Group #16

Machine Learning and Visual Analytics for Cell-based Biomarker Diagnostics

Team Members: Isabella Garcia and Pranav Siva
Mentor: Dr. Yu “Max” Qian, J. Craig Venter Institute

Abstract

Leukemia is a common blood cancer of the white blood cells, with subtypes exhibiting significant phenotypic heterogeneity that leads to irreproducible diagnosis and classification by manual gating. Analyses of the flow cytometry pipeline through normalization and visual analytic methods were conducted and quantified through the development and usage of a novel machine learning approach that requires no manual intervention. Testing revealed that the best model can accurately classify chronic lymphocytic leukemia in a 16-channel panel approximately 81% of the time, with F1-score 0.8. We conclude that normalization through cyCombine has an inconclusive effect on this classifier, that UMAPs demonstrate sample heterogeneity, and that a classifier trained on distribution statistics can achieve good performance.