

DEVELOPED A MACHINE LEARNING BASED SOLUTION TO CATEGORIZE PRODUCTS AND IMPROVE SEARCH ON THE WEBSITE

For a Leading Distributor



ABOUT THE CLIENT

The client is a leading plumbing and HVAC distributor based in the US.

CHALLENGE

Client had an inventory of 3 million SKUs out of which only ~10% had a product hierarchy attached. Categorizing the remaining SKUs was a slow manual process which directly led to loss of revenue and a poor search experience on the website. The task was to automate the process to reduce revenue loss and improve user-experience.

APPROACH

To achieve this, we had to:

- Extract data elements like product description, vendor name etc. that influence category
- Cleanse data by removing stop words, special characters and perform lemmatization
- Use text mining to create term document metric and evaluate model accuracy
- Add confidence threshold to classify predicted classification as 'Low' or 'High' confidence
- Add business rules and improve model accuracy iteratively by analyzing 'Low' confidence predictions

KEY BENEFITS

- ✓ *The high confidence classified product hierarchy with Machine Learning based algorithm gives the client the ability to take feedback and improve classification over time*
- ✓ *Man-hours are now spent on only validating the 'Low Confidence' outputs. The corrected output from these is fed back into the model's learning algorithm to reduce similar errors in the future.*
- ✓ *The customers are now able to search for these classified products on the website since they can be indexed post-classification and enrichment. This directly improves the search experience, reduces related inbound customer care calls and generates incremental revenue.*

RESULTS

- 30X faster categorization than the existing manual process
- 99% accuracy for 'High Confidence' predictions and overall accuracy of 95%
- \$250K per year cost reduction in 3rd party expenses for manual categorization
- Better search experience on the website and incremental revenue