Brake pad recipe creation using Gen AI

Mantek Singh

ML engineer @ Google Ads

Vineet Shah

ML engineer @ Google Bard

Prateek Karnal

Software engineer @ Databricks

Gagan Ganapathy

Software engineer @ sumo logic

Team Elleman

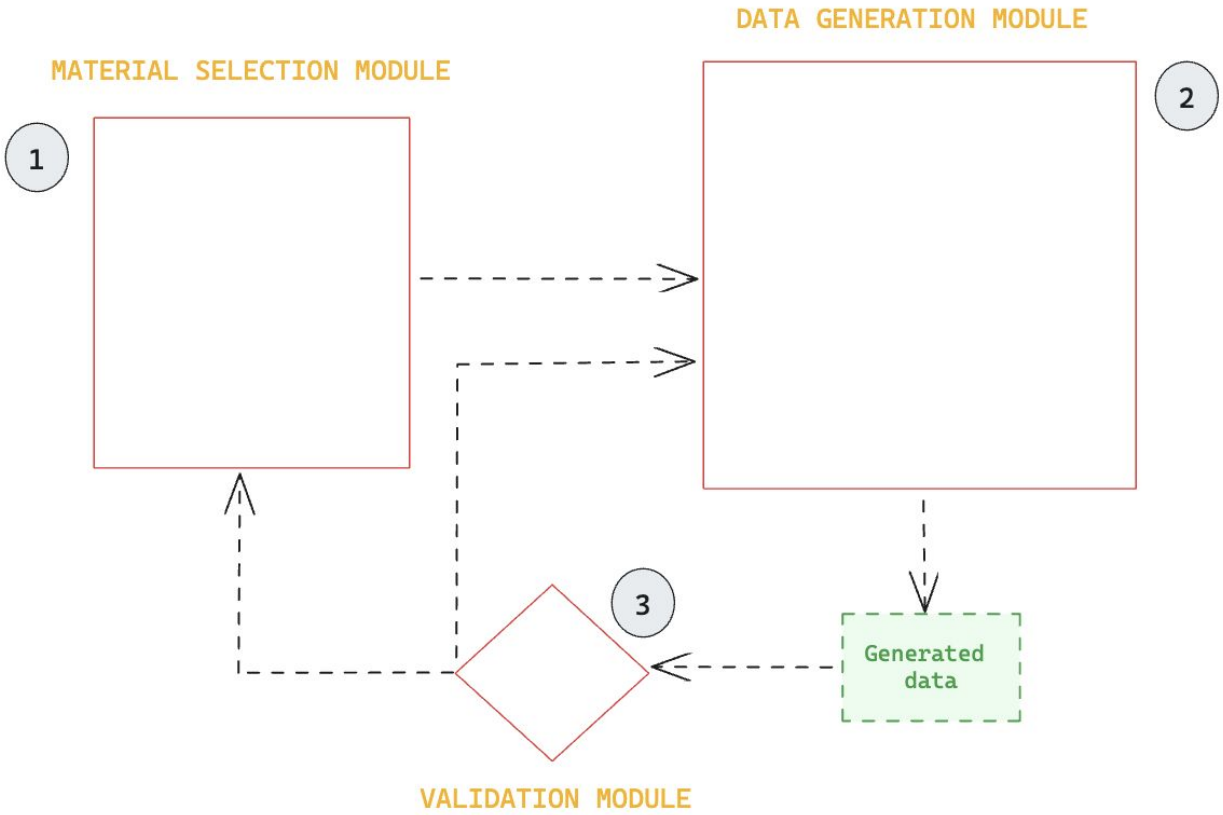
Motivation

Potential for high Impact

Large Language Models (LLMs) perform good on time series

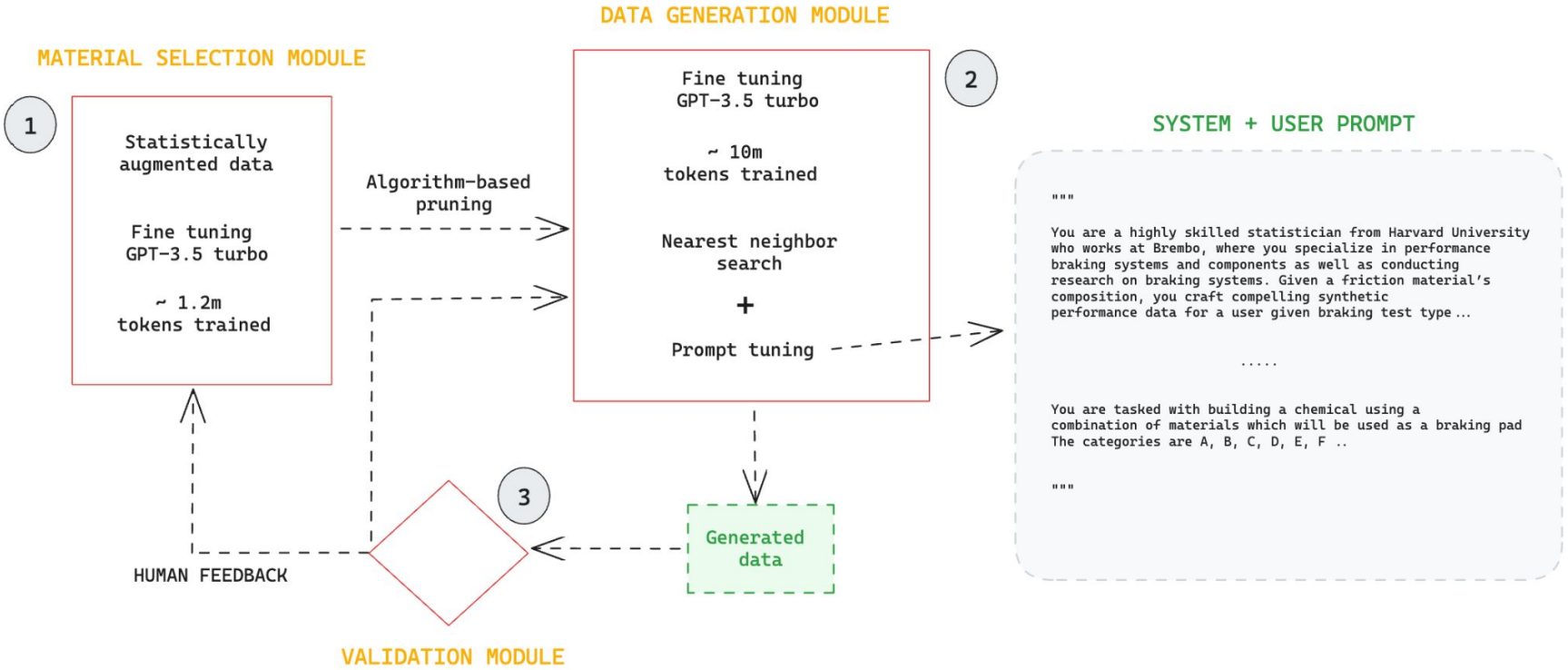
Challenging!

Our system in a nutshell

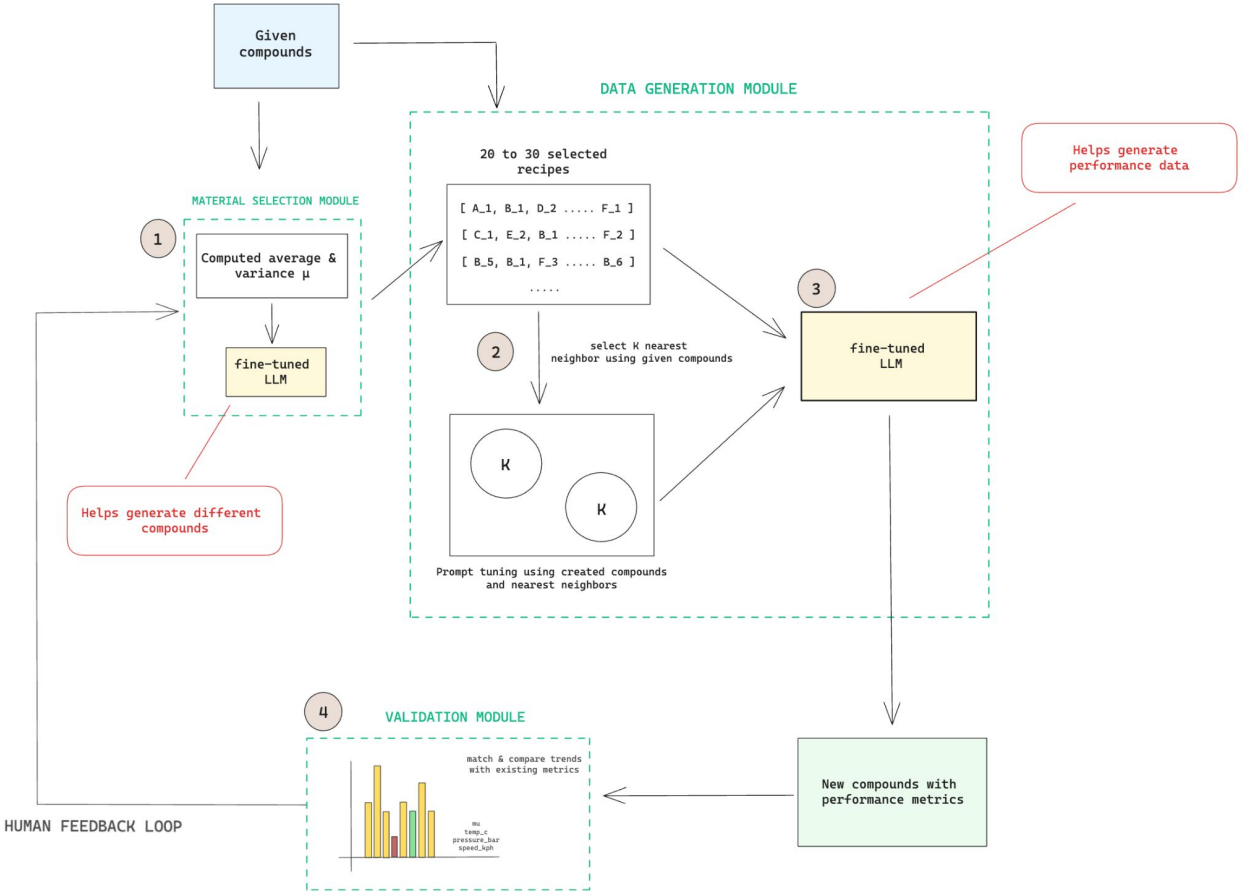


Simple Much?
Let’s Zoom In

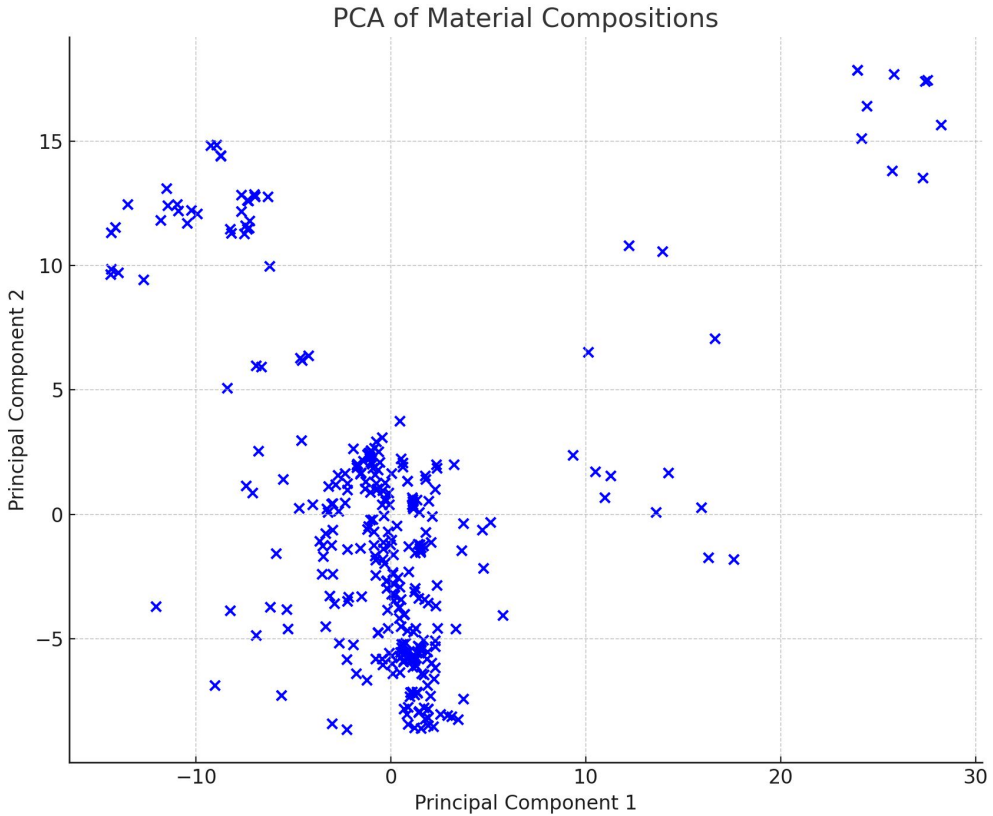
Techniques used



Diving deeper into the system

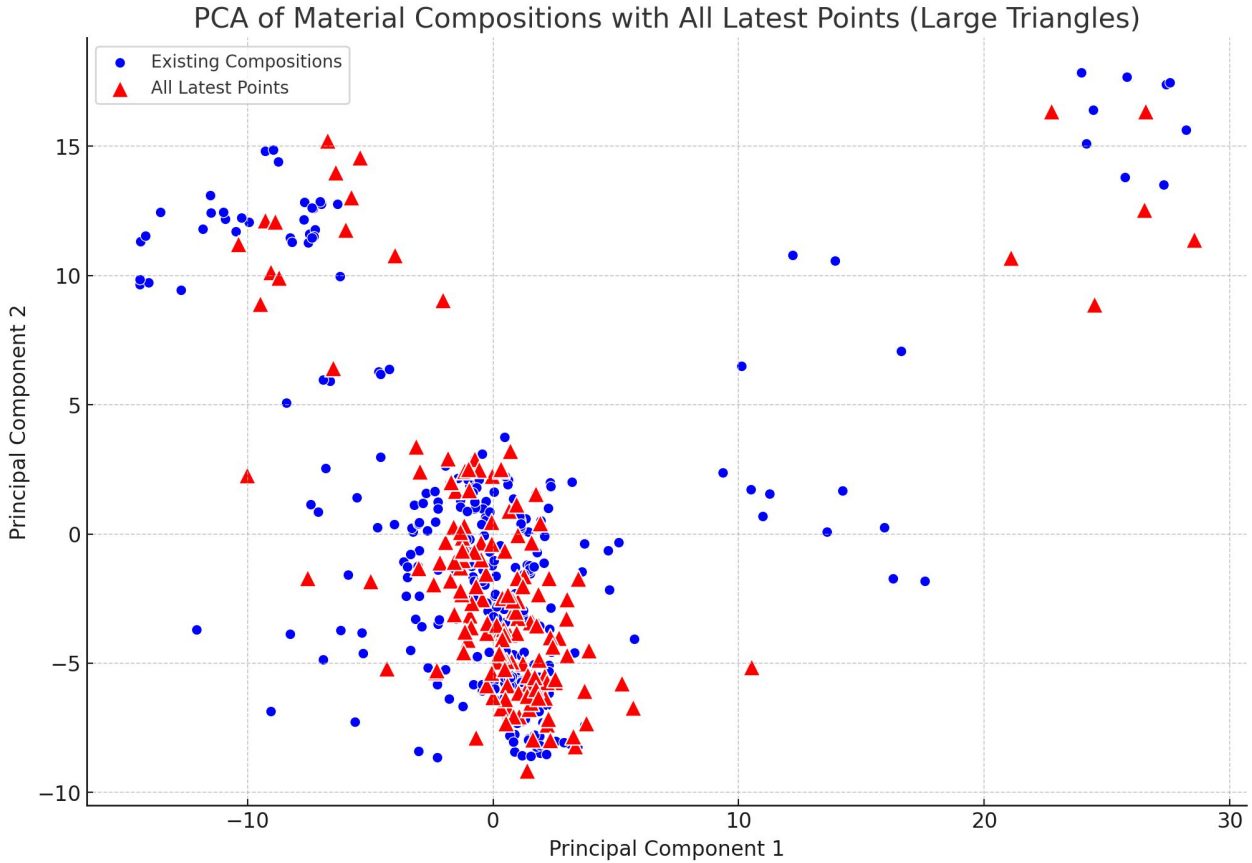


Initial data analysis



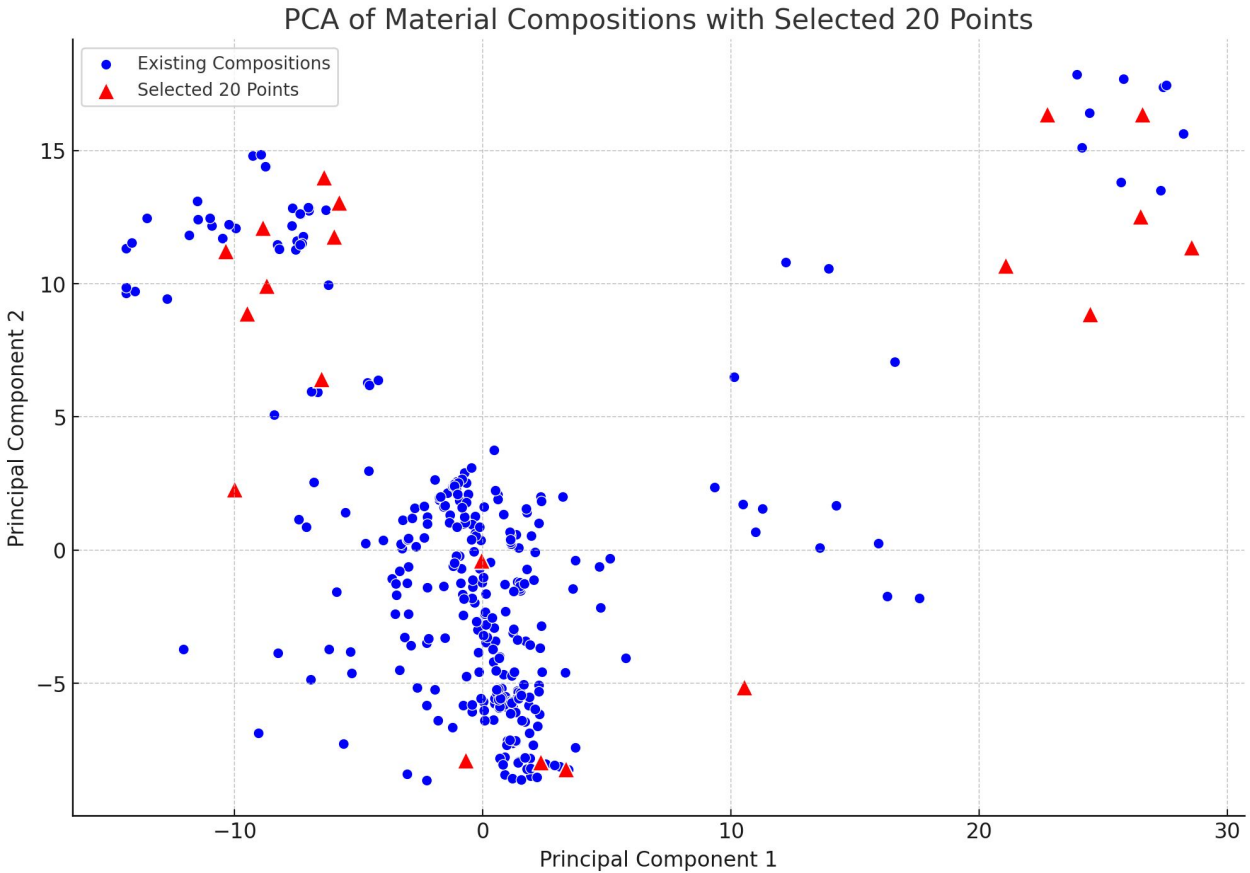
This is a visual representation of the existing material compositions in the dataset.

Creating new compounds



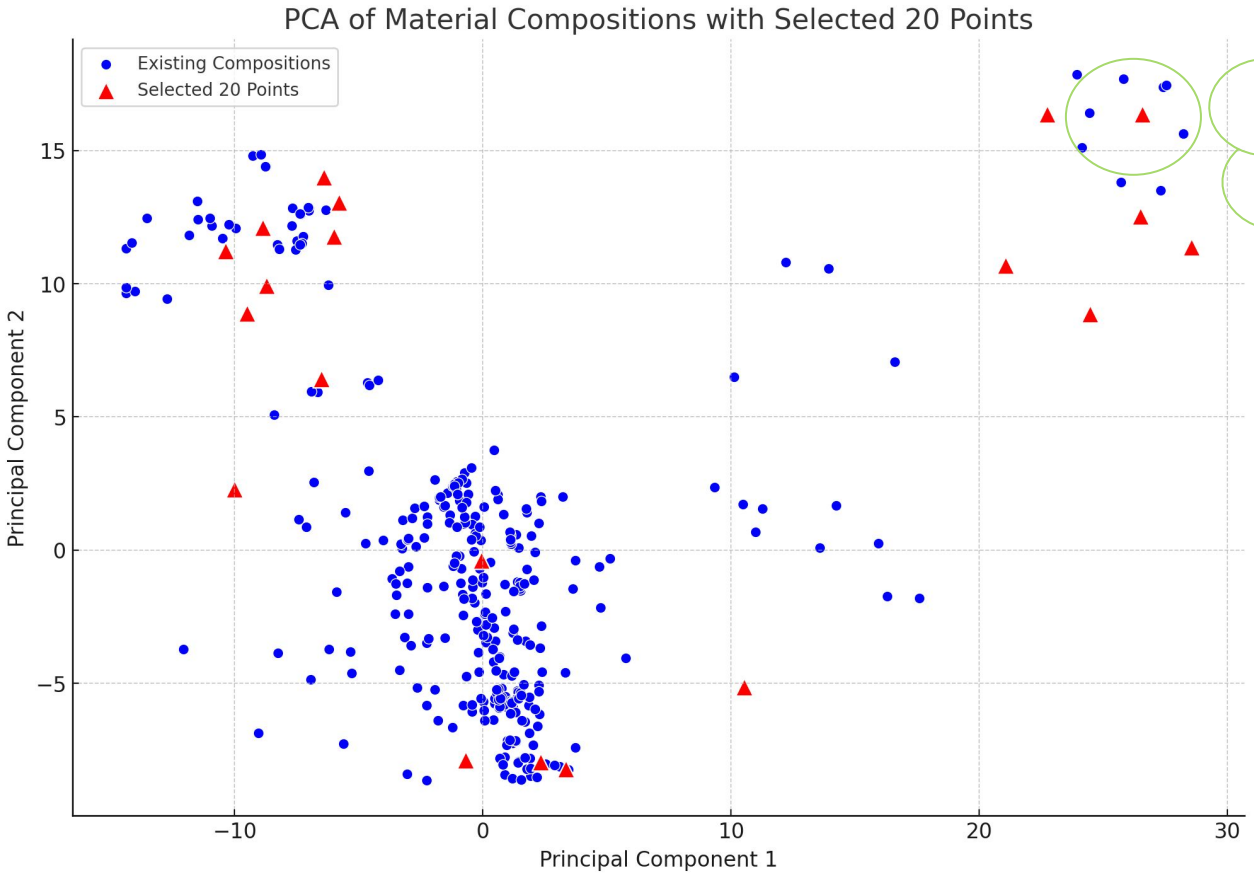
Red points are new materials generated by our AI model

Pruning for variability in compounds



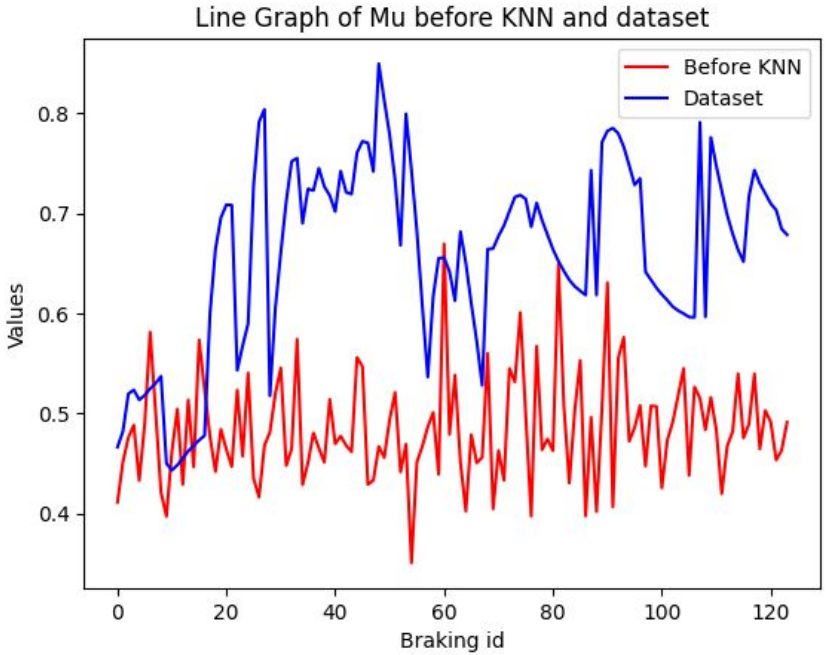
A subset of the red points were selected to optimise for variance

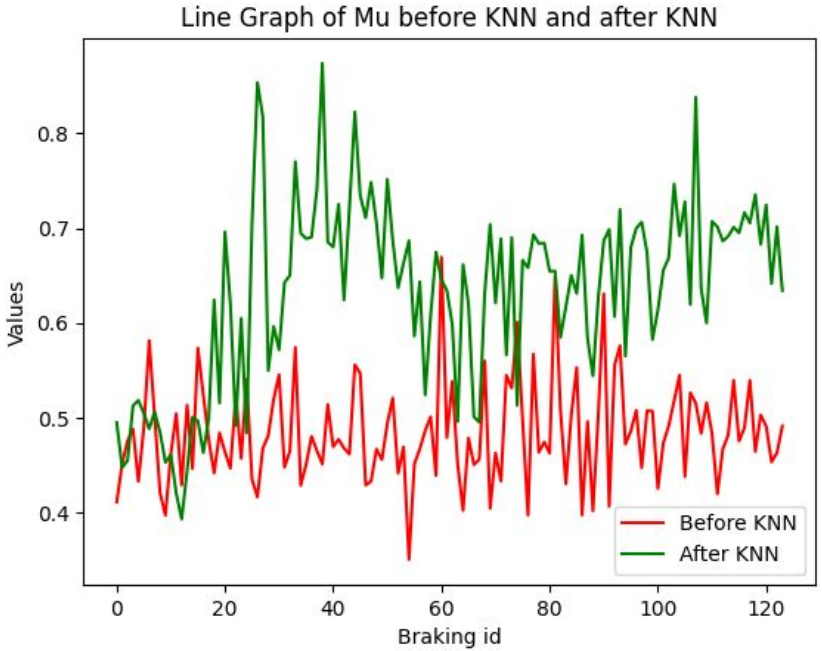
How did we improve performance?



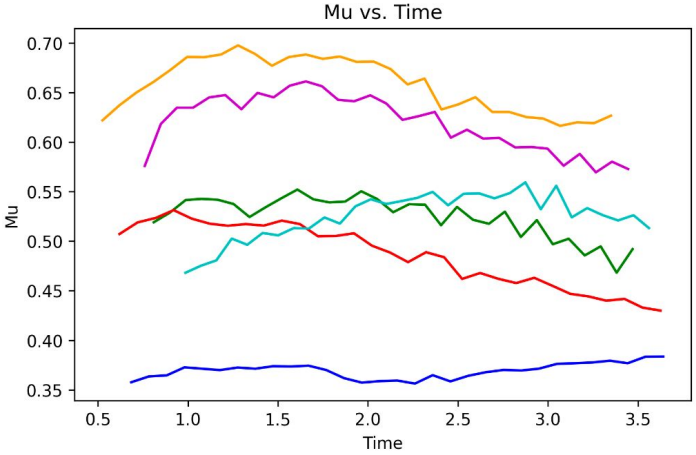
To get time series data for the red compound, we prompt our model the data of the blue ones

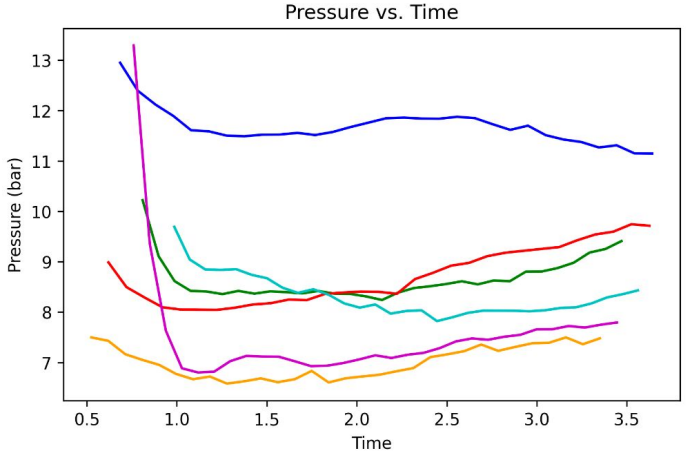
How did we improve performance?

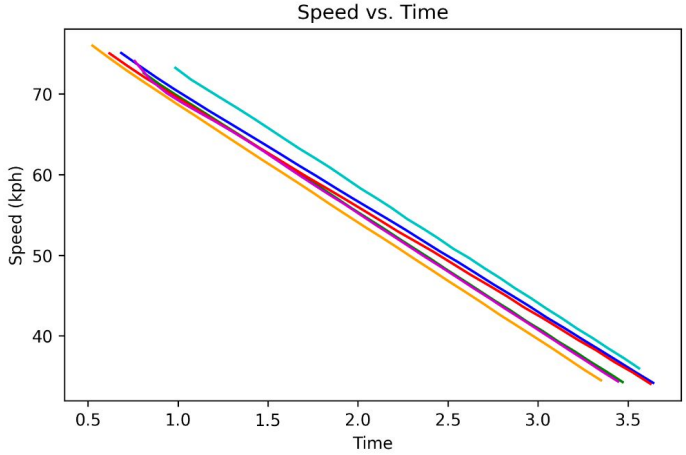


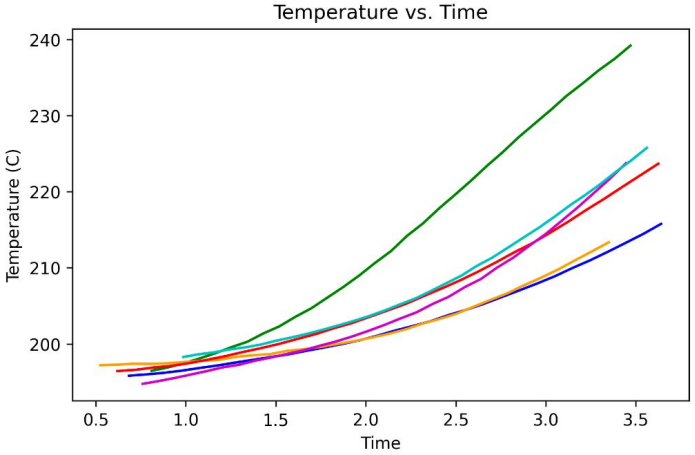


Results: Guess what's real and what's generated.









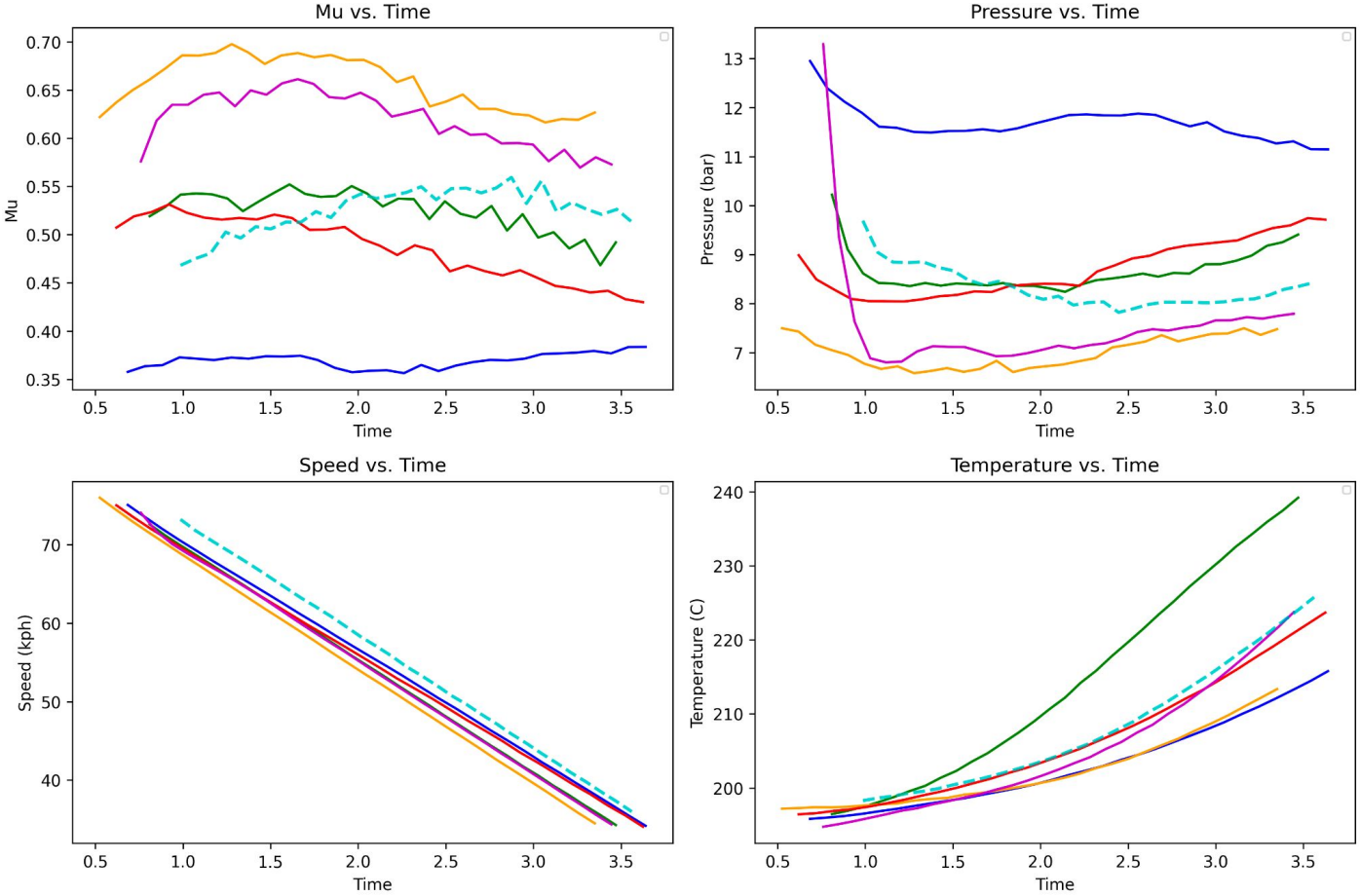
5 of these graphs are from existing data.

One has been generated by our model.

Can you guess which one?

Look closely!

Were you able to guess correctly?



The light blue one (dotted) is ours!

Future Work

Fine tune our models using more data (Increases physical relevance)

Better selection of input materials for fine tuning (Increases physical relevance)

RLHF – Automatically incorporate feedback from validator through prompt tuning

Inject some non-determinism into material generator to get varied set of materials (Increases variance)

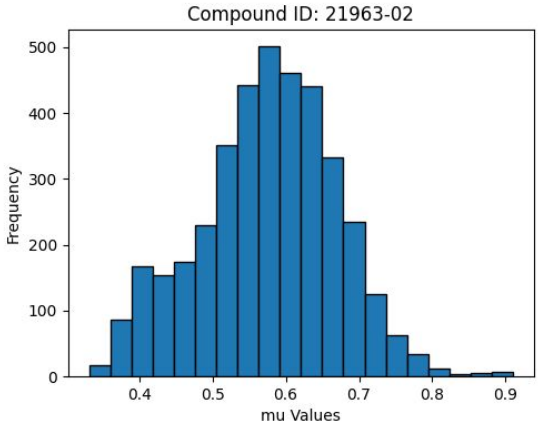
Cheers!

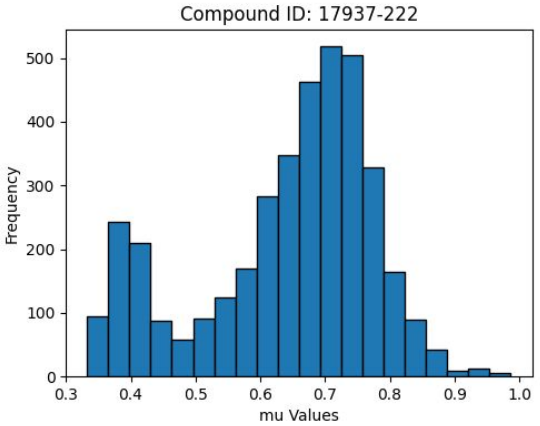
Appendix

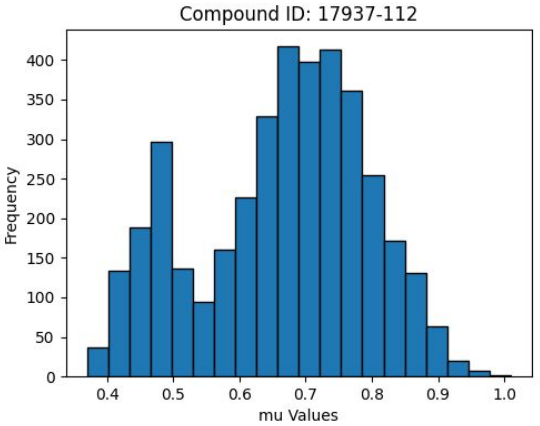
Why use GenAI?

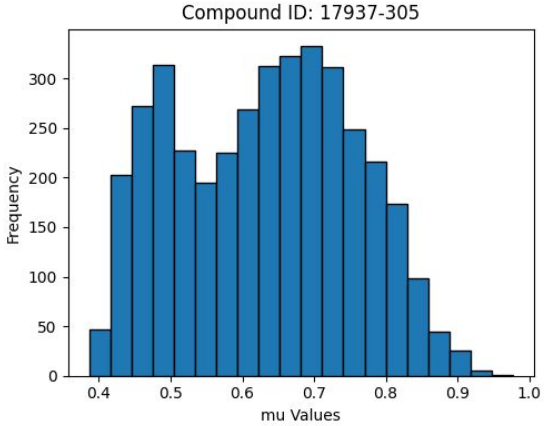
- GenAI is known to perform better on time series forecasting than traditional ML techniques (research paper).
- After analysing the given data, we also incorporated traditional ML techniques (like KNN, K means clustering) into our data generation pipelines to achieve great results.

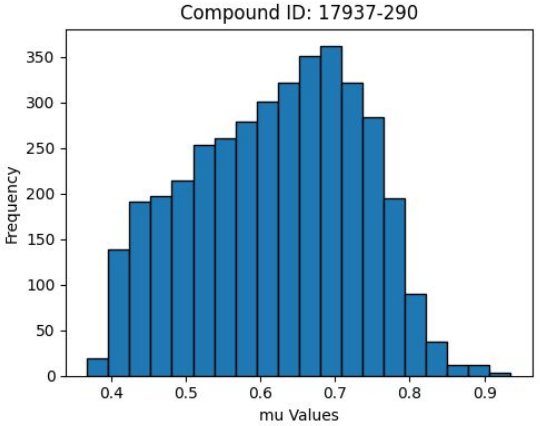
Initial data analysis











Plotting out the range of μ values across different compounds