Identifying Consumer Design Preferences Using Social Media

Alexander Gibbs-Poe, Avi Patel, Ayush Krishnamony, Benjamin Huang, Bhargavi Katuru, Darby Blum, Marcus Orciuch, Michelle Dunn, Prabodh Sonalkar, Oscar Su, Shelja Sarin

INTRODUCTION

Delta Faucet Company is the world's leading innovator in faucets, flush valves and related accessories.

Goal: Identify current and future consumer design/feature preferences for kitchen and bathroom faucets using social media and other websites. Based on the scope of the project, it is divided into three problem spaces –

Problem Space 1:

- Scan and summarize latest feature trends
- Identify market and design opportunity using Social Media sites (Twitter)

Problem Space 2:

- Use online social and other data to scrape product information
- Analyze the scraped data to prioritize existing or trending features/concepts

Problem Space 3:

- Analyze customer satisfaction data to identify areas of opportunities
- Categorize and prioritize the opportunities to improve customer service

METHODOLOGY

Problem Space 1: Our project is split into the following phases:

Figure 1.1: Project Workflow

Data Collection: Used Tweepy API through Python

- Utilized keywords, date, and location to filter
- Extracted username, images, number of likes and comments

Data Cleaning: Removed irrelevant tweets

METHODOLOGY

Data Analysis: Sentiment Analysis and engagement analytics to rank tweets based on popularity

Identify:

- Zero-shot text classification model
- Basic image classification model

Compile Results: Retrieved topics from the most popular tweets

Problem Space 2:

Web Scrapers:

- Utilized BeautifulSoup and Selenium
- Scraped product specifications for various retailer's websites e.g. – Amazon, Home Depot etc.
- Analyzed the data for popular trends

Automation of web scrapers:

- Scrapers feed system with data continuously
- Back-end connects database to system
- Front-end displays results to Delta Faucet employees

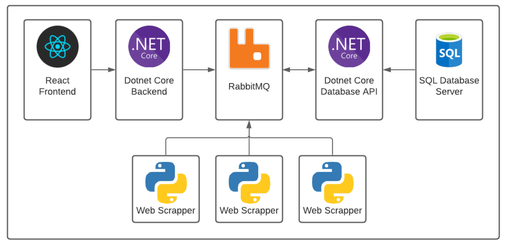

Figure 2.1: Diagram of Automation System

Problem Space 3:

- Utilized Python's BeautifulSoup and Selenium libraries
- Scikit-learn library for supervised classification.
- Built a dataset of scraped reviews.
- Built web app to display the dataset as well as some basic analytics and periodically scrape the web for additional reviews.

RESULTS

Problem Space 1:

Data Collection:

- Scraped 60+ tweets, with their text, images, username, location, number of likes, & number of retweets

	text	name	location	favorite	retweet
17	Hey friends that have a good sense for design/...	jessedriftwood1	NaN	39	1
15	@kareem_carr @arthur_affect I was in one of th...	anng27	Seattle, WA, USA	9	0
31	The organically shaped ZA #faucet matches the ...	TOTOUSA	NaN	7	3
39	We're here to say it and spray it. We've been ...	deltafaucet	Indianapolis, IN	4	1
33	Eleganza is a contemporary collection with min...	LacavaBathroom	Chicago, IL	3	0

Figure 1.2: Table of Data Collected

Sentiment Analysis:

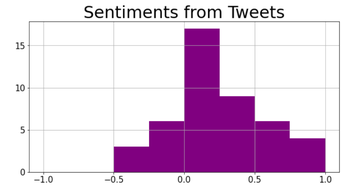

Figure 1.3: Distribution of Sentiment Score of Text from Tweets (Negative: -1.0 to Positive: 1.0)

Text Classification:

- Built a "**Zero Shot" Classification Model** using HuggingFace
  - Used it to filter relevant tweets
  - "Zero Shot" model is a
  - Form of unsupervised learning
  - Categorized tweets based on different topics

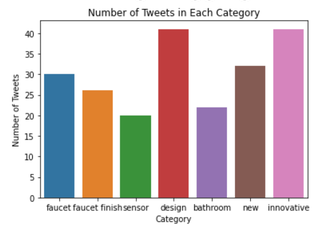

Figure 1.4: Tweets Per Topic

Image Classification:

- Built an image classification model on a test dataset using Keras and TensorFlow to serve as a building block for next year's students

Figure 1.5: Some Images from Tweets Collected

Identifying Consumer Design Preferences Using Social Media

Alexander Gibbs-Poe, Avi Patel, Ayush Krishnamony, Benjamin Huang, Bhargavi Katuru, Darby Blum, Marcus Orciuch, Michelle Dunn, Prabodh Sonalkar, Oscar Su, Shelja Sarin

RESULTS

Problem Space 2:

Web Scraping:

- Pull out spray wand, advanced spray and secure docking are some of the most popular features
- Pull down type faucets are most popular. However, standard faucets are preferred over pull out ones
- Matte Black is the most popular finish followed by Arctic steel and Champagne Bronze

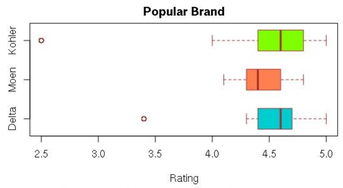

Figure 2.2: Brand vs Rating

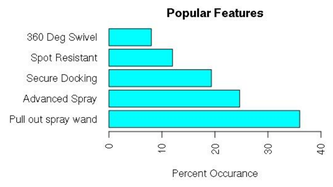

Figure 2.3: Popular Feature

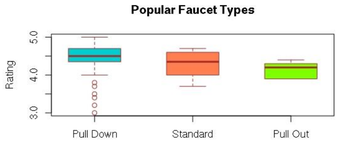

Figure 2.4: Popular Type

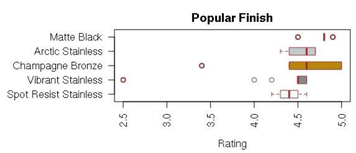

Figure 2.5: Popular Finish

Automation of web scrapers:

- The front-end, back-end, and scraper integration templates are nearing completion.

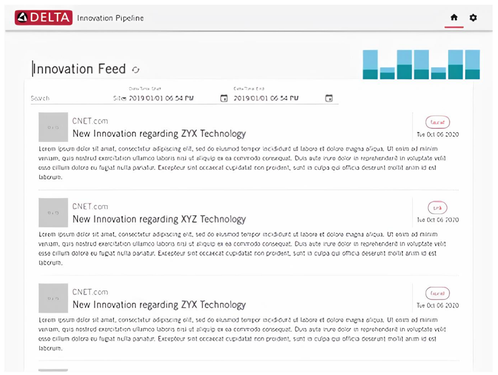

Figure 2.6: Front-end Interface

Problem Space 3:

- Built dataset of 300 reviews from 3 online review websites
- Trained prediction model in scikit-learn using Delta Faucet CSAT data (Fig 3.1)
- Model scores text from 1-5 (1 denotes high difficulty) (Fig 3.2)
- Model currently runs at 68% accuracy (Fig 3.1)
- Construct a web scraper app using the review dataset to display dataset & new reviews

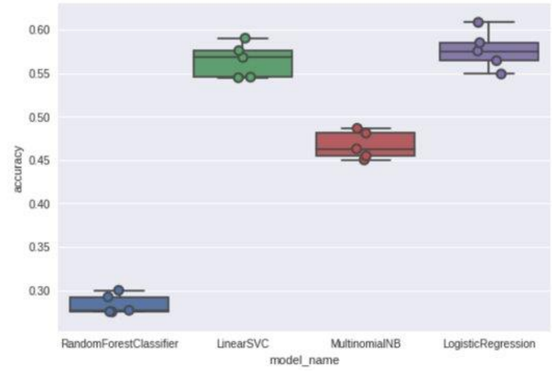

Figure 3.1: Graph of Prediction Model Accuracy

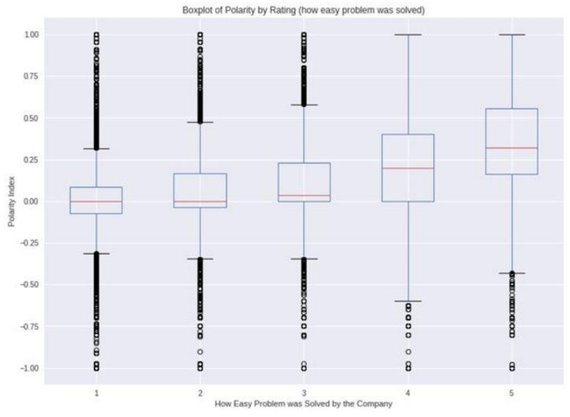

Figure 3.2: Sentiment Analysis Polarity vs Customer Rating

CONCLUSION

- With the analysis models that we have built, Delta Faucet can look at the top faucet features being talked about and the trending topics in different regions of the US.
- Delta Faucet should focus on –
  - Pull down type faucets
  - Pull out spray wand, advanced spray and secure docking
  - Matte black and stainless steel finishes
- The automation platform will show the current innovations in the market.
- Delta Faucet can predict the difficulty of a customer support case based on customer feedback with reasonable accuracy, and created a web app to scrape and display online reviews.

FUTURE GOALS

Problem Space 1:

- Develop a working computer vision model that can identify faucets and their different features, such as finish
- Gain access to other social media data, such as Pinterest, to get more relevant data.

Problem Space 2:

- Scrape more websites and integrate all the scrapers into the automated feed

Problem Space 3:

- Continue working on our web app and look into hosting it using AWS or some similar webservice.

ACKNOWLEDGEMENT

We would like thank both the Data Mine and Delta Faucet Company for the opportunity to work on this project. Specifically, the guidance from Dr. Mark Daniel Ward, Margaret Betz, David Glass and Liz Kershner was critical for the success of the project.