Assignment 1 (AI-611 ADMEP-EL1) RollNo\_25201317

# Q1) You work for a retail company that wants to improve its marketing strategies by targeting specific customer segments.

In order to effectively target customer segments, the company must leverage advanced data management, engineering, and preprocessing techniques. The steps are as follows:

1. Data Collection and Integration

* Gather customer data from multiple sources:
  + Transactional data (purchase history, frequency, spending amount).
  + Demographic data (age, gender, location, income group).
  + Behavioural data (website/app interactions, product views, abandoned carts).
  + External data (social media, loyalty programs, seasonal trends).
* Use ETL pipelines (Extract, Transform, Load) to consolidate these sources into a centralized data warehouse or data lake.

2. Data Cleaning and Preprocessing

* Handle missing values (e.g., imputation or removal).
* Standardize and normalize numerical attributes such as income, purchase frequency, etc.
* Encode categorical variables like gender, region, and customer type.
* Remove outliers that could skew marketing insights.
* Ensure data quality management through validation checks and duplicate removal.

3. Feature Engineering

* Create new features to capture customer behavior more effectively:
  + RFM Features (Recency, Frequency, Monetary value).
  + Customer Lifetime Value (CLV).
  + Seasonal or time-based purchase patterns.
* Aggregate features at the customer level (e.g., average spend per month).
* Use dimensionality reduction techniques (PCA, t-SNE) if data is very high-dimensional.

4. Data Management for Scalability

* Store structured data in relational databases (SQL).
* Use NoSQL or big data platforms (Hadoop, Spark) for large-scale unstructured/semi-structured data.
* Apply data governance practices for privacy, compliance (e.g., GDPR), and security.

5. Segmentation Techniques

* Apply unsupervised learning methods such as:
  + Clustering (K-Means, DBSCAN, Hierarchical clustering).
  + Customer segmentation based on RFM scoring.
* Create customer personas representing each segment (e.g., “high-value loyal customers”, “discount-seekers”, “seasonal buyers”).

6. Model Preparation and Preprocessing for AI

* Split data into training, validation, and test sets.
* Balance datasets if certain segments are underrepresented.
* Use data augmentation for sparse groups (e.g., synthetic samples via SMOTE).
* Apply scaling techniques (standardization, min-max scaling) to prepare for ML models.

7. Actionable Marketing Strategy

* Personalized recommendations using recommendation engines (collaborative filtering, deep learning).
* Targeted promotions for high-value clusters.
* Churn prediction models to retain at-risk customers.
* Campaign performance monitoring via dashboards and KPIs.

8. Continuous Data Management and Monitoring

* Automate data pipelines for real-time updates.
* Monitor data drift and customer behaviour changes.
* Continuously retrain AI models with fresh data.

# Conclusion

By applying advanced data management, engineering, and preprocessing, the retail company can transform raw customer data into meaningful insights. This enables the creation of data-driven customer segments, which improve marketing efficiency, personalization, and customer satisfaction.