Naive Bayes Artificial Intelligence Diagnostic Tool for Lung Nodule Cancer Detection

Authors: Noah Black, Mathew Alexander

Herbert Henry Dow High School, MI, Michigan Southeastern JSHS;

Teacher/Mentor/Sponsor: Marc Black, Dow Chemical Company

Lung nodules pose a substantial risk to patients. Lung nodules are ovular masses of tissue that propagate in the lungs of many patients. It’s key that malignant nodules are detected early because mortality rates rise dramatically with the size of the tumor. Typical diagnoses can take months to a year of repeated CT scans. Our research aims to curtail the time from recognition to diagnoses by using a cost-effective protein analysis of peripheral blood. By analyzing the blood protein expression levels of patients with malignant and benign lung nodules with a publicly accessible dataset on Gene Expression Omnibus, we created a Naive Bayes artificial intelligence (AI) capable of predicting the cancer status of patients with a lung nodule. The data was organized in R and Python. And, by using third party packages, we developed the AI in python. Ultimately, the AI had an efficacy of over 80%, which as a diagnostic tool categorizes it from “good” to “very good.”