CSE499A.15

Group 05

Student Name and ID

1. Emon Hossen 2211106042 (Herb2.0)

2. Faheem Hasnat 2211721642 (Dermnet)

3. Kazi Tanora Akther 2132580642 (Skincon)

Paragraph1:(about preprocessing)

In this stage, we focus specifically on the HERB 2.0 dataset, realizing that only the herb information file is relevant to our VLM project since it directly supports studying plant-based effects on skin diseases. We extract data solely from the herb information CSV, then filter and retain only the rows related to skin-related conditions while removing all unrelated entries for precise analysis.

Paragraph2:(purpose)

This refinement ensures the dataset aligns strictly with the project's objective understanding the connection between medicinal herbs and skin diseases. By filtering out irrelevant data, the model focuses on domain-specific information, improving accuracy and relevance for multimodal learning.