Capstone 2 Project Proposal:
Skin Lesion Classification For Cancer Detection

Problem:
　　　　We are trying to classify skin lesions to determine if they are cancerous or not. Raw probabilities can be made available instead of a simple binary classification.

Context:
　　　　"Skin cancer is a major public health problem, with over 5,000,000 newly diagnosed cases in the United States every year. Melanoma is the deadliest form of skin cancer, responsible for an overwhelming majority of skin cancer deaths. In 2015, the global incidence of melanoma was estimated to be over 350,000 cases, with almost 60,000 deaths. Although the mortality is significant, when detected early, melanoma survival exceeds 95%." (ISIC 2018 | ISIC 2018: Skin Lesion Analysis towards Melanoma Detection)

　　　　Easy access to high quality automated screening could potentially save many lives by allowing people to get an assessment that is less expensive in both time and money than a doctor's visit. A computer vision model could easily fill this role if it were available through a smartphone.

Criteria of Success:
　　　　Producing a model that can be used to provide a probability distribution across classes with good enough accuracy and cross-entropy to be practically useful.

Scope of Solution Space:
　　　　We'll be creating one or more deep learning models using PyTorch for the classification of skin lesions. Ideally, this model should be accessible using a simple GUI that allows users to run inference on their own images.

Constraints in Solution Space:
　　　　Images in our dataset were taken using a specific procedure by medical professionals. As a result, laypeople taking pictures of their own skin lesions may produce images that lie outside the distribution of images in our dataset.

Key Data Sources:
　　　　We'll be using the HAM10000 dataset, which contains over 10,000 labeled images of skin lesions in 7 classes, some of which are malignant and others are benign.

The labels were determined by medical experts. More information about this dataset can be found below.

https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DBW86T
https://arxiv.org/pdf/1803.10417.pdf

Citations:
ISIC 2018 | ISIC 2018: Skin Lesion Analysis towards Melanoma Detection. challenge2018.isic-archive.com/.