

CS410 Fall 2021 Project Progress @14-Nov-2021

Project Title : Retrieving & ranking customer communication/complaints for optimal routing & resolution of customer concerns

# of team members	: 1 Team Name : KMX
Captain	: Sathish Rama
Team Member	: Sathish Rama
NetID	: sbrama2

Status Update: Identified a public dataset on Kaggle with 5000+ customer complaints. There were lot of invalid rows so did the cleaning. The dataset doesn't have the complaints by product line so working on extracting product lines based on the notes in the complaint.

Blockers : None

Task Id	Task Description	Approx. Efforts in hours	Status
1	Identify publicly available customer concerns dataset that can be used for this project	1	Completed
2	Define/design retrieval strategy - query terms & ranking rules based on the dataset (say based on product lines, categories, define common query terms, based on product lines define ranking rules etc)	3	In Progress
3	Dataset cleanup, feature engineering – remove duplicates, blanks, invalid rows etc	4	Completed
4	Develop the core retrieval & ranking algorithm using query terms and ranking rules. Expected Outcome: Retrieve top complaints based on query & ranking rules and display in console.	10	
5	Evaluation : Label (query & ranks) set of complaints and use this as reference to evaluate the output on non labelled set of complaints	2	
6	Project Documentation & Submission	4	
7	Stretch Goal (if time permits) : Create a conversational complaint search & retrieval interface to key in query & display ranked results as response using IBM Watson chat or AWS tools. Host the core retrieval & ranking algorithm on cloud as function and involve it from IBM Watson chat.	10-20 (Optional)	