Technical Challenge – Data Science

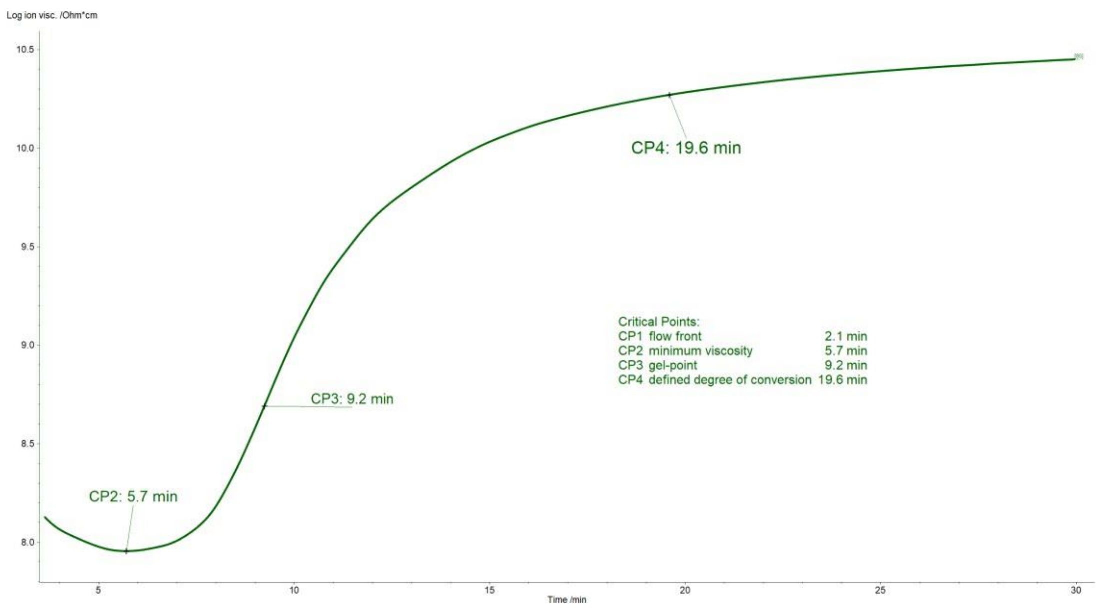
The provided data files include DEA (Dielectric Analysis) measurement data of an epoxy resin system.

Each file contains the data of a single measurement with the following information:

- **time_/min:** Measurement time in minutes
- **temp/°C:** Temperature in °C
- **DSCalpha:** Degree of cure; it describes the conversion achieved during crosslinking reactions (curing) and is defined here as a number between 0 and 1.
- **Impedance 1.78 kHz/Ohm:** DEA signal impedance (frequency 1,78 kHz); the impedance signal gives information about the material behavior during the measurement

(**Tg/C** and **Derivative Y1** are <u>not</u> relevant for the tasks.)

During the curing process, we are interested in specific points on the DEA measurement curves that give us information about the material behavior. We define these points as **Critical Points (CP)**. In the following, we want to focus on three of these points:



CP2: minimum of the viscosity of a material (simplified definition)

CP3: to simplify the task, we define CP3 as the inflection point after the minimum of the DEA curve

CP4: indicates when the degree of cure (**DSCalpha)** of 95% has been reached

Task 1: Data Analysis

- Determine CP2 of the impedance curve (As mentioned in the description above we assume that CP2 is the minimum of the curve.)
- Find CP3 on impedance curve (Consider noise in the signal and handle it appropriately)
- Plot the impedance curve based on time with a mark at CP2 and CP3 for ten different measurements.

Task 2: ML Modelling

When curing a polymer material, it is important to get the timing right. CP2 represents approximately the beginning of the curing (or hardening) process, while CP4 is at the end of it. We provide our customer with a recommendation for the duration of the curing process. To do that, we predict the time difference between CP2 and CP4 using the data of the time window starting at CP2 and ending at CP3. Your task is to come up with a model that does just that!

A More Formal Definition

$\Delta T = t_{cp4} - t_{cp2}$ denotes the time difference between CP4 and CP2.

$X_{cp2,3}$ is the data in the time window between CP2 and CP3.

Your task is to develop a data-driven model that predicts ΔT based on $X_{cp2,3}$. In other words, your model should approximate a function F such that $\Delta T \cong F(X_{cp2,3})$ using a data-driven approach (e.g. machine learning). Please use at least impedance data as an input of your model, but feel free to include the other dimensions of the dataset (e.g. temperature) as well.

Note:

- You are free to choose the type of model. It should be appropriate to the problem in terms of complexity, memory usage, and performance.
- Evaluate the quality of your model using best practices and present the results.
- If you did not finish task 1, use a fixed-length time window starting at CP2 as the input of the model.

Submission

Please implement your solution in a stable version of Python. You are free to use third party libraries.

The code you submit should be runnable under Windows. Please include instructions how to setup the environment and run your solution. Remember to provide a requirements file specifying the versions of your external dependencies.