Ski Genius Final Presentation

Ashwin, Chakita, Kyle (1), and Kyle (2)

Background

Our project is a recommendation platform where a user inputs information to a web interface, and the tool recommends the best ski resorts for that user.

Ski resort characteristics considered:

- Location
- Difficulty level
- Cost
- Snow quality

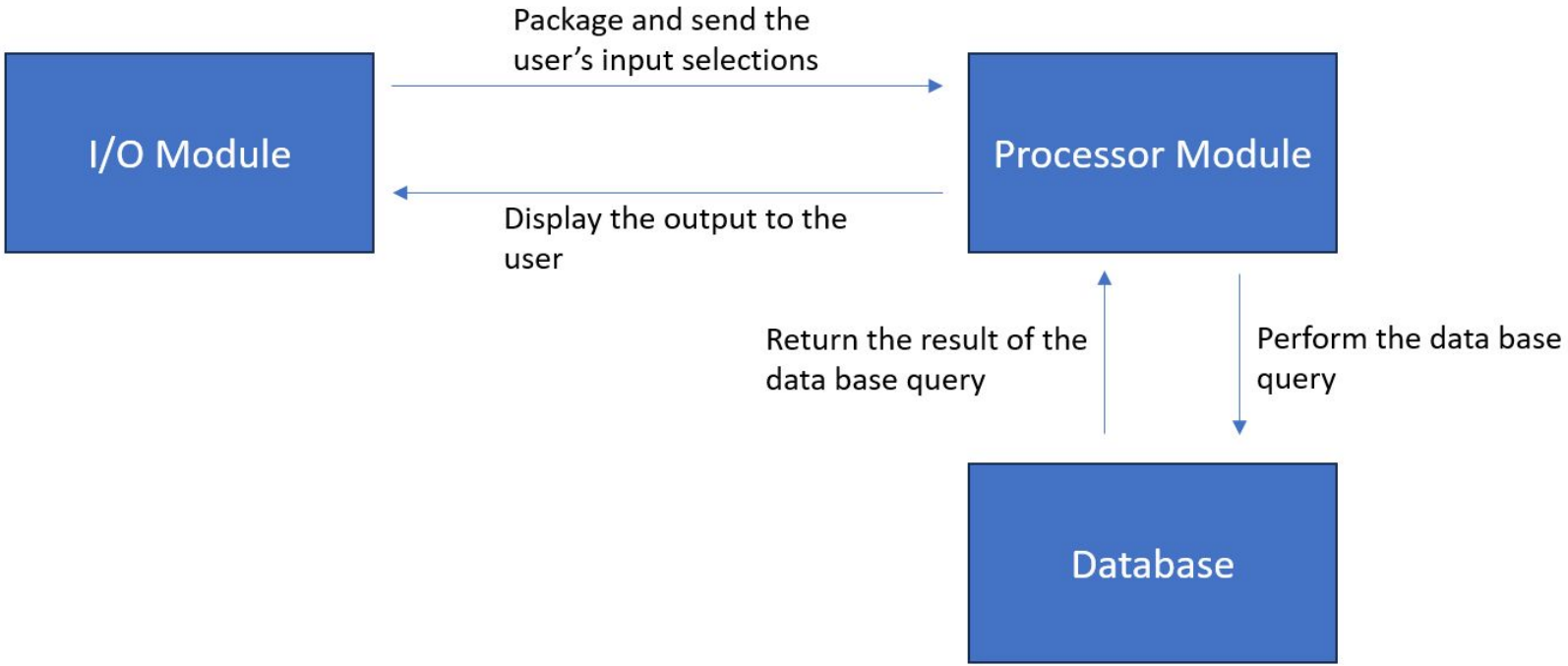
Data

- Global dataset of ski resort names, locations, and other characteristics (https://www.kaggle.com/datasets/ulrikthygepedersen/ski-resorts?select=resorts.csv)
- Nationwide dataset of ski resort coordinates from NSAA (https://services2.arcgis.com/C8EMgrsFcRFL6LrL/arcgis/rest/services/Ski_Resorts/FeatureServer/0)
- National snowfall dataset from NOHRSC (https://www.nohrsc.noaa.gov/snowfall_v2/)

Use Case

- User: Access webpage for tool
- System: Display options for user to enter ski area preferences
- User: Selects desired options
- System: Shows a list and map of the top ski areas sorted by how well they match the user’s preferences (includes resort name, snow summary, estimated trip price)

Design


Package and send the user's input selections
I/O Module
Processor Module
Display the output to the user
Return the result of the data base query
Perform the data base query
Database

Design: I/O Module

-Options for user to enter in their location, desired difficulty level, and financial constraints

-After the processor module queries database, displays output list of recommended ski resorts in user-friendly format

-Implemented in Streamlit

Design: Processor Module

-Turns user inputs into a vector

-Queries database using pymongo MQL language

-Performs k-nearest-neighbor search algorithm to find ski resorts in database which most closely resemble user input vector

-Returns top ski resorts in a format that is usable for the I/O module

Design: Database

- Stores data on ski resorts
- Implemented in MongoDB
- Vectorize the rows so we can perform a distance calculation

Lessons Learned

- Be careful about data sources
  - Coordinate problems with global ski resort dataset
- Iterate on query search methods
  - Location was underweighted at first, leading to strange results
- Good testing fundamentals
  - Mocking
  - Writing test cases in parallel
- CI/CD experience

Demo

https://skigenius.streamlit.app/


Questions?