

AHEAD Medicine Interview Questions

Programming Task (1-2 hours on a laptop)

You have familiarized yourself with the FCS data files and relevant documents we shared with you prior to the interview. As well as the Python library FlowCal to read and manipulate the FCS data. Please put your answers on GitHub and share the link with us.

TASK: There are a group of patients who were diagnosed either COVID-19 positive (sick) or negative (healthy). Each FCS file represents the specimen collected from one patient. Build an automatic predictor using a ML model of your selection, the labels provided in the

“EU_label.xlsx” as ground truth, and marker-channels with “use” = 1 in

“EU_marker_channel_mapping.xlsx” as data features.

NOTE: This dataset is NOT PERFECT. Do not spend too much time tuning the model for performance. We focus more on how you build the workflow and get your results.

Bonus Question

1. Please explain the fundamental principles of flow cytometry and walk through the step-by-step process of how it works? Additionally, highlight some common applications of flow cytometry in scientific research and clinical settings.
2. Below are plots of selected cell surface biomarkers of blood cell samples. Researchers are interested in picking out cells marked in yellow (occupying a high-density chunk at the bottom-right) for further analysis. How would you suggest a method to automatically identify these cells?



