**DATA/MSML 650**

**Final Project Proposal**

**Title:** E-Commerce Predictive Analysis using AWS

**Team Members:** Tausif Khan, Varun Jain, Hsi-Chun Wang, Sidhant Chanana, Allen Mathews

**Description**

E-commerce predictive analytics uses data analysis and machine learning algorithms to forecast customer behavior, trends, and outcomes in online retail settings. By analyzing past data and patterns, businesses can make informed decisions regarding inventory management, pricing strategies, and personalized marketing campaigns to optimize sales and enhance customer satisfaction.

**Method**

To build this project, we begin by grasping ETL on Big Data and the concepts of staging and Data Lake. We set up IAM Roles and Policies, and then analyze the dataset. Then we configure AWS CLI and comprehend Data Streams, Amazon Kinesis, and Apache Flink. We create a Kinesis Data Analytics Application and utilize Glue to define the Partition Key. Lambda functions are created for DynamoDB and SNS integration. We integrate Lambda with Kinesis with DynamoDB Data Modeling. Finally, we implement QuickSight Dashboards for comprehensive data visualization and analysis.

**Project Breakdown**

* Set up and configure the architecture including IAM roles/policies, VPC configurations and manage access control
* Setting up the ETL pipeline into the data lake for analysis
* Implement backend logic for serverless computation and data interaction between services
* Creating a Kinesis data analysis application
* Creating Lambda functions
* Integration of Lambda Functions with Kinesis
* Analyze and visualize customer trends
* Performing ETL for Parquet format using Glue DataBrew and Spark
* Creating QuickSight Dashboards