Project Design Phase-I
Solution Architecture

Date	14 October 2022
Team ID	PNT2022TMID33389
Project Name	Solution Architecture
Maximum Marks	4 Marks

Solution Architecture:

The data drift monitoring task continuously profiles the input data, compares it with baseline, and the results are captured in cloud watch. This tasks runs on its own computation resources using Deequ, which checks that the monitoring job does not slow down your ML inference flow and scales with the data. The frequency of running this task can be adjusted to control cost, which can depend on how rapidly you anticipate that the data may change.

The model quality monitoring task computes model performance metrics from actuals and predicted values. The origin of these data points depends on the use case. Demand forecasting use cases naturally capture actuals that can be used to validate past predictions. Other use cases can require extra steps to acquire ground-truth data.

Sage Maker Clarify provides greater visibility into your training data and models, helping identify and limit bias and explain predictions. For example, the trained models may consider some features more strongly than others when generating predictions. Compare the feature importance and bias between model-provided versions for a better understanding of the changes.

As companies continue to use data analytics and ML to inform daily activity, data drift may become a more common occurrence. Recognizing that drift can have a direct impact on models and production-ready applications, it is important to architect to identify potential data drift and avoid downgrading the models and negatively impacting results. Failure to capture changes in data can result in loss of process confidence, downgraded model accuracy, or a bottom-line impact to the business.

Example - Solution Architecture Diagram: