Project 2

Due March 25th, 2024

11:59 PM 100 points

CS 4432/5599

MATH 4463/5563

Data Science and Applied Machine Learning

Polymer Melt Flow Rate

Polymer properties such as density, melt index, and melt flow rate must be kept within tight specifications for each grade. This project is to analyze polymer production data to predict melt flow rate.

Background: There are gas phase and liquid slurry reactors that create polymers (polyethylene, polypropylene, polystyrene, and others) from chemical building blocks known as monomers (C2=, C3=, C4=, iC5=, and others). A catalyst is injected with the monomers under carefully controlled temperature and pressure conditions to cause a reaction that grows the polymer chains. Hydrogen is a chain transfer agent to stop the growth of the polymer chain. If the polymer chains grow too long then the polymer is too viscous for manufacturing in films, injection molding, or other applications.

This case study focuses on measurements of Melt Flow Rate (MFR) to determine the polymer viscosity based on reactor conditions. An accurate model is desirable so that the infrequent lab samples (every 2-8 hours) are supplemented with a virtual and continuous "soft sensor". A model that runs in real-time simulation alongside the physical reactor is called a digital twin.

Label	Data File Tag	Description
Time		Timestamp of the measurements
C3	513FC31103.pv	Propylene (C3=) Feed Rate (kg/hr)
H2R	513HC31114-5.mv	Hydrogen to C3= Ratio
Pressure	513PC31201.pv	Reactor Pressure (bar)
Level	513LC31202.pv	Reactor Bed Level (m)
C2	513FC31409.pv	Ethylene (C2=) Flow (kg/hr)
Cat	513FC31114-5.pv	Catalyst Feed Rate (kg/hr)
Temp	513TC31220.pv	Reactor Temperature
MFR	MFR	Melt Flow Rate (gm/10min)

References

Hedengren, J. D. (2021, December 16). Polymer Melt Flow Rate, Machine Learning for Engineers. Retrieved from http://apmonitor.com/pds/index.php/Main/PolymerMeltFlowRate

1. Using your team's knowledge of on basic statistics, regression, and machine learning procedure to perform an analytical study on the PolymerMeltFlowRate dataset posted to Moodle. Prepare a 10-minute presentation based on your analysis. You are welcome to use any additional data/resources, but make sure to discuss them in the presentation and include them in your submission. The scope and focus of the presentation are up to you as a team, however, keep the following guidelines in mind:
    a. All team members must be part of the presentation; any team member who does not actively participate in the presentation will not receive credit.
    b. Each individual team member is required to present for a minimum of 5 minutes to ensure equitable distribution of presentation responsibilities.
    c. Your presentation should state the goals of your research, i.e. what is the objective(s) of your analysis.
    d. Your analysis should at least include the following components:
        i. Exploratory Data Analysis and Preprocessing
        ii. Feature Engineering/Selection
        iii. Regression Model(s)
        iv. Training and Testing of the Model(s)
        v. Evaluation and Comparison of the Models
    e. Teams are encouraged to explore and utilize different methods, such as polynomial, Ridge, Lasso regressions, or combined techniques.
    f. Presentations should include the following sections:
        i. Introduction/Background
        ii. Goals of Research (What specific insights or questions do you aim to address through your analysis?)
        iii. Methodology (i.e. what analysis did you preform)
        iv. Results
        v. Analysis of the Results
        vi. Conclusions/Reflection (i.e. what did you learn from the analysis)
2. Submit a video with your presentation, any material you used during the presentation, such as PowerPoint slides, any code that was used to perform your study (well-documented and commented) and if necessary, any other data/resources used (discussed in the presentation).

Notes:

1. All four professors will score your presentation out of 25 points, your team's final grade will be the sum of those scores.
2. Your grade will factor the quality of your presentation, analysis, and visualizations, as well as the depth of your analysis.

3. After all submissions have been graded, your presentations will be posted for all students to view on Moodle.
4. This project will account for 10% of the final grade in this class.
5. All data analysis/visualization must be done using Python.
6. Teams have been assigned for this project (see attached file).