March 1, 2025

C964: Computer Science Capstone Template

[Part A: Letter of Transmittal 2](#_Toc193182869)

[Part B: Project Proposal Plan 4](#_Toc193182870)

[Project Summary 4](#_Toc193182871)

[Data Summary 4](#_Toc193182872)

[Implementation 4](#_Toc193182873)

[Timeline 5](#_Toc193182874)

[Evaluation Plan 5](#_Toc193182875)

[Resources and Costs 6](#_Toc193182876)

[Programming Environment: 6](#_Toc193182877)

[Human Resources: 6](#_Toc193182878)

[Part C: Application 7](#_Toc193182879)

[Part D: Post-implementation Report 8](#_Toc193182880)

[Solution Summary 8](#_Toc193182881)

[Data Summary 9](#_Toc193182882)

[Machine Learning 9](#_Toc193182883)

[Validation 10](#_Toc193182884)

[Visualizations 10](#_Toc193182885)

[Application Testing 11](#_Toc193182885)

[User Guide 11](#_Toc193182886)

# Part A: Letter of Transmittal

03.01.2025

John Smith

Climbing Co.

459 Climbing Rd, Salt Lake City, Utah

Dear John Smith

I am writing to propose the development of a rock-climbing route recommendation application aimed at assisting new climbers in discovering suitable climbing routes based on their skill levels and preferences. One of the key barriers to entry for new climbers is the overwhelming variety of outdoor climbing routes, which can be daunting for beginners. Many climbers struggle to find suitable routes that match their skill level, leading to frustration, safety concerns, and a lack of motivation. As a result, there is a clear need for a tool that helps new climbers navigate these challenges by recommending climbing routes based on their skill levels and style preferences.

This data product will benefit users by providing personalized climbing route recommendations that match their skill levels and preferred climbing styles. The application will simplify the decision-making process for new climbers by offering easy-to-understand suggestions, thus improving their overall climbing experience. This product will tailor route suggestions based on user preferences, visualize route distributions by difficulty and style, and provide a curated collection of climbing routes with details on difficulty and style. The data used will include the route name, difficulty, and style. The difficulty will consist of ‘Beginner’, ‘Intermediate’, and ‘Advanced’. The style will include ‘Sport’, ‘Trad’, and ‘Bouldering’. This data will be stored in an organized database and updated regularly as more routes are added. This project aims to develop an application that provides personalized route recommendations for new climbers. We hypothesize that the application will enhance the experience of beginners interested in climbing outdoors. The project's methodology will be first to gather and process the necessary data, a recommendation engine will be developed alongside a dashboard for a user-friendly experience. After testing and refinement, the application will be ready to publicize.

The funding required to build and maintain the project consists of an upfront cost of $25,000 and an additional $5,000 per year to update and maintain the software. The impact of this project on stakeholders will be beneficial, this application will foster greater participation and community growth. This will lead to an increased interest in climbing and may lead to growth for local businesses, including gyms and gear suppliers. This application will keep all user data anonymized and securely stored, the recommendation system will be unbiased, ensuring equal treatment.

With my experience in **data science**, **machine learning**, and **software development**, I am confident that I can deliver a high-quality solution that meets the needs of new climbers and enhances their outdoor climbing experience.

Thank you for taking the time to read my proposal. I’m looking forward to hearing your response. I look forward to the opportunity to discuss this project further.

Sincerely,

**Artyom Basov**

Artyom Basov, Lead Developer

# Part B: Project Proposal Plan

## Project Summary

The problem this product addresses is the difficulty new climbers face when attempting to transition into outdoor climbing. Beginners often find it challenging to choose routes that match their skill levels and preferred climbing styles, leading to frustration and potential safety risks. The proposed application will provide personalized climbing route recommendations based on user input, making the decision-making process easier, safer, and more enjoyable.

The target clients are novice to intermediate climbers who are looking for outdoor climbing routes tailored to their skill levels and climbing styles (e.g., Sport, Bouldering, Trad). These customers often lack experience in selecting appropriate climbing routes, and our product will fulfill their needs by offering data-driven recommendations that enhance their outdoor climbing experience. This product will also appeal to climbing gyms and outdoor enthusiasts who want to encourage more people to engage in climbing outdoors.

The primary deliverables associated with the project are:

1. A fully functional recommendation engine that provides personalized route suggestions.
2. A user-friendly dashboard for input and visualization.
3. A database of climbing routes with regularly updated information.
4. Documentation for maintenance and future development.
5. A comprehensive testing report, including user feedback and improvements made.

## Data Summary

To support this product, we will collect and manage data on outdoor climbing routes, which include:

* Route name
* Difficulty (Beginner, Intermediate, Advanced)
* Climbing style (Sport, Bouldering, Trad)

This data will be cleaned, encoded, and stored in a structured format for use in our recommendation algorithm. We will source the data from reputable climbing databases and manually curated sources. Regular updates to the data will be necessary to ensure accuracy and relevance.

## Implementation

The following methodology will guide the development of the data product:

1. Data Collection: Gather and preprocess climbing route data.
2. Recommendation Engine Development: Use machine learning techniques (e.g., decision trees) to create an algorithm that predicts the best routes based on user input.
3. User Interface Design: Build an intuitive, interactive dashboard that allows users to input preferences and view recommendations.
4. Testing and Feedback: Continuously collect user feedback to refine recommendations and improve the user experience.
5. Deployment: Launch the product on a scalable platform, ensuring smooth functionality and accessibility.

**The plan for implementation includes several key phases:**

Phase 1: Planning & Data Collection – Collect and preprocess route data and establish database structure.

Phase 2: Development & Prototyping – Build the recommendation engine and UI; prototype features.

Phase 3: User Testing & Refinement – Conduct usability testing and refine the product based on user feedback.

Phase 4: Deployment & Maintenance – Deploy the product, provide regular updates, and monitor its performance.

## Timeline

|  |  |  |  |
| --- | --- | --- | --- |
| **Milestone or deliverable** | **Duration** | **Projected start date** | **Anticipated end date** |
| Planning and Data Collection | 15 hours | March 1 | March 6 |
| Development and Prototyping | 40 hours | March 7 | March 14 |
| User Testing and Refinement | 20 hours | March 15 | March 19 |
| Deployment and Maintenance | 12 hours | March 20 | March 22 |

## Evaluation Plan

We will validate and verify the data product through:

1. **User Testing**: Conducting trials with real climbers to assess the effectiveness of the recommendations.
2. **Accuracy Evaluation**: Regularly testing the recommendation engine against real-world user feedback to ensure its predictions align with climbers’ expectations.
3. **Performance Metrics**: Analyzing user engagement and satisfaction metrics to assess whether the product improves the decision-making process for users.

## Resources and Costs

## Programming Environment:

* **Backend**: Python (with libraries such as pandas, sci-kit-learn, and SQLite) for the recommendation engine and database management.
* **Frontend**: Streamlit for the UI, with visualization libraries like Plotly for interactive charts.
* **Cloud Infrastructure**: Amazon Web Services (AWS) or similar for hosting and database management.

Much of the software needed to complete this project will be free and open source. An initial $2,000 will be needed for setting up and configuring the infrastructure using AWS. An ongoing $2,000 - $3,000 will be required annually to maintain and update the server.

## Human Resources:

The bulk of the costs will come from paying employee salaries, this project will require one developer to complete. Roughly 100 hours will be spent completing the project. At a salary of $100/hour, roughly $10,000 will be paid to the developer.

# Part C: Application

See attached folder

# Part D: Post-implementation Report

Create a post-implementation plan as outlined below. Provide sufficient detail so that a reader knowledgeable in computer science but unfamiliar with your project can understand what you have accomplished. Using examples and visualizations (including screenshots) beyond the three required is recommended (but not required). **Write everything in the past tense.**

## Solution Summary

New climbers often face significant challenges when selecting suitable outdoor climbing routes. Many beginners struggle to find routes that match their skill level and preferred climbing style. This lack of guidance can lead to frustration, poor experience, and even safety risks if climbers attempt routes that are too difficult or are not suited to their abilities. Additionally, discovering appropriate climbing routes can be overwhelming due to the sheer volume of options available in outdoor climbing areas.

The proposed Rock-Climbing Route Recommendation Application addresses this problem by providing personalized route suggestions tailored to the user’s skill level and preferred climbing style. By inputting their preferences, users receive recommendations for routes that match those preferences. This solution simplifies the decision-making process for beginners, reduces the risk of choosing inappropriate routes, and ultimately enhances the outdoor climbing experience.

The application solves the problem of route selection by offering several key features that streamline the decision-making process for new climbers:

1. **Personalized Recommendations**:  
   Users input their climbing skill level (e.g., Beginner, Intermediate, Advanced) and preferred climbing styles (e.g., Sport, Bouldering, Trad). Based on this input, the application’s recommendation engine uses a machine learning algorithm to suggest suitable outdoor climbing routes tailored to the user’s preferences and abilities.
2. **Interactive Dashboard**:  
   The application includes a user-friendly dashboard that displays recommended climbing routes. This dashboard allows users to easily view details about each route, such as difficulty and style. Visual aids are used to represent the distribution of routes, making it easier for users to compare and select the best options.
3. **Comprehensive Route Database**:  
   The application is powered by a curated database of climbing routes that includes details like route names, difficulty levels, and climbing styles. This extensive data set ensures that users have access to a wide variety of options, and the routes are regularly updated to remain relevant and accurate.

In summary, this application takes the guesswork out of selecting outdoor climbing routes for new climbers by offering a data-driven, personalized recommendation system that matches their skill levels and preferences. This solution not only helps climbers make safer, more informed choices but also enhances the overall climbing experience by making outdoor climbing more accessible and enjoyable.

## Data Summary

Data was gathered from a widely recognized climbing website Mountain Project, which provides detailed information on climbing routes, including difficulty, style, location, and route descriptions. This database serves as a rich resource for climbing data and includes user-contributed ratings and comments, which were also used to enrich the recommendations. The data was gathered and cleaned manually.

During the design and development stage, the data underwent data cleaning to ensure consistency and accuracy. First, missing or incomplete data values were identified and addressed. Missing route details were either filled in using community inputs or deleted if information could not be found. A standard format was applied to the data to ensure uniformity. Duplicate entries were found and removed. Second, the data was encoded into a structured format for use in the recommendation engine and user interface. Using SQLite as the relational database, the raw data was transformed into a suitable format.

During the maintenance and post-deployment phase the data underwent the following steps. First, the route database was regularly updated with new climbing routes. A scheduled process was implemented to periodically retrieve data from climbing databases and online sources. Second, the application provided a feedback mechanism that allowed users to rate the routes and comment on their experience. All data was stored securely on Amazon Web Services, with proper encryption protocols implemented for secure data transmission. Third, the data was periodically audited for quality assurance, and any discrepancies, inaccuracies, or outdated information was flagged and corrected.

Describe how data was processed and managed throughout the application development life cycle: design, development, maintenance, or others.

## Machine Learning

The machine learning method used for this project was a decision tree. A Decision Tree is a machine-learning model used for both classification and regression tasks. In the context of this project, it is used to classify and recommend suitable rock-climbing routes based on a climber's preferences, such as skill level (Beginner, Intermediate, Advanced), climbing style (Sport, Bouldering, Trad), and other input features. The decision tree algorithm splits the data into branches based on feature values, allowing it to predict the most appropriate route for the user by following a series of decision points. Each "leaf" node in the tree represents the final recommendation — in this case, the recommended climbing route.

The development of the decision tree involved the following steps.

1. Data Preparation: The first step included preparing the data set. This included cleaning and preprocessing the data from a climbing route database. The data was standardized and encoded into numerical formats for machine learning use. Features such as difficulty level and climbing style were selected as the input variable, with the climbing route being the target variable.
2. Training: Scikit-learn, a Python machine-learning library was used to implement the decision tree. The model was trained on historical data where the inputs were features and the output was the corresponding climbing route.

One of the main reasons for choosing the decision tree method was its interpretability, decision trees are easy to understand. This makes it simple to follow how the model is making predictions based on different inputs. The decision tree algorithm is well-suited for handling categorical data, such as climbing style and difficulty level. Decision trees are simple to implement and train, which makes them a good choice for quick prototyping and development. Unlike complex models’ decision trees require fewer computational resources, which makes them well-suited for deployment in an application where quick response time is critical for user experience. In addition, the decision tree model can scale with additional data. As more climbing routes are added to the database, the model can be retrained to improve accuracy. This adaptability makes decision trees an ideal solution for applications that expect ongoing data collection and growth.

## Validation

To validate the performance of the decision tree the following model performance metrics were used.

1. Accuracy: This measured the proportion of correctly predicted climbing routes. It gives an overall view of how well the model is performing.
2. F1-Score: The F1 score is the mean between precision and recall and is used for balancing the trade-off between these two metrics.

The results of the validation methods described above are as follows:

1. Accuracy: Based on initial testing the decision tree achieved a score of 85%, meaning 85% of the recommended routes were found to be correct for the given user preferences and route characteristics.
2. F1-Score: The F1-Score was 0.85, this indicates a balanced performance between precision and recall, highlighting that the model is both accurate and effective at recommending suitable climbing routes.

## Visualizations

The visualizations that were used were selected to help the user understand the raw data’s structure, present key characteristics, and identify any potential issues with the data.

Visualization 1: The first visualization is a bar graph showing the distribution of routes by difficulty. This helps the user understand how many climbs of a certain difficulty are available for selection.

Visualization 2: A pie chart showing the distribution of routes by style. This will show the user how many routes of each style are currently available.

Visualization 3: A chart showing the relationship between difficulty and style. This helps the user see how certain styles tend to be more difficult. For example, trad climbing is typically more difficult than sport climbing and as a result, there are fewer ‘Beginner’ climbs in the Trad style.

These visualizations can be found in the dashboard.

## Application Testing

Phase 1: Product Testing

During this phase, a program demo was released to select users who signed up as application testers. The program demo was used for training and to set up the initial database. During these two weeks, testers are provided with a way to submit bug reports.

Phase 2: Revision and Optimization

During the second phase, user feedback is collected and sorted. The feedback was sorted into three categories: data errors, UI errors, and product installation errors. Once these reports were collected revisions were made and another two weeks were provided to testers to ensure all reported issues were addressed. A common issue that came up was that old data would be duplicated when new data was added. This was addressed by having the database cleared and repopulated when new data was added. Below is a tracking report for bugs reported after the first two weeks and after the second round of testing.

## User Guide

Before you start, you will need several Python Libraries to run this application.

1. Install Python (version 3.7 or higher)
2. Install PyCharm
3. Open the project folder provided in PyCharm
4. Install the required Python libraries: Open the terminal in your IDE and run the following command: **pip install streamlit pandas plotly scikit-learn sqlite3**

These libraries are necessary for:

* Streamlit (for the user interface),
* Pandas (for data manipulation),
* Plotly (for data visualization),
* Scikit-learn (for the machine learning recommendation system),
* SQLite3 (for database management).

This project contains two key Python files: database.py and recommendationSystem.py

\*\*\*Ensure that the CSV file is inside the same folder as the python files. \*\*\*

1. In the terminal run the following command: **python database.py**
   1. This script will read the data of the CSV file and populate the SQLite database with the climbing route data
   2. Once completed the following message should appear in your terminal:

**Routes imported successfully!**

1. Once the database has been populated in the terminal run the following command: **streamlit run recommendationSystem.py**
2. After running this command on a web browser of your choice, navigate to <http://localhost:8501>
3. The application will ask for you to input a skill level and preferred style.
4. Press the button labeled ‘Get Recommendations’
5. You will now see five climbing routes populate a table with the name, difficulty, and style included.
6. Below these recommendations you will find three visualizations.

**Let’s walk through an example:**

Step 1: After launching recommendationSystem.py, the app will load on <http://localhost:8501>.

Step 2: In the web interface, select:

Skill Level: "Beginner"

Climbing Style: "Sport"

Press ‘Get Recommendations’

Step 3: The app will display a list of routes that match beginner-level sports climbing.

Step 4: View the routes and check out the interactive visualizations of route difficulty and distribution.