# MLE-10 One-Pager

| **Project Title** | Automating Explainability in Healthcare |
| --- | --- |
| **GitHub** | https://github.com/Timecapp/XAI\_Capstone |
| **Presentation** | https://docs.google.com/presentation/d/1rI4byBdGlJn\_1nwsFZLaZarMxek-hsHrtHivRO4l0WU/edit?usp=sharing |
| **Member names** | Pedram, Shai |
| **Submission Date** | February 5, 2022 |

## 

### Problem definition

Diagnosing breast cancer requires a gold standard histopathological confirmation; conducted by a pathologist who microscopically visualizes dissected tissue. The analysis of the specimen results in a diagnosis and stage if positive, with a justification of observed findings. AI models trained on object detection/segmentation tasks are currently used as a diagnostic tool, as they provide precise and consistent predictions. However, the explanations from these models, such as heatmaps, similar images and high level features, do not match those of an experienced pathologist. Explainable features that are suitable for histopathological image diagnosis need to be defined. The data annotations and choice of models needs to be designed to generate accurate diagnostic predictions with high quality explainable features. The annotation and explanation process involved could be more precise and consistent with accurate diagnostic algorithms that identify specific and measurable morphological abnormalities that reflect a carcinoma based on objective annotations.

### Objectives

The intention is to allow end-users to upload digital histopathology and acquire a diagnosis without bias (without personal identifiers in metadata), with an explainability model to optimize interpretability.

### Methodology

We began the dataset exploration with the Samsung issued Canine Breast Histopathology dataset, which we realized did work on an image feature selection basis, but only complicated the black box issue as there required a translation of histopathological features between canine to human species. We proceeded to establish a dataset source - which we did 2 of; the first with the Wisconsin Breast Cancer Dataset (csv) and the second with the BreakHis dataset (whole slide images). We conducted EDA on both and then proceeded to run a baseline (BreakHis) based on Yu Huang’s use of LIME to explain the image prediction results of a 2D Convolutional Neural Network (ConvNet) for the Invasive Ductal Carcinoma (IDC) (Huang, 2020). Once the explanation of the model prediction was obtained, we applied get\_image\_and\_mask() to get the template image and the corresponding mask image (super pixels) showing the boundary of the area of the IDC image in yellow. Then we applied a LIME image explainer to explain the IDC image prediction results of the 2D Convoluted Network model of IDC of the breast from digital pathology slides. Having completed this task and achieving 86.9% accuracy on image features and fulfilled LIME explainability, similar to the author’s results, we proceeded to run K fold logistic regression, and SVM on the BreakHis dataset and achieved a similar 86% accuracy in detection. In our EDA we explored univariate, bi variate and multivariate analyses as well as used heat maps and correlation matrices to establish what we were getting right and wrong. We then tried several models on the Wisconsin Breast Cancer Dataset and after EDA and baseline, we tried a Logistic Regression model with an accuracy of 97.6%, Naive Bayes at 94%, and SVM much higher at 98.2%, and Random FOREST AT 94.7%, where the Decision Tree training set accuracy was 0.99 and test set accuracy was 94.7%.

Then we tried to address the data imbalance, so we oversampled the minority class using the Synthetic Minority Oversampling Technique (SMOTE), and using Standard Scaler, and Log transformed skewed data, removed outliers with z-score, we achieved Random Forest accuracy is 96.3%, and successfully applied LIME, SHAP MOrris sensitivity partial sensitivity plots and gaussian ROC curves (0.99 AUC).

We are now deploying the model and writing up the presentation and associated notebook to share the comparison of our models’ outputs.