

Essay Question

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

1. Data Ingestion

The first part of the pipeline is to ingest the data and transform it into the appropriate format for our model. We can use AWS Glue for the data ingestion and transformation task and store the data in S3. We can also use AWS Glue to create a Data Catalog that will help us keep track of the data sources and the transformations applied to the data.

Reference: <https://docs.aws.amazon.com/glue/latest/dg/what-is-glue.html>

2. Data Quality Monitoring

Using Deequ, we can create a baseline for the initial input dataset that we provide. This baseline provides us with the baseline schema constraints and statistics for each feature so that we can schedule regular checks to ensure that the data quality is maintained.

Reference:

<https://docs.aws.amazon.com/sagemaker/latest/dg/model-monitor-data-quality.html>

3. Model Monitoring

For monitoring the performance of the model, we can use Amazon SageMaker Model Monitor to detect any drift in the model's predictions. A monitoring schedule can be set up to monitor the model's performance regularly and compare the model's predictions with the actual values to detect any drift. This is done by storing the model's real-time or batch inference in an S3 bucket and comparing it with the actual values. Model monitor compares the model's predictions with the actual values using a suitable metric for our ML problem type, e.g. mean square error for regression problems. These comparisons allow us to detect drift in the model's performance and can also be integrated with Amazon CloudWatch to get alerts when significant drift is detected.

Reference:

<https://docs.aws.amazon.com/sagemaker/latest/dg/model-monitor-model-quality.html>

4. Model Retraining

Using Amazon CloudWatch, when the performance metrics exceed our model-specific threshold, we can trigger a retraining job through an EventBridge rule to update the model with the latest data. Using Amazon SageMaker's automatic model tuning feature to automatically tune the hyperparameters of the model, we can find the best hyperparameters for the model by running multiple training jobs with different hyperparameters and selecting the best performing model. This will help us improve the model's performance and reduce the drift in the model's predictions.

Reference:

<https://docs.aws.amazon.com/sagemaker/latest/dg/automatic-model-tuning.html>