

Weekly Meeting with Dr. Hannah

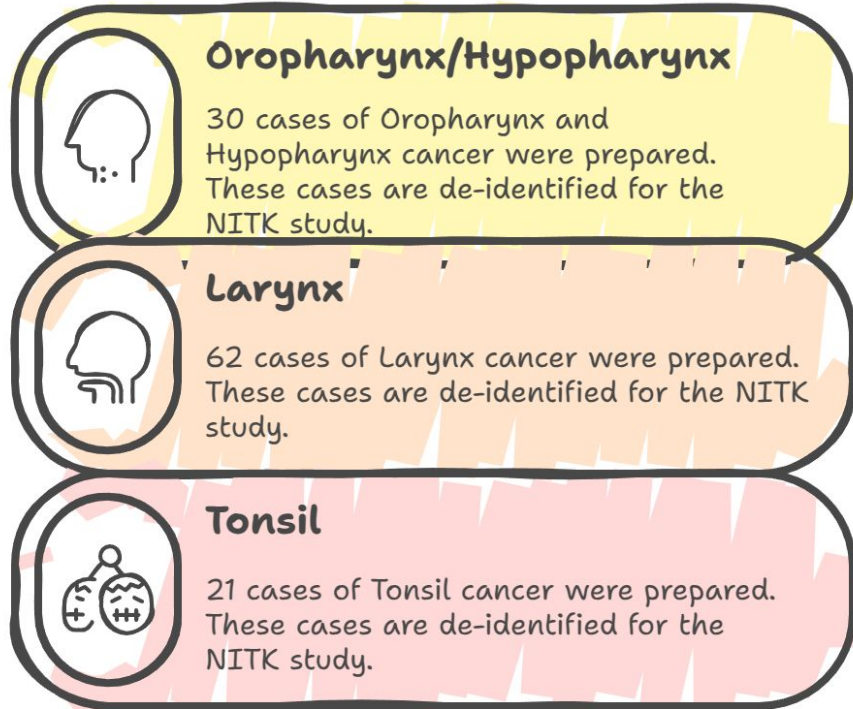
Presenter: Hasan Shaikh
Quantitative Imaging Research and
Artificial Intelligence Lab (QIRAIL)

ML Model Optimization Process

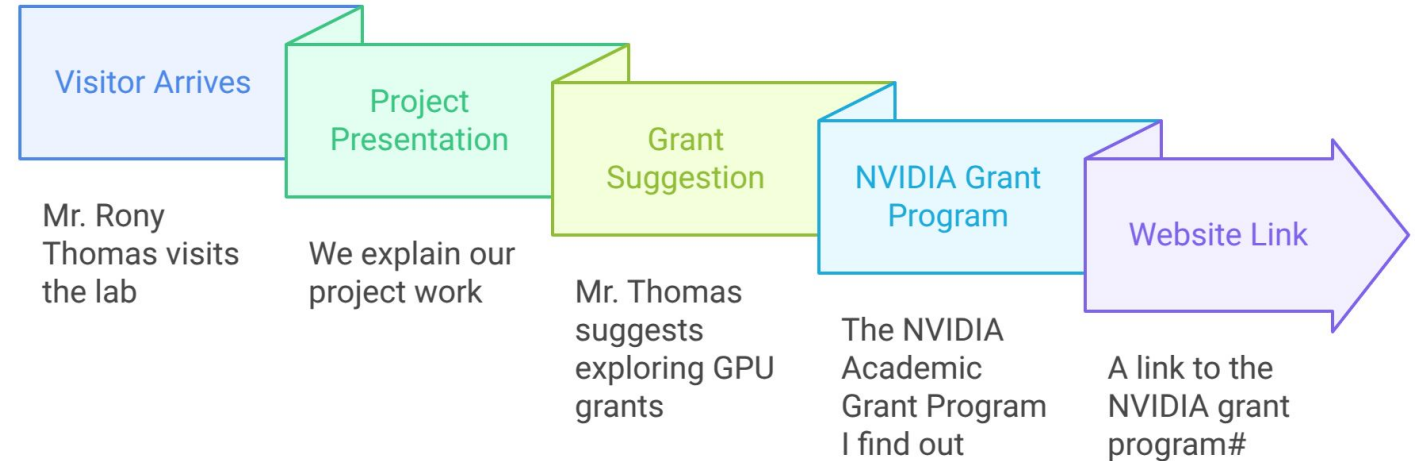
1	Implement Multi-Core Processing Speed up training by using multiple cores
2	Apply Grid Search Tune hyperparameters using a systematic search
3	Apply Bayesian Optimization Tune hyperparameters using a probabilistic approach
4	Test Additional Classifiers Evaluate performance of XGBoost, LightGBM, and MLPClassifier
5	Identify System Bottleneck Recognize system slowness as a limiting factor
6	Verify Data Split Ensure test data is excluded during training

* All codes have now been modularized into functions for streamlined workflow and clarity

De-Identified Head and Neck Cancer Data for NITK



NVIDIA Academic Grant Program



#Website link: <https://www.nvidia.com/en-in/industries/higher-education-research/academic-grant-program/>

Project Contributions Timeline: September 2024 to March 2025

Start of project contributions

September 2024

Completion of project contributions

March 2025



Selected Feature Techniques: Particle Swarm Optimization (PSO)

Classifier	Training ROC AUC	Training Accuracy	Test ROC AUC	Test Accuracy
Logistic Regression	0.63	0.54	0.57	0.59
Naive Bayes	0.77	0.51	0.58	0.42
Linear Support Vector Machine (Linear SVM)	0.87	0.65	0.48	0.63
Radial Basis Function Kernel SVM (RBF SVM)	0.87	0.64	0.44	0.63
Decision Tree	0.84	0.82	0.64	0.63
Random Forest	0.99	0.95	0.49	0.47
Voting Classifier	0.99	0.82	0.48	0.63



Planned Future Work – Classifiers vs Feature Selection Techniques

In the upcoming weeks, I plan to test the following classifiers:

1. Logistic Regression
2. Naive Bayes
3. Linear Support Vector Machine (Linear SVM)
4. Radial Basis Function Kernel SVM (RBF SVM)
5. Decision Tree
6. Random Forest
7. Gradient Boosting (e.g., XGBoost, LightGBM)
8. Voting Classifier
9. Stacking Classifier
10. Multi-Layer Perceptron Classifier (MLPClassifier)

These classifiers will be evaluated in combination with the following feature selection/optimization techniques:

1. SelectKBest (Univariate statistical selection)
2. Least Absolute Shrinkage and Selection Operator (LASSO)
3. Particle Swarm Optimization (PSO)
4. Whale Optimization Algorithm (WOA)
5. Grey Wolf Optimizer (GWO)
6. Genetic Algorithm (GA)
7. Simulated Annealing (SA)

This results in a total of **70 unique combinations (10 classifiers × 7 feature selection techniques)** that I plan to implement.

Selected Feature Techniques: Least Absolute Shrinkage and Selection Operator (LASSO)

Classifier	Training ROC AUC	Training Accuracy	Test ROC AUC	Test Accuracy
Logistic Regression	0.69	0.68	0.35	0.63
Naive Bayes	0.63	0.68	0.70	0.68
Linear Support Vector Machine (Linear SVM)	0.61	0.65	0.73	0.63
Radial Basis Function Kernel SVM (RBF SVM)	0.64	0.66	0.68	0.63
Decision Tree	0.87	0.80	0.27	0.47
Random Forest	0.99	0.92	0.33	0.53
Voting Classifier	0.99	0.82	0.48	0.63