neo4j

HELPING THE WORLD MAKE SENSE OF DATA

# GRAPH DATA SCIENCE USE CASES: RECOMMENDATIONS

*Jaimie Chung*

Product Manager, Graph Data Science

*Dr. Phani Dathar*

Data Science Solution Architect

# Problem

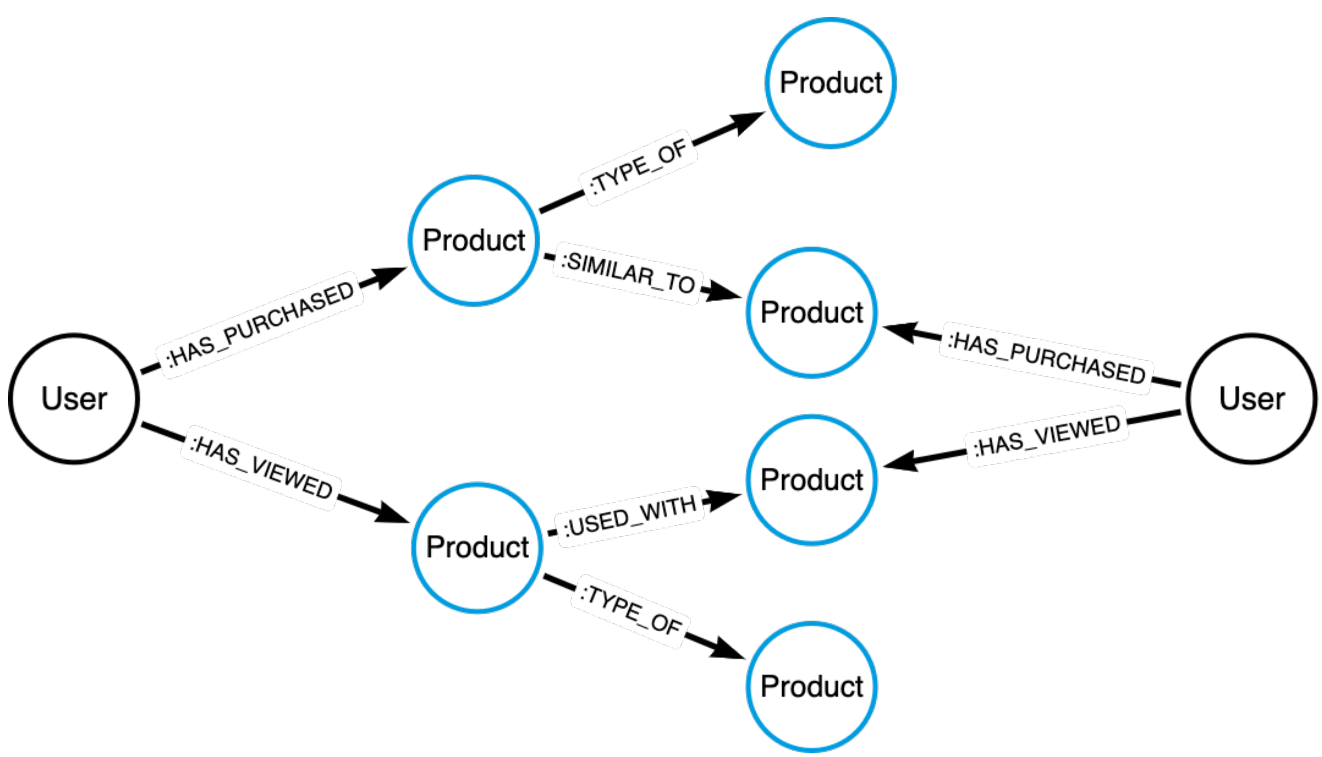
Recommender systems are ubiquitous and essential for increasing customer satisfaction and accelerating business growth. Recommendations powered by machine learning have come a long way since collaborative filtering, the likes of which Netflix utilized in the 2000s. Now, companies across industries filter based on content, demographics, products, and more. Moving beyond generic recommendations based on macro level insights, personalized recommendations are becoming especially relevant to ecommerce and retail businesses with large customer bases and product portfolios in the millions.

Companies must be concerned not only about the quality of their recommendations (are we suggesting the right thing?) but also about how quickly they can derive relevant recommendations and serve them to their users (are we suggesting the most up-to-date thing?). No one likes seeing an advertisement for a pair of shoes they bought three weeks ago following them around on the web.

Given the intricacies of recommender systems and the risks of serving bad recommendations, it is necessary to leverage advanced analytics to ensure that customers receive the most relevant and timely recommendations. With Neo4j's Graph Data Science (GDS) framework, you can use the connections in your data to power recommendation engines that fuel customer satisfaction and business growth, gaining otherwise unattainable insights from the relationships that exist in the data you already have.

# Data Model

One common use of recommendations is product recommendations based on user behavior and product ontologies. Users could take multiple actions on a website: browse through navigation for product categories, view products of interest, and add products to their cart.




*A graph data model for recommendations*

This data can be coupled with product information based on ontologies: products with similar uses, of the same type, or complementary products. While this example applies to ecommerce, the same could be said for a media publisher that wants to recommend content or a SaaS company that wants to recommend additional product features.

In the data model shown, a business could use query-based logic, use graph algorithms, or implement supervised machine learning techniques to serve a promising recommendation to the user.

Graph-based approaches enable you to power recommendations that take into account not only domain-specific knowledge of your product offering but also user-driven actions and behavior, leading to more accurate recommendations with higher conversion rates.

# Solution

Neo4j's Graph Data Science framework offers a variety of analytical approaches to make relevant recommendations based on the relationships within your data, ranging from localized query patterns to machine learning–based insights.

## Queries

Cypher is a powerful, intuitive, graph-optimized query language. For example, a content provider could conclude that users who are interested in content about business are also likely to be interested in economics. A simple recommender system could therefore include a Cypher query that serves articles from the economics section over content from less relevant sections.

## Graph Algorithms

Neo4j Graph Data Science offers out-of-the-box graph algorithms for similarity and community detection. These algorithms are useful on larger datasets where it's hard to know exactly what you're looking for. These techniques often identify candidates for recommendations that are not obvious with queries alone. Some examples of useful algorithms include:

- **Community detection algorithms**, like Louvain, for customer segmentation. Rather than profiling users based on demographic data, identify relevant peer groups based on buying behaviors and actions. These same algorithms can identify commonly co-purchased items across multiple transactions or users, improving standard market basket techniques.
- **Centrality algorithms**, like PageRank, use network structure to identify important items, so you can understand which product is most relevant – or which items to promote to increase total sales
- **Similarity algorithms**, like Node Similarity and KNN, use graph structure to find similar items – exploiting the network structure of actions across a dataset to make recommendations.

Together, these techniques yield powerful recommendations that consider both a product's characteristics and user buying patterns.

## Supervised Machine Learning

Machine learning on a graph encompasses two aspects: how you represent your data, and the predictions you want to make with it.

To characterize your data, **graph embeddings** learn to encode *your* graph in a simple, machine-readable format – specifically tuned to solve your problem. They simplify all the rich information in your full graph into a simple descriptor for each item or customer. That description then feeds into downstream machine learning tasks.

Techniques like **link prediction** and **node classification** can use graph embeddings to predict changes to the structure of your graph. You can train a model to predict which item a customer will buy (future relationships forming), or to identify which customers might churn (future labels).

# Results

Graphs enable recommendation engines to leverage the most powerful predictor of future behavior: relationships. Better recommendations mean a more personalized experience, higher conversion rates, and even the ability to forecast trends to manage inventory.

While graph-based recommendations created using manual Cypher queries may be a huge step forward in user experience for many, large volumes of data and uncertainty around user behavior mean they may not be the final solution. Graph algorithms and graph-based machine learning let you identify the important patterns in big data, and predict unusual or unexpected future behaviors. Accurate, timely recommendations are crucial in driving conversions and, ultimately, business growth.

## Customer Spotlight: Top Retailer

A top retailer with over $100 billion in yearly revenue relies on Neo4j to power their product recommendations. The retailer's portfolio has millions of products that fall into complex, overlapping product taxonomies, and their customers are increasingly expecting near natural language search capabilities. It was essential to go beyond a traditional recommendation engine.

The main challenges were the volume and variety of product and user data – which were not connected in the relational world – and the limited ability to improve search relevance and product ranking based on sparse and heterogeneous data. Their previous solution failed when there was not enough data about the user or when new products were introduced into the portfolio.

To solve this, the retailer uses Neo4j's Graph Data Science framework to derive similarity relationships on a graph that connects two years of customer interactions with product taxonomies, products, and search queries. Identifying users with similar behavior solves for users with no historical data, and identifying similar products increases coverage for new products. They also use node embeddings that account for complex relationships between users, sessions, products, and taxonomies to make recommendations for unusual and ambiguous search terms. Embeddings are critical to improving overall search relevance, ranking, and quality, even compared to state-of-the-art recommendation systems.

With Neo4j GDS, they are able to incorporate the connections in their big data that are crucial to understanding their customers' needs, and provide novel and relevant recommendations where even the best recommendation systems fall short. As a result, customers receive timely and helpful recommendations, which leads to more purchases and return visits.

## Conclusion

Neo4j's Graph Data Science framework enables you to make sense of your data –- at scale.

No matter whether your data challenge is streamlining your recommendations strategy, keeping readers onsite, or bolstering your search experience, using Neo4j Graph Data Science for recommendations provides fast and accurate results. Recommendations is just one of the many use cases enabled by graph data science.

Learn more about Neo4j's Graph Data Science framework at [neo4j.com/graph-data-science](neo4j.com/graph-data-science) or get started right away with a [Neo4j GDS Sandbox](Neo4j GDS Sandbox).