**Summary Report**

**Introduction**

**Objective**

This project aims to predict user churn for an e-commerce platform, utilizing event data such as views, carts, and purchases. Churn is the absence of a purchase or site visit within the past 30 days.

**Key Findings**

- We successfully defined churn using a recency-based threshold.
- Through feature engineering, we incorporated recency, frequency, monetary metrics, and session-based features.
- A Random Forest model was used for churn prediction, and performance metrics such as precision, recall, and F1 score were evaluated.

**Model Performance**

- The Random Forest model achieved a Precision of X%, Recall of Y%, and an F1 Score of Z%.

**Methodology**

**Data Preprocessing**

- Loaded the raw event data and ensured the proper formatting of event_time.
- Created user-level features based on RFM metrics (Recency, Frequency, and Monetary).
- Defined churn as a user not purchasing in the past 30 days.

**Feature Engineering**

- Developed additional features including session count, average session duration, and behavior patterns such as view-to-cart and cart-to-purchase ratios.
- Engineered a final set of features that were used in the model training.

**Model Building**

- Built a Random Forest model to predict churn.
- Hyperparameter tuning was performed using Grid SearchCV to optimize model performance.
- Model performance was evaluated using metrics such as Precision, Recall, and F1 Score.

**Feature Importance**

- The most important features in predicting churn were recency, frequency, and monetary, with additional insights from behavioral features like view-to-cart ratios.

**Business Insights**

**Key Drivers of ChurnI**

- Users who frequently view products but never add them to the cart are at a higher risk of churn.

- High recency values (i.e., no recent activity) strongly correlate with churn.

**Business Recommendations**

- **Personalized Offers**: For users who have a low view-to-purchase ratio, consider offering discounts or personalized promotions to incentivize purchases.
- **Engagement Campaigns**: Send re-engagement emails to users with high recency values (i.e., have not interacted with the platform in over 30 days).
- **Product/Brand Focus**: Products and categories with lower conversion rates could benefit from targeted marketing or product adjustments to reduce churn.

I LEARNED HOW TO KDE KERNEL DENSITY PLOT AND RANDOM FOREST  AND HOW TO TRAIN AND TEST THE DATA RFF (RECENCY, FREQUENCY, MONETARY)  AND BEHAVIORAL AND EDA