1) What is the overall goal of the project?

Skin cancer misdiagnosis by dermatologists is not uncommon. 25-33% of the skin cancers are incorrectly diagnosed as eczema or another less serious disease. The overall goal is to design and implement an application that can effectively classify whether a patient developed skin cancer based on photographs of skin lesions on dermatologic slides.

2) What data set you would use for the project with links?

Skin Cancer MNIST: HAM10000: 10015 dermatoscopic images https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000

3) Individual tasks?

Pre-processing of the images to obtain relevant features Ensemble of machine learning models Deep learning models Comparison of models using different metrics and limitations

4) How do you plan to integrate the project Collaboratively, work on pre-processing of the images to obtain relevant features and model performance or limitation of the model. One person can work on ensemble of machine learning model while other member can build deep learning models.

5) What will your novel aspect of investigations be beyond the basic ML pipeline. The application will run abnormal skin slides on a machine learning model and perform preliminary skin cancer checks to spot any abnormal lesion sites suspicious for skin malignancy, this will guide dermatologists and or pathologists to a focal point on the slide to allow for confirmatory diagnosis with greater confidence. In a way, this application facilitates in bringing attention to sites of immediate concerns, such as less differentiability of the cells through staining differences with the surrounding tissues, or shape irregularity.