Essay propose a model monitoring pipeline and describe how you would track model drift in 500 words.

Assume that the model has been implemented as a pytorch model. I would track model drift using ipython notebook(s) to evaluate, get feedback, visualise and process data.

Data collection

- Script to fetch evaluation data from database and put into csv if the data is not too much
- Or if the data is too large to fit into memory, write a dataloader class in pytorch to load from the dataset path
- Data should have the ground truth and model prediction.
- We should have a fixed collection of data to be used and loaded to evaluate against the current batch of data to detect if the underlying data inputs have changed to the extent that it is different from what the model was initially trained on. This data could be obtained from previous batches of trackings that we agree upon on how each sample should be.

Computation of evaluation metrics

- After loading the data, perform metric calculations on the metrics we want to track and evaluate the model against.
- Metrics like KL divergence can be computed to evaluate whether the new batch of data is different from the fixed collection of data
- We can keep track of how previous batches' data' metrics e.g. KL divergence were, to visualise the change in the metric over the different batches. If the metric have been consistently bad or sub-optimal for many batches, it could mean that there is data drift.

Visualisation

- Visualise the metrics computed over the different batches; if the metrics are consistently below a threshold, then investigate why, looking at error analysis.

Error analysis - identifying the errors and diagnosing the causes

- If the metrics are not what we expect them to be e.g. below a threshold, look at the samples where the error between the model prediction and ground truth is high.
- For these samples where there is error, decide whether this sample is anomaly (and whether to ignore or not), or if not anomaly, reason why the error is high.
- Document the error analysis findings for future reference in error analysis.