EXAMENSARBETE Subjective matching of products using Machine Learning
STUDENT Alexander Ekdahl & Joacim Åström
HANDLEDARE Elin Anna Topp (LTH)
EXAMINATOR Jacek Malec (LTH)

Using machine learning to predict what you value in home electronics

POPULÄRVETENSKAPLIG SAMMANFATTNING **Alexander Ekdahl & Joacim Åström**

Insurance companies manually find replacements for thousands of products every day. Can a machine replace the manual labor while remaining faithful to the consumers' desires?

Imagine your cell phone breaking and is no longer available on the market, in that case, you probably have a good replacement in mind. Now imagine your TV or laptop breaking as well which makes finding replacements more complicated. Finally, imagine hundreds of thousands of home electronics needing replacements all over the world, every day. This is what insurance companies struggle with on a daily basis, which we find important yet difficult to do. However, machine learning could potentially solve this problem without any trouble. We have investigated the use of machine learning as a tool to help insurance companies evaluate and replace damaged goods. This process is currently handled by an elaborate algorithm that requires day to day maintenance and tweaking to keep up with current market values, and there is great promise for an automated model to alleviate manual labor by learning from past data.

This is where machine learning comes into play. Using an extensive database of products and their attributes along with tens of thousands of unique matches: Can we teach a computer to do this process of matching products automatically? This task proved more challenging than expected. We have not found any existing machine learning technique suitable to perform this task.

Our most successful approach to this problem was to predict a hypothetical replacement for a given product. This hypothetical product would have features which captured the subjective attributes desired in a replacement product. Ultimately, a simple cosine similarity, often used in text searching, would be used to find the product that mostly resemble the hypothetical product thereby determining the best replacement product. For an example, the replacement for an iPhone 4 32GB in 2017 would be another Apple phone, approximately 5-inch screen, with improved camera and performance metrics. The closest matching product to this hypothetical phone would be an iPhone 6 32GB.

The results were promising; however, the original dataset contained many flaws and inconsistencies. Therefore, before any further progress can be made the dataset needs to be improved. Ultimately, the model was able to pick a suitable replacement product 79% of the cases it was given.