Empowering the Utility of Forestry Data Collections with OpenAI-Enabled Services in Dashboards

Degree programme : BSc in Computer Science
Thesis advisor : Prof. Dr. Ulrich Fiedler
Expert : Markus Nufer (Nufer Consulting AG Bern)

Clear and engaging dashboards facilitate data exploration, analysis and understanding. This makes it a valuable tool, helping to better apprehend the complexity of the forest ecosystem and its management. This bachelor thesis investigates the integration possibilities of OpenAI products in forest-related dashboards and examines the potential of use in the education of foresters.

Approach

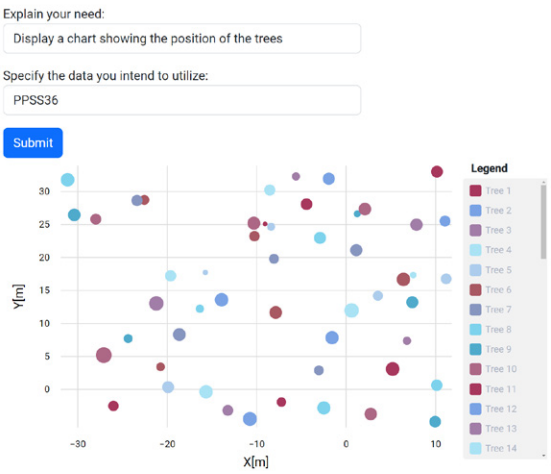
A comprehensive literature review was first conducted. Due to the rapid acceleration of the artificial intelligence research publications in the past 5 years, particular attention has been given to the most recent scientific publications and findings.
Based on the acquired knowledge, it became clear that OpenAI Large Language Models were the optimal choice for this thesis. The Generative Pre-trained Transformer (GPT) 3.5 models were chosen because of their high token limit, which is interesting since dashboards may require large data flows.
An Angular application was developed to assess the abilities of the GPT-3.5 models and to present a Proof of Concept, highlighting its innovative features and the value they add to dashboards. Most of the forest-related data utilized in the evaluation was obtained from the Martelage Sylvotheque (MSC) application database.

Assessment of GPT-3.5

Multiple test series were conducted and the first aimed to evaluate GPT arithmetic computation ability. It was asked to calculate the average diameter of trees from datasets of varying sizes and formats. GPT did compute correctly the average for small datasets in JSON and CSV formats. However, the task was more challenging for larger datasets in JSON and CSV formats, resulting in incorrect computations.
Another test series was performed to test GPT aptitudes to search for a word or a value in a CSV or JSON formatted dataset and to return the corresponding id(s). All tests succeeded when the dataset contained only one occurrence of the searched value/word but rarely when several values/words were present.
GPT showed good capabilities to determine which function and/or parameter needs to be used to process the user input. But it never responded correctly when a nested function call was required.

Proof of Concept Implementation and Evaluation

In the developed application, the user can write its need to display a tree map of a specific forest area. This input is then processed with the help of GPT-3.5 in order to display the appropriate chart. Two distinct requests are therefore made to the OpenAI API, each asking GPT to perform a specific task. The first task is to identify the desired tour within a dataset and to return its id. With this identifier, it is now possible to extract the corresponding tree data from the MSC database. This data is incorporated into the second request. The new task of GPT is to identify the function that should be called to display the chart. GPT is also asked to prepare and format the chart data needed by this function. This data and the name of the function are included in the response. Displaying the chart is then a straightforward process that only requires to call the function with the provided data.
This Proof of Concept showcases several benefits: users can directly input requirements via text, simplifying interactions with the UI and therefore improving user experience. It also simplified the UI development. However, the use of GPT-3.5 in more complex scenarios may be problematic, requiring the use of advanced techniques or additional tools.



Proof of Concept Example: Display the tree map of the forest area named «PPSS36»

David Paul Coutrot
Distributed Systems and IoT
d.coutrot@proton.me