**Performance demonstration :**

In this project, we explore different recommendation models to suggest the top three items for each customer on their next visit to the store (for example, a Leroy Merlin store). The first model is a **cluster-based approach**, where we apply feature engineering and dimensionality reduction, then group customers into clusters. We recommend the top three products based on aggregated preferences within each cluster.

The second model is **matrix factorization**, which we use as a **baseline**. It learns hidden patterns between customers and products to make personalized recommendations.

In addition, we test **graph-based models** such as GCN (Graph Convolutional Network), GAT (Graph Attention Network), and a weighted version of GCN. These models leverage the structure of customer-product interactions as a graph to improve recommendation quality.

Finally, we compare the performance of all models using various metrics, including both ranking-based and basket-level evaluation methods.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Recommendation Model | Recall@3 | NDCG@3 | MAP@3 | Lift@3 | Training time (s) | Inference time (ms) |
| Cluster-based | 0.016 | 0.029 | 0.020 | 1.89 | 0.029 | - |
| MF | 0.038 | 0.071 | 0.051 | 4.49 | 0.08 | 1 |
| GAT | 0.049 | 0.088 | 0.062 | 5.79 | 20.9 | 94 |
| GCN | 0.067 | 0.128 | 0.09 | 7.84 | 8.2 | 16 |
| Weighted GCN | 0.084 | 0.151 | 0.110 | 9.818 | 10.3 | 23 |

The graph-based recommendation models show a clear and significant improvement over the Matrix Factorization (MF) baseline across all evaluation metrics. While MF achieves a Recall@3 of 0.038, graph models such as GAT, GCN, and especially Weighted GCN perform substantially better. Weighted GCN, in particular, more than doubles the Recall@3, reaching 0.084—representing a 121% increase over MF. Similar gains are observed in other metrics: NDCG@3 rises from 0.071 with MF to 0.151 with Weighted GCN (+113%), and MAP@3 improves from 0.051 to 0.110 (+115%). These results indicate that graph-based models, by leveraging the structural relationships between customers and products, are more effective at ranking relevant items higher.

The Lift@3 metric, which measures improvement over a random baseline, also increases notably—from 4.49 for MF to 9.82 for Weighted GCN—highlighting the stronger discriminative power of graph models. While this performance boost comes at the cost of higher training and inference times, the trade-off is reasonable. For instance, Weighted GCN has an inference time of 23 milliseconds compared to MF’s 1 millisecond, which remains acceptable in many real-world applications where recommendation quality is critical.

**Overall, the graph-based models, and particularly Weighted GCN, demonstrate a substantial performance gain, justifying their additional computational cost by delivering much more accurate and relevant recommendations**.