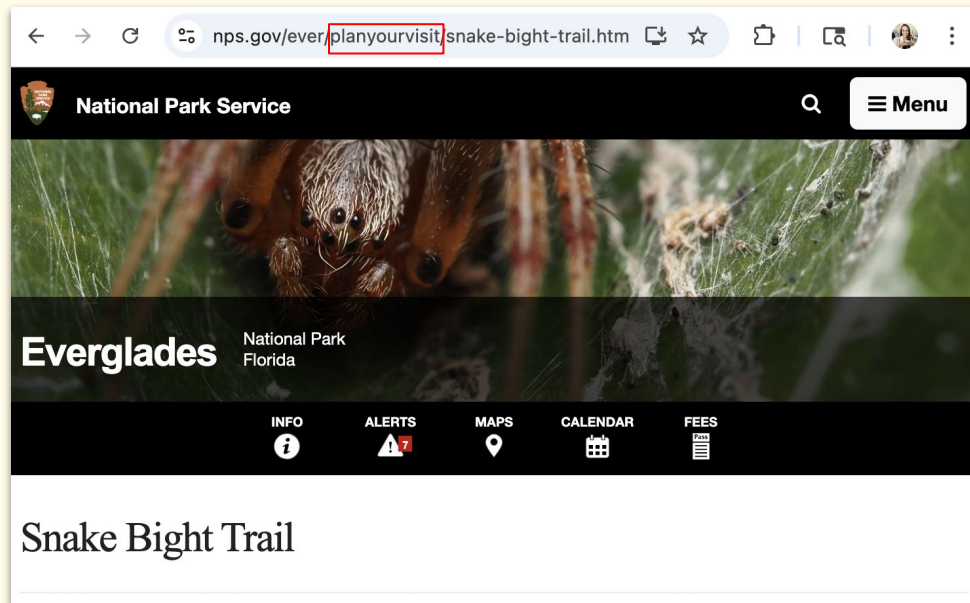
1. **Input:** Extracts trail longDescription via the NPS API
2. **Processing:**
  - Removes HTML tags
  - Detects elevation gain using regex
  - Transforms the text into a TF-IDF vector + binary elevation flag
  - Feeds both into a logistic regression model
3. **Output:** Returns predicted energy level for the description



*Regex formula identifies elevation gain in description*

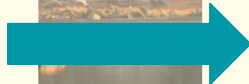
# Limitations & Enhancements

- **Limited API data**  
Only 37 of the 63 national parks appeared in the API due to inconsistent folder structure on the website
- **Regex formula refinement**  
Formula could be further refined to better capture variations in phrasing and units



*Everglades National Park hiking trails live under “Plan Your Visit” rather than “Things to Do”*

SCAN ME



yesitsmary.shinyapps.io



## Find a Hiking Trail

Select your energy level:

Challenging



Select National Park:

Grand Teton



### Valley Trail - Phelps Lake Overlook

Travel from Teton Village to the Phelps Lake Overlook on this strenuous hike.

Duration: 7-9 Hours

[More details](#)





A scenic landscape featuring a calm lake reflecting the surrounding forest and distant mountains. In the foreground, a wooden bench sits on a grassy bank. The scene is framed by large green shapes on the left and top right, which contain the text 'You' and 'Thank' respectively.

# Thank

# You