

Machine learning approach in MALDI-TOF SARS-COV-2 detection

Kadyrova Irina

Kazakhstan

Machine learning methods (ML) are often used to solve medical problems. We used ML to search the mass spectrum for structures corresponding to the distribution of molecular peaks characteristic of SARS-CoV-2 in order to expressly detect SARS-CoV-2. However, RT-PCR is considered as the gold standard for detecting viral RNA in nasopharyngeal swabs (NS), but it is time-consuming and requires constant changes in the primers composition due to the mutation of SARS-CoV-2 strains. We propose a method for the detection of SARS-CoV-2 in NS using MALDI-TOF MS and the ML approach. Methods. NS from patients with PCR-confirmed Covid-19 and control participants were tested (130 and 80 swabs respectively) with MALDI-TOF MS MicroFlex LT (v 3.1, Bruker Daltonics, Bremen, Germany) using HCCA matrix. MALDI spectra were preprocessed in R package using wavelet smoothing, SNIP algorithm baseline removing, filtering of the samples on intensities and signal-to-noise ratio; calibration was done on the PQN normalization. Peak detection was performed with MAD method. Peak alignment was performed using 70% minFreq and 0.005 tolerance. Results. ML approaches included GLM, RT, RuleFeat Ensemble Model. Obtained models was characterized with specificity, sensitivity and F1 score. GLM (specificity= 1, sensitivity=0.5) model showed low F1 score (0.4). However RT and RuleFeat Ensemble Models demonstrated sensitivity, specificity and F1 score equaling 1 after optimization. Conclusion. Thus, we propose a screening method for SARS-CoV-2 detection (sensitivity 1, specificity 1). The developed methodology combines the analysis of NS samples using the MALDI-TOF-MS with ML approaches, is suitable for screening patients with COVID-19 at the first stages of diagnosis. RT and RuleFeat Ensemble Models demonstrated high sensitivity, specificity and F1 scores.

Biography

Irina Kadyrova has her own experience in valuation and passion for ML and data. The research team built this model after many years of experience in research, evaluation, work in both hospitals and scientific laboratories. This approach meets all the requirements for precise, specific, sensitive diagnostics.

irina.adilevna@gmail.com

Abstract received : July 19, 2022 | Abstract accepted : July 19, 2022 | Abstract published : 13-12-2022