**Proposal: WikiMood Search Tool**

**My vision for the project**

Imaging: you’re sitting in lecture. You’re listening, but you just want something to scroll on your laptop while you listen. Or maybe you’re just bored at home on a lazy Sunday with nothing to do; your friends are all busy, you’ve done all your homework, there’s nothing good on Tik Tok. You just need something to do to pass the time that’s mildly interesting. Enter: Wikipedia Mood Search Tool. This tool allows you to search for interesting Wikipedia articles to read based on your current mood. For example, sometimes you need a little pick-me-up, search for articles that are uplifting. You want to cry? Search for articles that are depressing. The possibilities are endless.

**What is exactly the function of your tool? That is, what will it do?**

My tool will allow a user to find interesting Wikipedia articles to read based on their current mood. Sometimes, you just want to read interesting articles, but finding articles that match your current mood is a hassle. Say you are feeling sad and you want to read sad articles (like how people listen to sad music when they are sad); my tool will allow you to find articles that are sad. Or, you’re feeling sad, but you want a pick-me-up. My tool can find articles that have an uplifting mood. Beyond the mood function, the user can also search for articles that pertain to a topic, for example they can search “baseball” with the “satisfying” mood to read about stories of players overcoming challenges.

**Why would we need such a tool and who would you expect to use it and benefit from it?**

This tool would provide an easier way to find Wikipedia articles. Its purpose is for user enjoyment, so it is something that can help make peoples’ days a little easier, which is something I believe everyone can appreciate. I searched, but I did not find any similar tools online, so this should be the first tool of its kind. I believe this tool should not be too complicated to build. The challenge will mainly come from how to correlate the different moods the user can pick from to articles that actually match that mood. I am planning on doing this by using keywords that are in the articles as trigger words for each mood, such as ‘triumph’ for uplifting, ‘death toll’ for depressing, and so on.

**How do you plan to build it? You should mention the data you will use and the core algorithm that you will implement.**

I plan to build this tool as a web application using Angular to create the user interface, and Flask to create the backend service which handles doing the document ranking and data processing. I will make use of the MediaWiki API to fetch around 100 random articles from Wikipedia, which I will then rank using BM25. I will also create a ‘mood score’ for each article based on keywords in the article so that the system can return the articles which best match the mood the user picks. The system will fetch around 100 articles to avoid having to store many GBs of Wikipedia articles locally, but still have the ability to rank them with BM25. If I fetched thousands of articles for example, doing BM25 would take a very long time. Whereas if I stored articles locally, I could do BM25 much faster, but I would need a lot of storage in my cloud hosting service once I deploy the project. Fetching 100 random articles should still give a good mix of moods and topics, while avoiding the previously mentioned downsides of the other implementations.

**What existing resources can you use?**

I will make use of Angular for the UI and Flask for the backend, along with the MediaWiki API to fetch Wikipedia articles, and Python libraries to help with processing and indexing the articles.

**How will you demonstrate the usefulness of your tool?**

I will demonstrate how the mood selected by the user affects the articles that my tool returns. I will also do user testing with multiple people to make sure the mood of the returned articles matches the mood that the user was expecting to see. I plan to have users rank the accuracy of the mood of returned documents on a scale of 1-5, and repeat this for multiple trials to ensure that the returned articles consistently matches the desired mood, even with the randomness of the articles fetched by the WikiMedia API.