About improvements:
1. First of all, we should focus on something. We need to know what we want. Do we want to find these mountain names in some local groups and chats where conversations are casual, or do we want to find them from some scans of scientific books? These really change the approach of creating the dataset, and as a result, we will get much better results as our focus won't be split.
2. Secondly, we shouldn't use ChatGPT as a dataset creator. It is okay, but at some point in time, he starts doing the same sentences with the same structure as before. Moreover, it does poorly on indexing mountain names, so it is not the best choice. What is more, it uses all the same mountain names. And prompts to change will not really work as for the new prompt GPT will again start with popular ones. As a result, the model overfits specific data and specific names. So we should at least find and review develop/test datasets so we would know how our model is actually working. But that point actually requires to know some initial information about why we are doing this project.

All in all, these changes should improve our model and make it viable in use.