1. **Link: <https://www.kaggle.com/c/otto-group-product-classification-challenge>**

**Discovery Business Problem Framed**

The Otto Group is one of the world’s biggest e-commerce companies, with subsidiaries in more than 20 countries, including Crate & Barrel (USA), Otto.de (Germany) and 3 Suisses (France). This company is selling millions of products worldwide every day, with several thousand products being added to our product line.

A consistent analysis of the performance of products is crucial. However, due to diverse global infrastructure, many identical products get classified differently. Therefore, the quality of product analysis depends heavily on the ability to accurately cluster similar products. The better the classification, the more insights company can generate about product range.

Define the model with high accuracy for identifying future product categories. Identify most influential features



|  |  |
| --- | --- |
| Student X | Problem search, discovery of necessary methods, analytic plan |
| Student Y | Code with technical documentation |
| Student Z | 2 presentations for sponsors and analytics |

**3) Analytic plan**

**Initial Hypothesis**

Better use most influential features during classification

**Data**

More than 200,000 the Otto Group products with 93 features, each feature is categorical variable

**Model Planning - Analytic Technique**

Gradient boosting, decision tree, random forest

**Result & Key Findings**

1. Built a predictive model to distinguish main product categories
2. Model’s predictive power is at least as good as currently used techniques by the company
3. Model is invariant to adding new product categories

**Business Impact**

Running the suggested model can help to rise e-commerce revenues at least by 8.8% annually and to reach 6.5 billion euros