Runtime Terror
Alexis Babb, Taizhou Qing, Ashish Roy, Zhiyu Zhou

Project Link Here

Project Motivation / Project Importance
Invasive ductal carcinoma (IDC) is - with ~ 80 % of cases - one of the most common types of breast cancer. It's malicious and able to form metastases which makes it especially dangerous. Often a biopsy is done to remove small tissue samples. Then a pathologist has to decide whether a patient has IDC, another type of breast cancer or is healthy. In addition sick cells need to be located to find out how advanced the disease is and which grade should be assigned. This has to be done manually and is a time consuming process. Furthermore the decision depends on the expertise of the pathologist and his or her equipment.

This project was motivated by our desire to leverage deep learning to automatically detect and locate tumor tissue cells and to speed up the process of breast cancer diagnosis. In order to exploit the full potential in the future, one could build a pipeline using massive amounts of tissue image data of various hospitals that were evaluated by different experts. This way, the medical community would be able to overcome the dependence on the pathologist which would be especially useful in regions where no experts are available.

Project Goals
From a mission-driven perspective, our primary goal was to:
  ● Develop a deep learning model which can determine if a breast lump is cancerous or not (without having to perform a biopsy or rely on a doctor's expertise).

From a data science perspective, our goals were to:
  ● Select a performance metric algorithm (Selected: AUC, chosen as it is a way to evaluate classifier performance and takes into account different threshold values, thus providing)
  ● Build three deep learning models (Selected: CNN, ANN, RNN)
  ● Tune hyperparameters of the deep learning models (Selected: batch size & epochs, optimized via GridSearchCV)
  ● Visualize the results of the models (Selected: Confusion matrices and AUC-ROC curves)
  ● Evaluate the models and select the one with the best performance (Selected: CNN, due to highest AUC)