Thesis Work

MACHINE LEARNING FOR AUTOMATED CATEGORIZATION OF PRODUCT ARTICLES

Background

Caspeco is an Uppsala based company that provides tools and services for salary handling, resource scheduling, financial analysis and budget mainly for the hospitality business, and especially restaurants.

With hundreds of customers, Caspeco is continuously accumulating vast amounts of data. Endless potential lie in analyzing this data to provide business insights, but there are many challenges that need to be addressed. We are currently in the process of categorizing the product articles of the sales transactions we collect every day from our customer’s different cash systems, in order to be able to create benchmarks and do large scale market analysis of trends and buyer behaviour. This is currently being done manually, which is neither scalable nor cost-efficient.

With the amount of data at hand, and the availability of manually processed items, we believe this could be a good candidate for getting Caspeco working with Machine Learning.

Purpose

The purpose of the thesis is to:

1. Evaluate different approaches to providing automated, adaptive categorization of product articles.
2. Implement a practical solution to this categorization problem and plug it into the data pipeline.

Applications

We look for computer science or information systems (B.Sc. or M.Sc.) with an interest in machine learning, natural language processing, data analysis and programming. Knowledge of C# or Azure is considered a plus, but not a requirement.

The work will be performed at Caspeco’s office central in Uppsala, and the scope of the task can be adapted to cover both a Master and a Bachelor thesis. Start date is flexible and applications can be sent to tomas.johansson@caspeco.se