Characterizing Intent Using Customer Journey: a Sequential and Graphical Model Approach

Simon Weill
Arnaud Sarfati

What are the different personas in Wayfair customer base?

Identifying and clustering customer behaviors is a first step towards personalizing the customer journey.

Wayfair clickstream data logs every action customers take on the website and has the potential to provide insight into consumers behaviors.

We use K-Means algorithm to cluster the embeddings obtained. We figure out the optimal number of clusters by analysing the evolution of Sum of Squared Errors (SSE) and Silhouette coefficient as the number of clusters increase. We find that the optimal number of clusters is 5.

Unsupervised Clustering

We observe strong discrimination between clusters which links back to the business intuition: Different conversion rates, Different Add-To-Cart rates, Different typical subsequences, Different channel distribution, Different visitor types.

• By improving customer journey for specific personas, post purchase touchpoints can be reduced by 10X, and reduction of customer service cost and increased customer return rate can translate overall into $100M additional revenues
• The insights from the model will be used to do back-testing and AB test different summarized insights from our work.
• On a technical standpoint, model improvement can be considered by integrating new features about customers, catalogue specific details on pages and other graphical features that can be used to improve clustering results.