Identifying Types of Skin Cancer from Image Data

CSC 371 (Machine Learning) Final Project by Ian Hall and Eric Xu
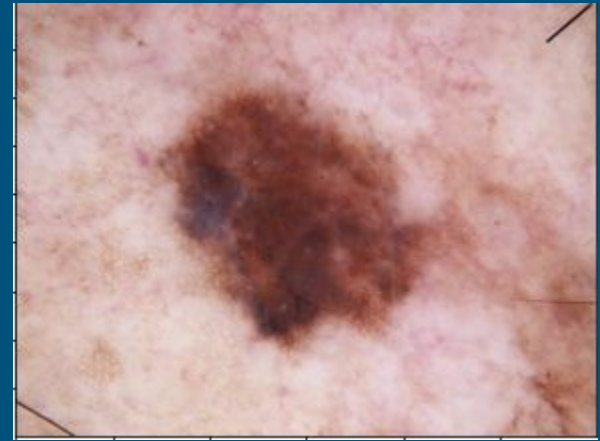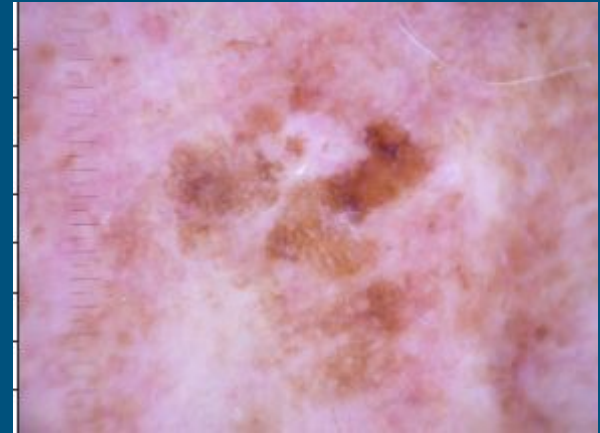
Overview of Project

Identifying type of skin cancer from image data

Typically, identifying the type of skin cancer requires the use of microscopes and general consensus from a group of doctors

Goal: create a model that can accurately identify skin cancer type from an image to simplify the problem
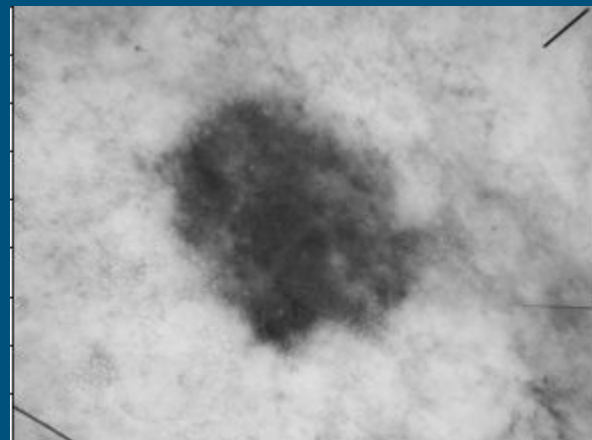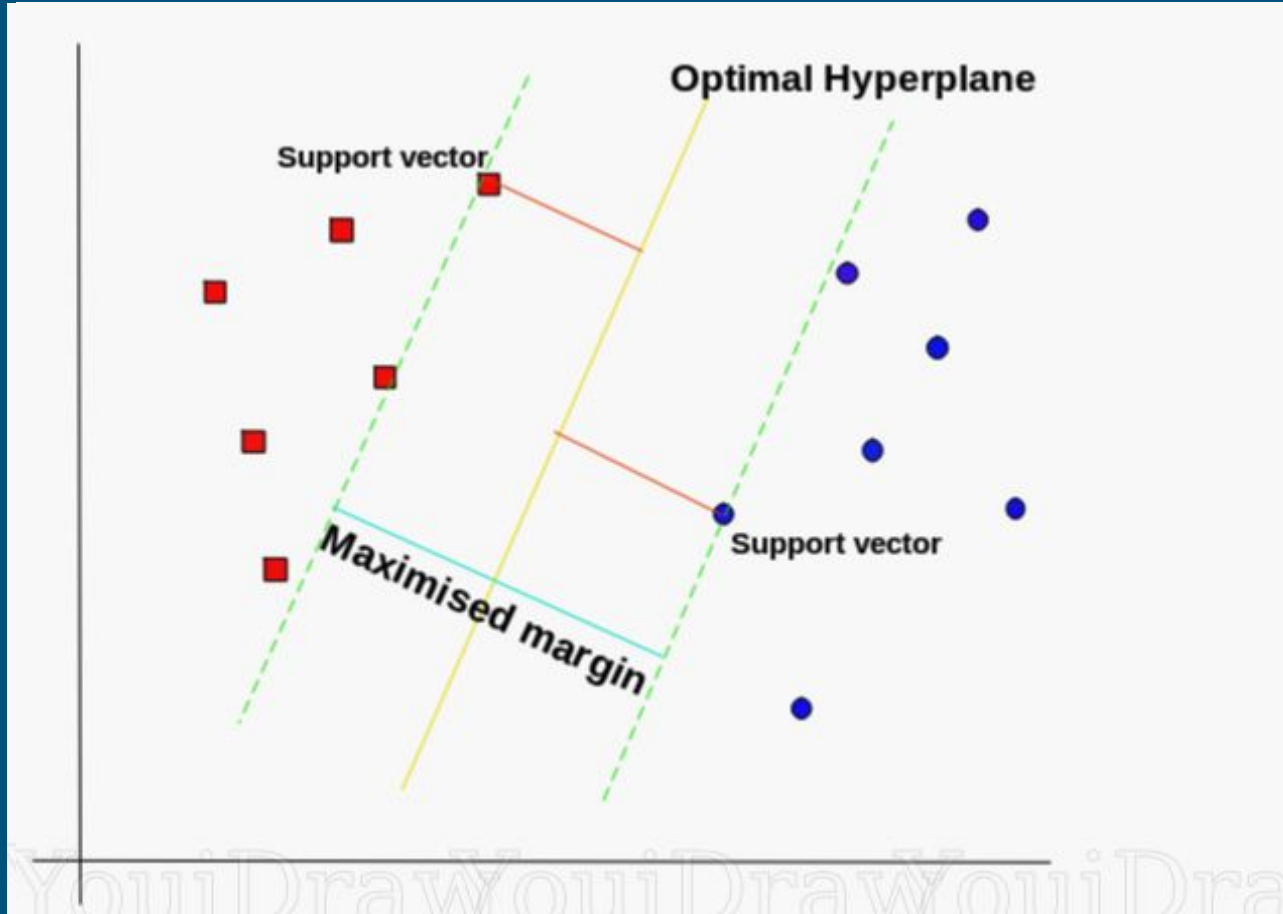

Melanoma example

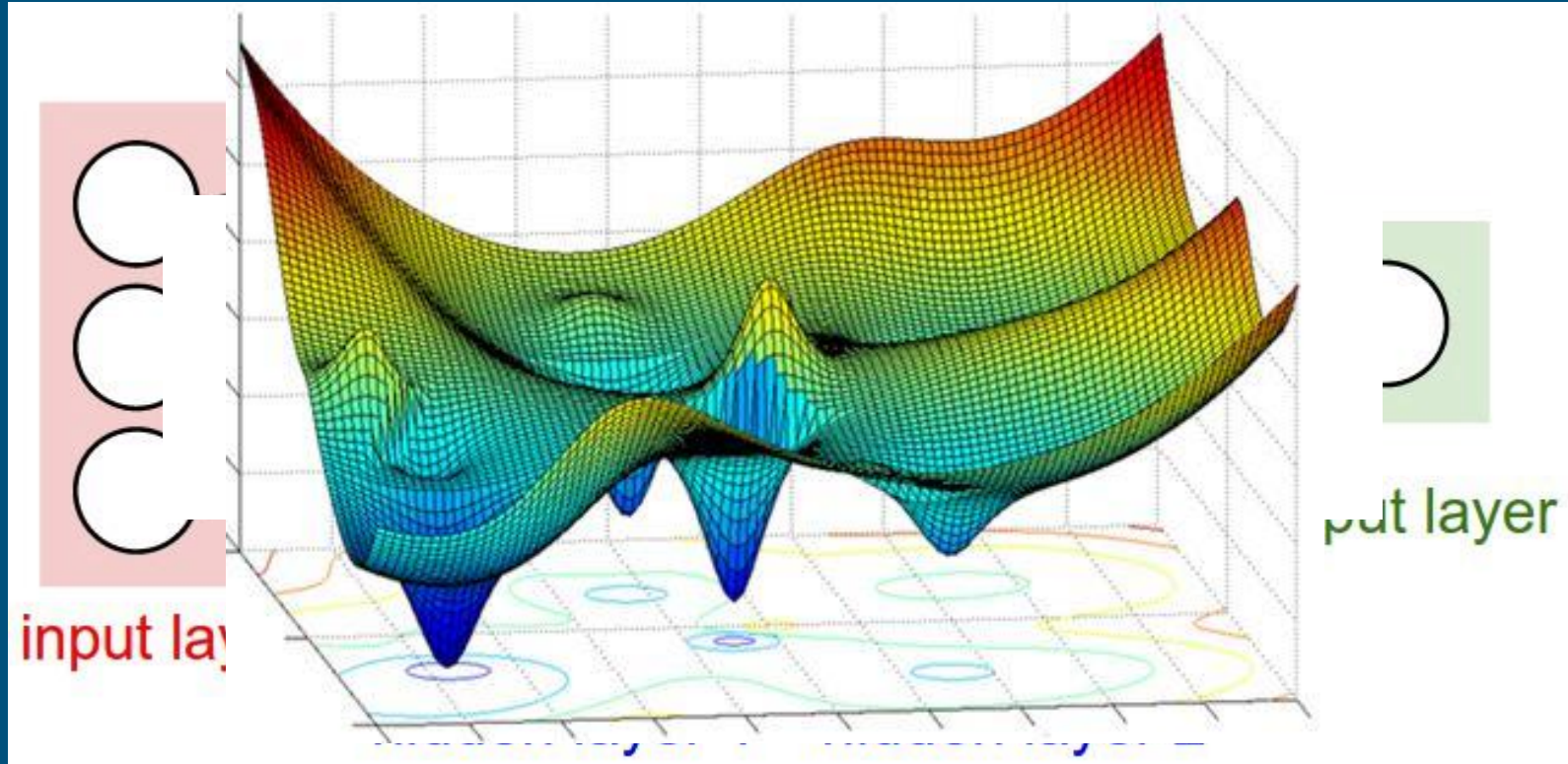benign keratosis-like lesions example

Data

- 10015 unique data points
- Each data point consists of:
  - Patient ID Number
  - Picture of the Skin Cancer: 450x600
  - Type of Skin Cancer (7 types)
  - How the Skin Cancer was identified
  - Age of Patient
  - Sex of Patient
  - Where the Skin Cancer was located
- Focus: Purely, on Image Data

# Support Vector Machines

Neural Networks

Preliminary Results

	Support Vector Machine	Neural Network
Accuracy on Validation Set	60.4%	73.2%

- Created preliminary models for both methods without any hyper-parameter tuning.
- Appears that Neural Networks work much better
- More experiments needed to produce stronger models

Future Work

- Hyper-parameter tuning
  - Regularization
  - Convergence
  - Learning Rate
- Transfer Learning
  - Using ResNet image classification model

Source

- https://towardsdatascience.com/https-medium-com-pupalerushikesh-svm-f4b42800e989
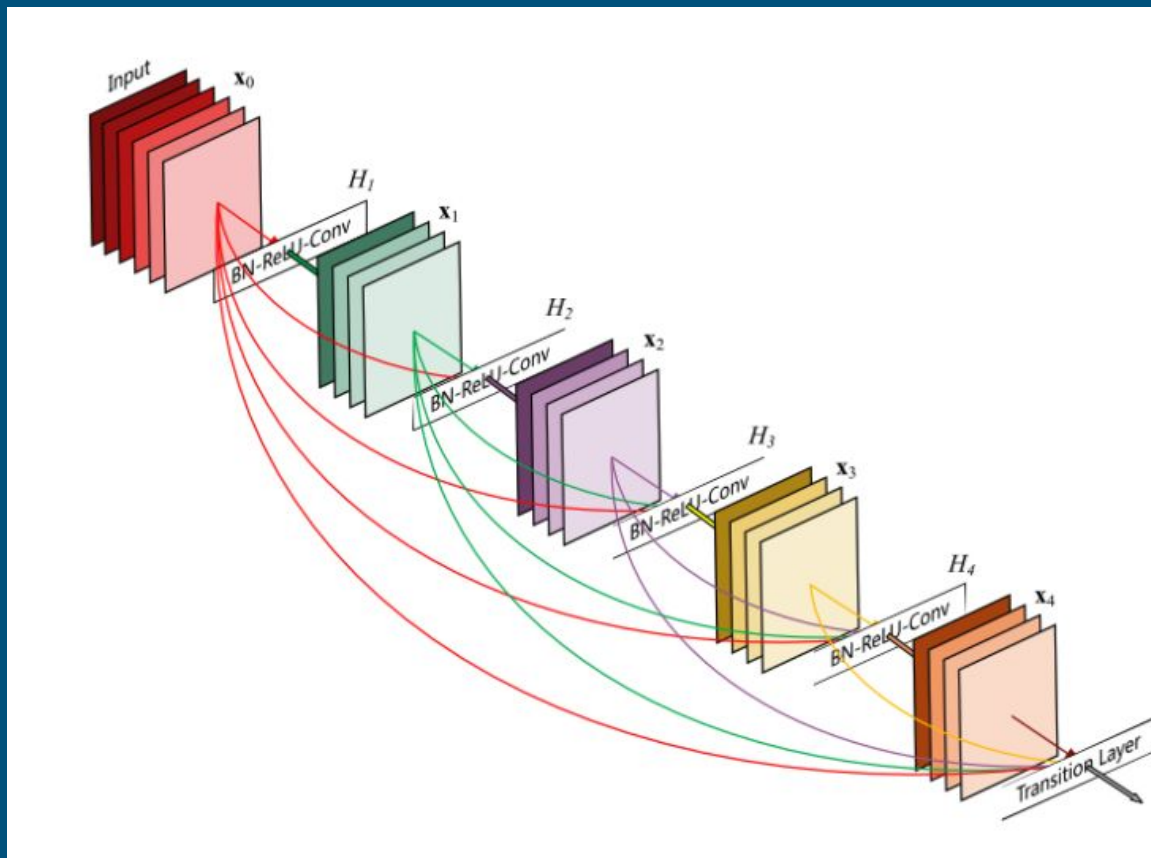
- https://www.sas.com/en_us/insights/analytics/neural-networks.html#:~:text=How%20Neural%20Networks%20Work,neuron%20in%20a%20human%20brain.

- https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DBW86T

- https://towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035