Executive Summary

This project focuses on building a personalized hybrid recommendation system for a cannabis e-commerce platform. I worked with simulated user browsing data that involves their interaction with the platform's products (clicks, saves, carts, and purchases) and detailed product attributes like strain, potency, price, and category to understand browsing behavior and engagement patterns.

To generate recommendations, I combined graph-based modeling and NLP techniques:

- A Node2Vec graph embedding approach captures co-engagement patterns between users and products, making it easy to recommend similar items based on behavior.

- A TF-IDF text embedding layer analyzes product names, descriptions, and tags for content-based recommendations.

I blended both methods into a **hybrid ranking system** and also built classification models (Logistic Regression, Random Forest, XGBoost) to predict purchase intent and using **SMOTE** to handle class imbalance. This setup not only improved recommendation quality but also highlighted features most predictive of conversions.

Overall, this project shows how combining behavioral data and product text can create more relevant recommendations for cannabis marketplaces, while also giving insights into pricing, product positioning, and engagement trends. Future steps could include conditional probabilities on purchasing behavior to reflect real world dynamics, A/B testing with real-time recommendations and experimenting with **graph neural networks** for deeper personalization.