



Prediction of SARS-CoV-2 infection status and COVID-19 disease severity based on clinical data with Machine Learning

Zhakparov D^{1,2,3}, Schmid M⁴, Styrzynski F⁵, Roelke H⁴, Solek J⁵, Krzysztof J⁵, Lukasik Z⁵, Roqueiro D⁶, Makowska J⁵, Sokolowska M^{1,2} and Baerenfaller K^{1,2,3}.

¹ *Swiss Institute of Allergy and Asthma Research, 7065 Davos Wolfgang, Switzerland*

² *University of Zurich, 8006 Zurich, Switzerland*

³ *Swiss Institute of Bioinformatics, Lausanne, Switzerland*

⁴ *University of Applied Sciences of Grisons, SII, 7000 Chur, Switzerland*

⁵ *Department of Rheumatology, Medical University of Lodz, Lodz, Poland*

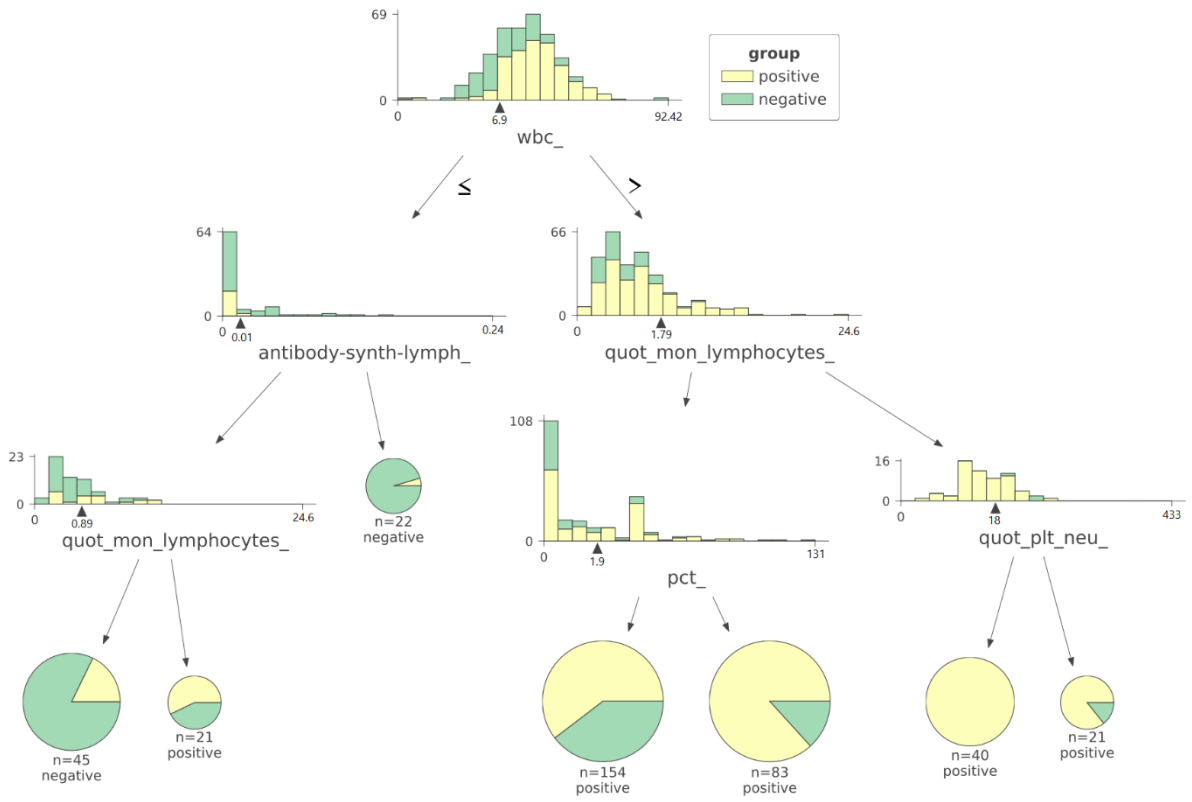
⁶ *Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland*

In the current COVID-19 pandemics, it is of critical importance that SARS-CoV-2 positive patients can be reliably identified already at hospital admission. Moreover, it is important that disease progression of COVID-19 patients is monitored to allow for appropriate and timely treatments.

The aim of our study is to build a decision tree based on clinical data using Machine Learning to find predictors of a possible SARS-CoV-2 infection upon hospital admission. Additionally, we want to see the progression of features that lead to severe illness and a negative outcome.

We trained several models based on different estimators on a dataset containing patient information on comorbidities, pre-medication and on symptoms and laboratory parameters at hospital admission and over the course of the disease for 201 SARS-CoV-2 positive and 314 negative subjects with a COVID-19-like clinical picture at hospital admission. The longitudinal clinical laboratory data was parametrized to observe how these patterns are associated with disease severity and outcome.

We have identified a set of important features that are able to differentiate between SARS-CoV-2 positive and negative patients, and between COVID-19 patients who survived or deceased. Our study therefore managed to provide clinicians with valuable information on a patient's possible status, based on which the treatment can be planned.



Medical decision tree based on the most important features to predict the SARS-CoV-2 infection status