Next-Click Prediction System for Fashion Campus

Enhancing Product Recommendations, Through Sequential Modeling

Problem Statement & Objectives

Problem:

- Users frequently leave without completing purchases due to lack of personalized recommendations.

Objectives:

- Understand user navigation behavior.
- Predict the next-clicked product using user sessions.
- Compare Markov chains and GRU (Gated Recurrent Unit) in recommendation accuracy.
- Visualize user journey and product relationships.

Data Used & Feature Engineering

Datasets:

- Clickstream: 12.8M records (100K sampled records for efficiency)
- Customer: 100K
- Product Metadata: 44K
- Transactions: 850K

Key Features Extracted:

- Product sequences by session.
- Time-based features: hour, weekday, time since last click.
- Click counts, session metrics.
- Embedding-based product similarities.

Methodology & System Architecture

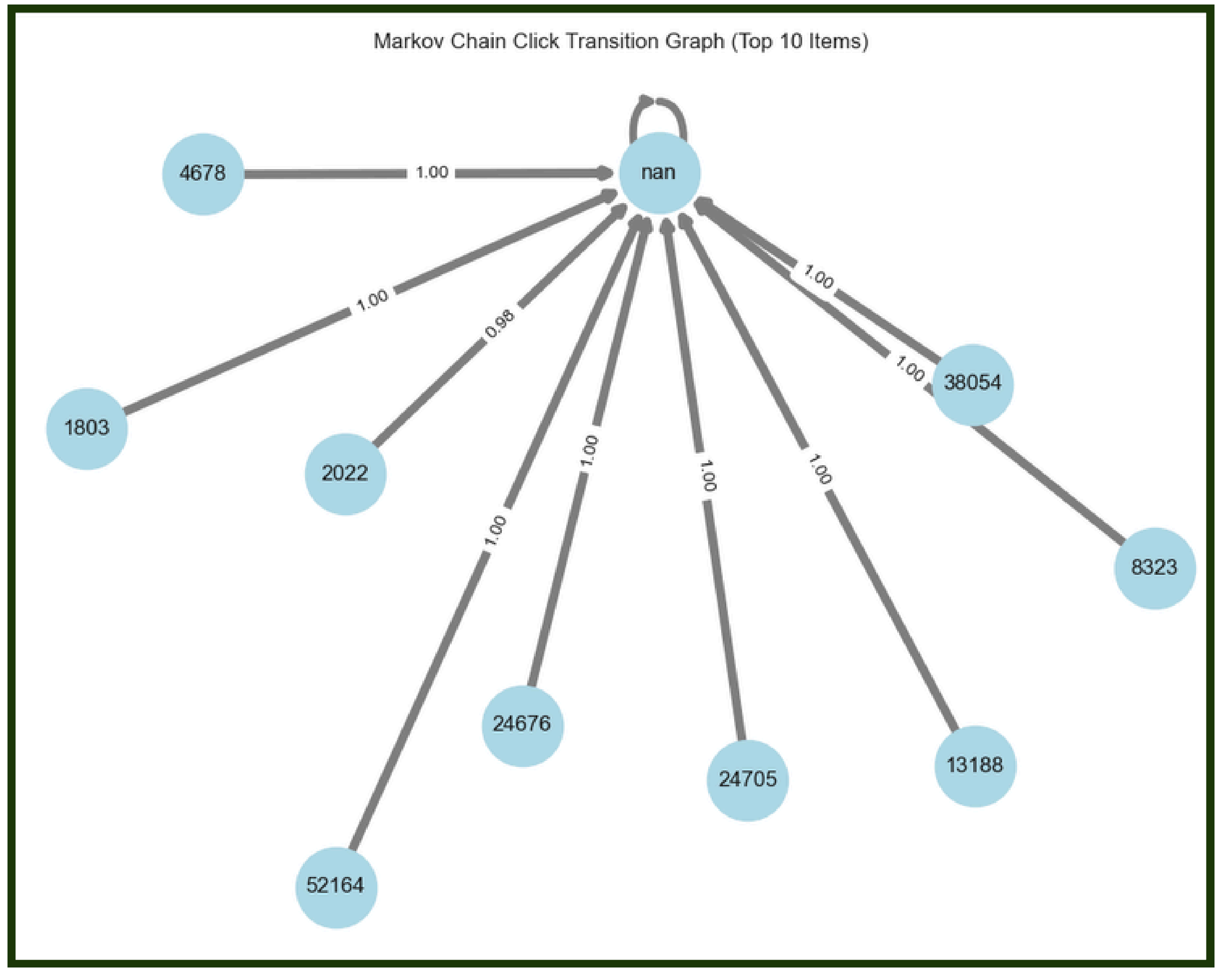
Pipelines:

- Preprocess → Sequence Creation
- Word2Vec: Learn product embeddings
- t-SNE: Visualize product relationships
- Markov Chain: Simple next-click modeling
- GRU (Gated Recurrent Unit): Sequential deep learning
- Simple Input-Output Setup: User interaction for next-click prediction

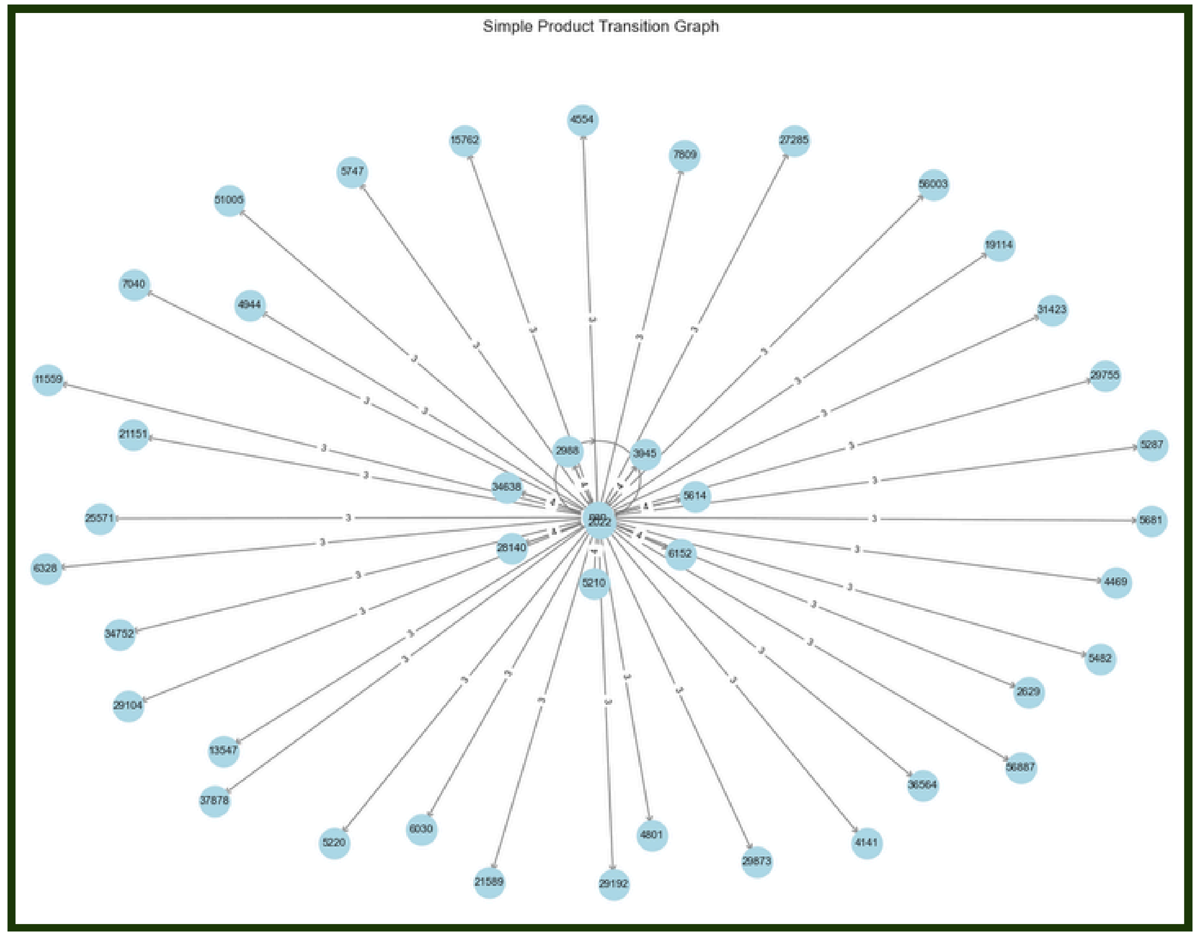
Tools & Libraries:

- Python, Pandas, Gensim, TensorFlow, Keras, Seaborn, NetworkX

Model Evaluation & Results



Model	Top-1 Accuracy	Top-3 Accuracy
Markov Chain	88.68%	90.07%
GRU	87.16%	88.51%



User interaction for next-click prediction

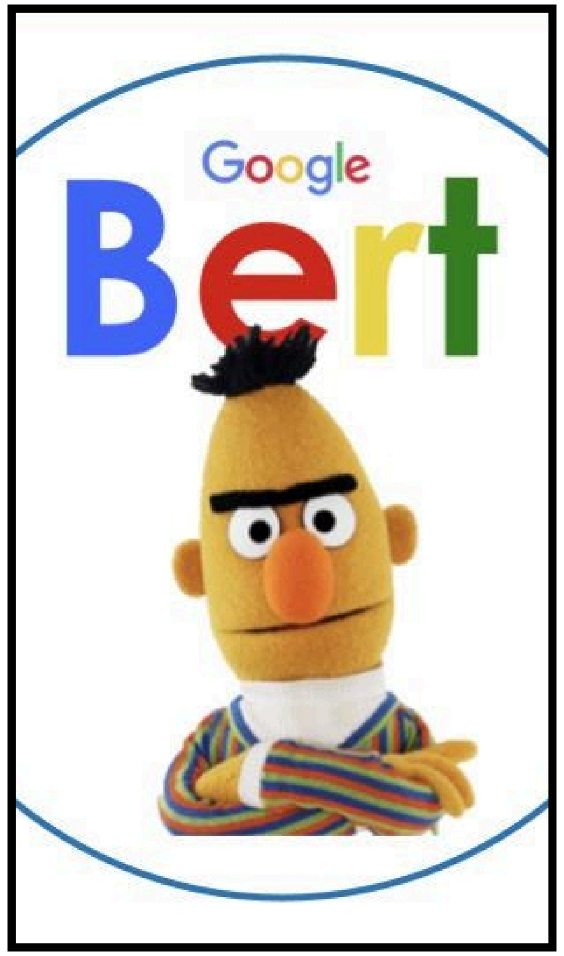
```
    print("\n[!] No similar products found.")

 ===  Next-Click Prediction ===

 Available Product IDs (sample): ['nan', '2022', '2988', '5210', '5614', '6152', '5217', '3945', '28140', '34638'] ...
Enter your last clicked Product ID: ↑↓ for history. Search history with c-↑/c-↓

]:
```

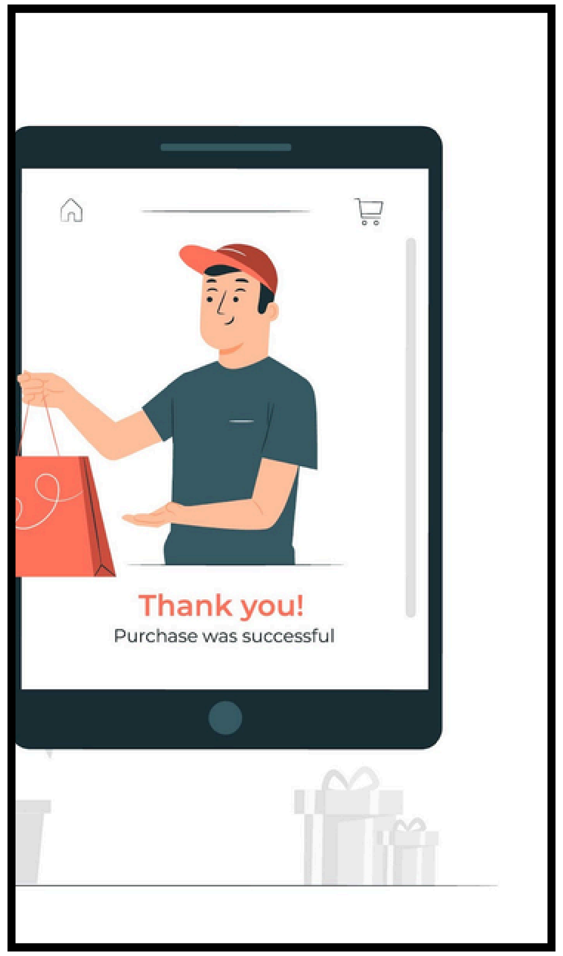
Conclusion:

- Built a scalable Next-Click Prediction System for Fashion Campus.
- Combined classical (Markov Chain) and deep learning (GRU) methods.
- Used Word2Vec to capture product semantics and user behavior.
- GRU performed well on complex sessions; Markov excelled on simpler ones.
- Implemented a simple CLI for real-time user interaction.
- Showcased the power of embeddings and sequence modeling in e-commerce.

Future Work

- Integrate user demographics & real-time feedback
- Use Transformers (like BERT4Rec) for deeper context
- Deploy on cloud (AWS/GCP) for real-time scalability
- Personalized recommendations via reinforcement learning





References

- Hwangbo H. (2022). Sequence-aware recommenders for fashion e-commerce. ElectronicCommerce Research, 22(3), 587–605. Retrieved from https://doi.org/10.1007/s10660-022-09627-
- Jiang, J., Wang, Z., Fang, F., & Zhao, B. (2020). TPG-DNN: A method for user intent prediction based on total probability formula and GRU loss with multi-task learning. Retrieved from https://arxiv.org/abs/2008.02122
- Shobiyya, L. H. (n.d.). Retrieved from https://www.kaggle.com/datasets/latifahhukma/fashion-campus/data
- Zhang, H.,& Chen, Y. (2022). A deep Markov model for clickstream analytics in onlineshopping. In Proceedings of the 2022 Web Conference (pp. 3243–3253).ACM

Fashion Campus

Kaggle is the world’s largest data science community with powerful tools and resources to help you achieve your data science goals.

k kaggle.com

Thank You