Unsupervised Customer Segmentation using Knowledge Graphs

# What is Customer Segmentation and how is it useful?

Customers are the key for a business to continue, so grouping them together for marketing and targeted actions can be financially beneficial to the business. One methodology of performing this can be to use the customers’ behaviour to look for similarities between them, what products they purchased at stores, and how they rated the products. Dealing with the sequences of customer reviews and purchases has previously been performed using neural networks that specialise in sequences or recommendations. Pai et al wrote an article on instead using knowledge graphs to obtain groupable embeddings of the customers, which will be analysed here.

# Knowledge Graph?

The above image is a graph that displays the connections Jane, James and Jerry have with each other. A graph is made up of nodes/vertices connected by edges. The above graph is also a knowledge graph, which is a graph where the nodes are entities, and the edges are relationships. Entities and relationships need to be writable as a set of facts. This graph can allow the reader to quickly see the connections that James and Jerry have, such as both being programmers and both having the same favourite colour.

# Embeddings

## Word embeddings

In multiple data scenarios, a lot of potential input objects can be involved, for instance, in natural language processing there are many different words that occur in any language and so for inputs to any machine learning model, the model must be able to deal with these. A common method is to perform one hot encoding, where a one in an input vector represents a word. Whilst it is easy to read for a human, it is sparse, and a vector is made for each word in a sentence with the length of each vector the number of words in the vocabulary. Instead of taking this approach, another model can be produced that learns a low-dimensional vector output, the output can then be used as the input for other model. As the models learn about the words it is trained on the similarities or differences between these words become encoded into the vectors that are output, a commonly found property is that similar words become closer in the vector space so words like king and queen are more likely to be closer together than the words king and grass would.

https://kawine.github.io/blog/nlp/2019/06/21/word-analogies.html

## Knowledge Graph Embeddings

Word embeddings and knowledge graph embeddings are not the same, but they share a similar concept as trying to encode the information of the entities, relationships their attributes and the semantic structure into real valued vectors. Unfortunately, when it comes to real world knowledge graphs training them is a lot harder than Word2Vec and other word embedding models, the embedding models must be trained on a machine and there are no pretrained models as the knowledge graphs are unique.

In this blog post we will be replicating the results of Pai et al in his paper “Unsupervised Customer Segmentation with Knowledge Graph Embeddings”. In this paper Pai et al setup Rate Beers reviews as a graph instead of just tabled sequential data and use the embeddings of the users later to segment the customers from each other.

You can follow along using our repository RateBeer’s Stanford SNAP-RateBeer dataset https://www.ratebeer.com/api.asp and our git repository located at: https://github.com/ambiata/UnsupervisedCustomerClustering. We have also pretrained a TransE Knowledge graph embedding model located here (https://drive.google.com/drive/folders/1q4EMs9Y7OtdHElObRa2Dig-xVBoI1lWJ?usp=sharing).

## TransE

# Turning the Reviews into a Graph

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Review | Customer | Aroma | Palate | Taste | Appearance | OverallScore | Date | Beer | Style |
| 1 | HeyThere | 3/5 | 4/5 | 5/5 | 1/5 | 13/20 | 1/12/22 | 655 | Ale |
| 2 | Coco | 4/5 | 3/5 | 4/5 | 3/5 | 14/20 | 4/12/22 | 599 | Stout |
| 3 | HeyThere | 2/5 | 5/5 | 4/5 | 2/5 | 13/20 | 6/12/22 | 100 | Ale |

Pai et al proposed using a knowledge graph on being able to segment customers purchasing consumables, objects that could be purchased multiple times like milk and eggs. To

Admittedly it is hard to find a product that satisfies this requirement at the same time as being something that has data that we can follow along with but Pai used the reviews that Rate Beer (www.ratebeer.com) for his article and we will also use this data. Each of the reviews that the people make is joined to their previous reviews by a succeeding and a preceding relationship edge between each other.

In this article we will be using PyTorch for our neural networks, the reason for this is that for graph neural networks PyTorch tends to be more often used and so that has resulted in modules being written ready for its use.

# Bibliography

Unsupervised Customer Segmentation with Knowledge Graph ...." <https://www.semanticscholar.org/paper/Unsupervised-Customer-Segmentation-with-Knowledge-Pai-Brennan/745a7286116ea56492904133a1dab05d7baa56d8>.