Customer Intent Based Recommendation System

Mani Garlapati, Souradip Chakraborthy, and Sunil Kumar Potnuru

***Abstract*—A system to provide customer intent recommendations by leveraging voice of the customer like user reviews/feedback from various platforms (e.g. social media, retail portal etc.). The system uses deep learning and hierarchical clustering to categorize the reviews/feedback from the customer into intent categories by understanding the context and customer sentiment and recommend items from intent categories. Currently retailers are not leveraging customers voice from different channels to make personalized recommendations based on their intent or need. Our algorithm helps to solve this problem by leveraging machine learning and deep learning techniques to identify customer interest categories and interest items in each category. Accordingly, these intent items can be added to the cart based on customers interest.**

[]

***Index Terms*—LSTM, Collaborative Filtering, Cosine Similarity, Customer Interest, Context Matching**

I. Introduction

There is enormous amount of information, which retailers tend to skip, which can be very valuable to business. Once it is captured and analyzed it can improve business and customer satisfaction.

One such information is the customer comments being posted on social networking sites, customer feedback, call center, blogs etc. A variety of platforms record the voice of customers every day which describe their taste, sense, choice, preferences, current and future interests which can help business.

Our process involves capturing the voice of customers and extracting the sentiment and important features from the content which explains customers intent and helps to predict which items would different individuals like and recommend the same from retailers catalog items. We have defined a three step process. Firstly we classify retailer item catalog into different intent categories using hierarchical clustering. Secondly we parse voice of the customer through these intent categories and identify top 3 intent categories for the customer. In the last step we recommend items from these 3 intent categories based on the context and customer sentiment.

II. Overview of methodology

This section presents overview of our methodology. Our methodology consists of the following high level steps:

1) Creation of intent categories from item catalog
2) Mapping voice of customer(VOC) to intent categories
3) Recommending items to an individual

The workflow presented in Figure 1 presents overview of methodology.

Mani Garlapati, Souradip Chakraborthy, Sunil Kumar Potnuru are with WalmartLabs, Bangalore, Karnataka, India

A. Creation of Intent categories from item catalog

Item catalog of a retailer can be used to create intent categories. We use the descriptions of items to create intent categories. This process includes the following steps:

1) Creation and processing of text corpus corresponding to items
2) Hierarchical Clustering of items into intent categories
3) Classifying a new item description to an intent category using deep learning model

B. Mapping voice of customer to intent categories

Identifying interest of an individual can be achieved by utilizing the data from various sources. In order to identify interest of an individual, the following process has been applied:

1) Extraction and collection of voice of customer
2) Pre-Processing and sentiment analysis on tweets [1]
3) Identifying potential categories of interest for an individual using deep learning model

C. Recommending items to an individual

Once we identify potential interest of an individual, we use those identified intent categories to recommend items using two methods.

1) Recommendation Using Word-Word Match and Cosine Similarity
2) Recommendation based on context using LSTM deep learning model
3) Traditional Recommendation if no relevant information is present [2]

III. Methodology and Implementation

In this section high level steps of methodology are explained in detail.

Figure 2 presents the detailed workflow of the whole methodology and implementation.

A. Creation of Intent categories from item catalog

1) Creation of text corpus corresponding to items: Signage descriptions and item descriptions within a category are combined together to form larger descriptions. Category information is also combined with descriptions to form a text corpus. Corpus thus formed will serve as the features for the clustering algorithm at later stage and hence it's important to provide appropriate weightage to the important features. Using fuzzy intelligence and the weight(w) has been optimized.

Table I presents a sample combined description.

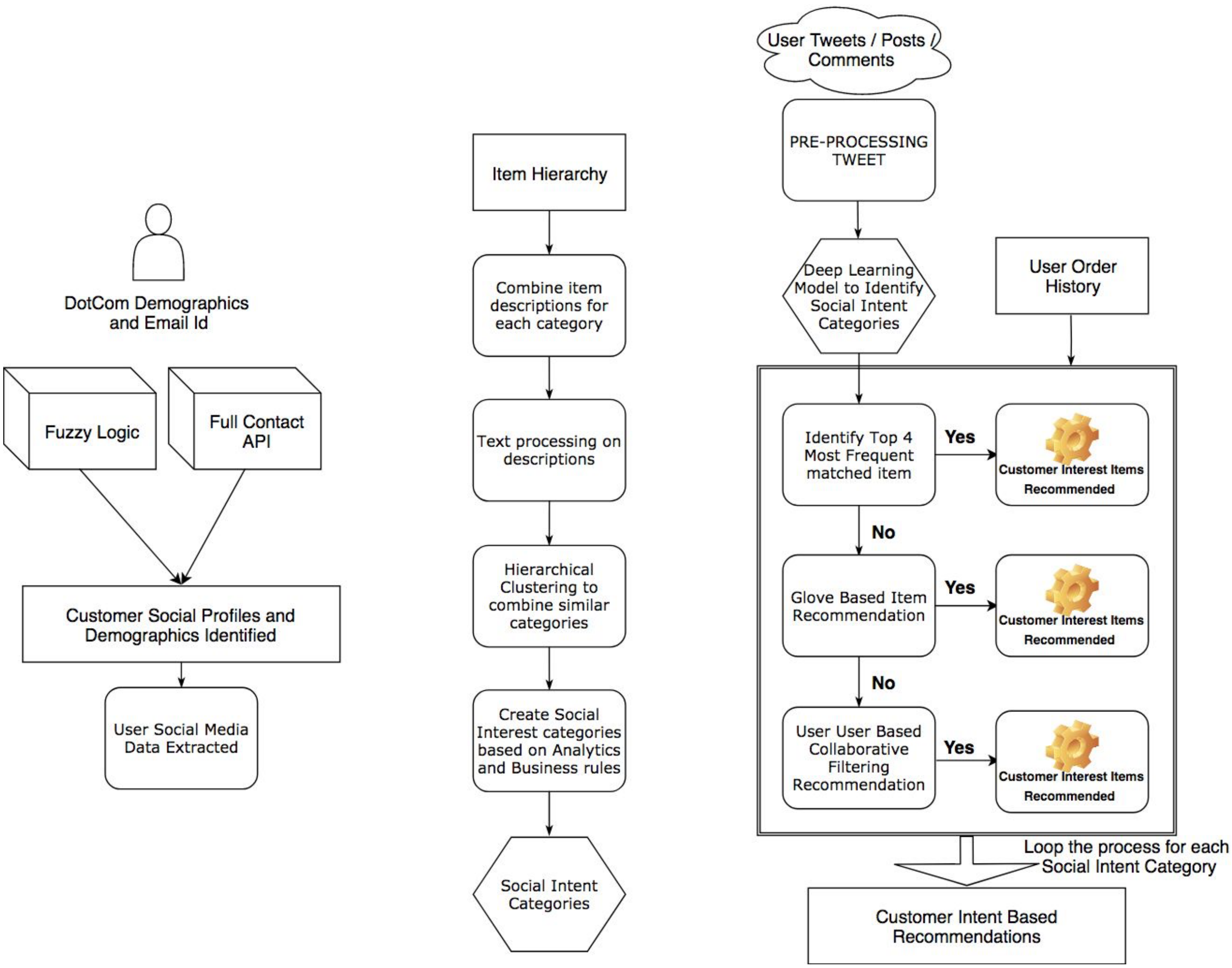

Fig. 1. High level workflow of the process

TABLE I
EXAMPLE OF A COMBINED ITEM DESCRIPTION

CATEGORY NAME	COMBINED DESCRIPTION
BAGGED SOCKS	HANES P10 NO SHOWGB P10 ANKLEFOL P6 GY LOWCUTGB P3 NO SHOW CLRHTGB P12 NS ARCHFRUIT OF THE LOOM LAGB P10 LOW CUT EXTFOL P6 CT CREWHANES P6 NO SHOWNN NO PROMHANES WOMENS SOCKSHANES-GIRLS/WOMENFOL P6 CMFT LWCTGB P3 ANKLE CLRBKCREWFOL P6 ANKLEHANES NO SHOW LADIES SOCKSGILDANGILDAN P3 NO SHOWFOL P10 ANKLE SOCKSNN SOFT & SENSIBLE NO SHOW 10PR BAGHANES P6+2 NO SHOW, WHITE, SZ 5-9HNS 12 PK BUNDLED CREW BLACK SIZE 5-9FOL P6 CMFT LOWCUT SOCKSGILDAN P3 ANKLEGILDAN BONUS PACKSFOL P6 ANKLE SOCKS

Stop-words are removed from the text created text corpus. After removing stop-words, corpus is provided as input to a weighted TF-IDF Vectorizer to obtain a TF-IDF matrix containing vectors only for the important features.

Table II presents the parameters and their values for the TF-IDF vectorizer.

where,

N-gram Range : The lower and upper boundary of the range of n-values for different n-grams to be extracted

Max_df : When building the vocabulary ignore terms that have a document frequency strictly higher than the given threshold

TABLE II
TERM FREQUENCY INVERSE DOCUMENT FREQUENCY MODEL SPECIFICATIONS

Model Parameters	Value Chosen
Ngrams Range	1,2
Max_df	0.9
Min_df	0.1
No of features	2000

Min_df : When building the vocabulary ignore terms that

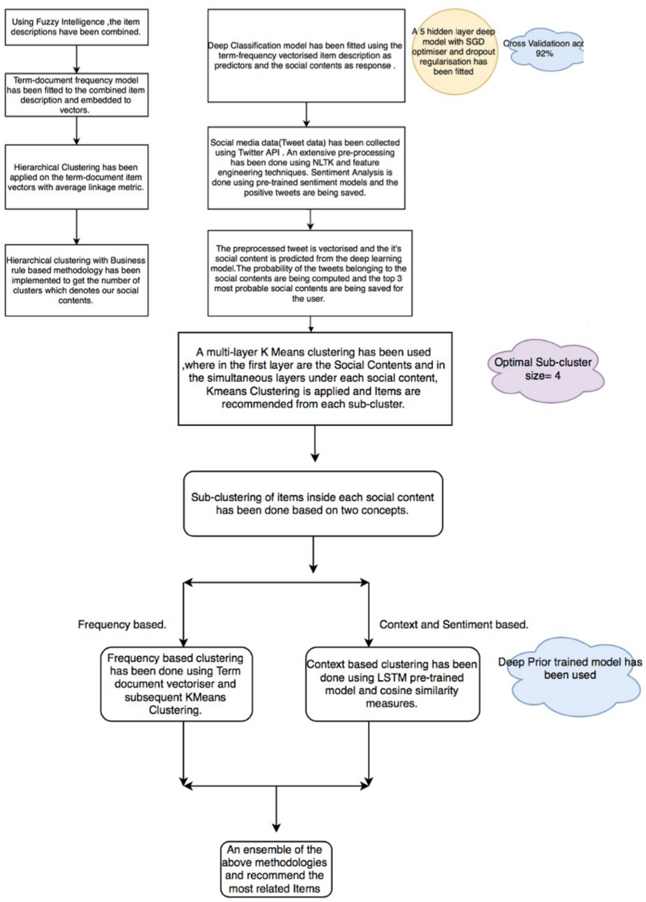

Fig. 2. Detailed methodology of workflow

have a document frequency strictly lower than the given threshold

No of features : Top 2000 features ordered by term frequency across the corpus.

The obtained vector is highly sparse and very large in dimension. It takes lot of time to form clusters using the obtained matrix. We need a denser matrix which could provide information almost similar to the sparse matrix and of smaller in size. In order to achieve that, the sparse matrix is decomposed using Singular Valued Decomposition(SVD).

2) Hierarchical Clustering of items into intent categories: Hierarchical clustering has been used on the corpus vectors to find out the similarity between the different item descriptions and merge the similar descriptions. We perform Top-Down hierarchical clustering on the matrix we obtain from the SVD. The optimal number of clusters were decided using Silhouette Coefficient Plot.

Silhouette Coefficient :

The Silhouette Coefficient [3] is calculated using the mean intra-cluster distance (a) and the mean nearest-cluster distance (b) for each sample.

The Silhouette Coefficient for a sample is (b — a) / max(a, b). The best value is 1 and the worst value is -1. Values near 0 indicate overlapping clusters. Table III presents no of clusters and theri corresponding average Silhouette coefficients.

TABLE III
AVERAGE SILHOUETTE COEFFICIENTS FOR OPTIMALCLUSTER DETERMINATION

No. of Clusters	Average Silhouette Coefficients
3.00	0.0906
6.00	0.0940
9.00	0.1031
12.00	0.1041
15.00	0.1164
18.00	0.1375
27.00	0.1881
30.00	0.2054
33.00	0.2176
36.00	0.2268
39.00	0.2328
42.00	0.2383
45.00	0.2393
48.00	0.2469
51.00	0.2430
54.00	0.2434
57.00	0.2504

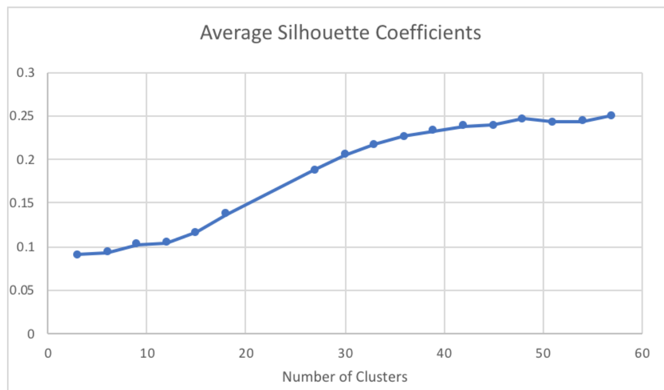

Fig. 3. Average Silhouette Coefficient Plot for Optimal Number of Cluster determination

As it can be seen from the Figure 3 , the average silhouette coefficient reaches its maximal point when the number of clusters is between 48-50 and at 49, it reaches the peak. Hence, the optimal clusters or Customer Intent Categories have been taken which is in line with business sense.

The resulted clusters are named based on the most frequent working term within each cluster. In this case, we have clustered items of apparel category into various customer intent categories.

Output Of Hierarchical Clustering and Customer Intent Clusters :

Tables IV, V, VI, VII, VIII and IX present example of few customer intent categories. Table X presents list of intent categories.

TABLE IV
CUSTOMER INTENT CATEGORY-EXAMPLE 1

Existing Category Descriptions	Customer Intent Category
Boy Boots	
Boy Winter Boots	
Girls Boots	
Ladies Boots	Shoes & Boots
Ladies Winter Boots	
OCC Slip Resistant	
Occ Work shoes	

TABLE V
CUSTOMER INTENT CATEGORY-EXAMPLE 2

Existing Category Descriptions	Customer Intent Category
Bagged Socks	
Athletic Socks	
Socks(D23)	
Wellness Socks	Socks
Casual Socks	
Seasonal Socks	
Normal Socks	

TABLE VI
CUSTOMER INTENT CATEGORY-EXAMPLE 3

Existing Category Descriptions	Customer Intent Category
Scrubs	
Watches	
Backpack	
Bedding and Accessories	
Accessories(D23)	
Handbags	Accessories
Belt	
Leather goods	
Hats	
Accessories(D33)	
Scarves	

TABLE VII
CUSTOMER INTENT CATEGORY-EXAMPLE 4

Existing Category Descriptions	Customer Intent Category
Men's Athletic	
Sports Elastic	
Ladies Athletic Sportswear	Sports &Athletics
Athletes Tights	
Sporty Tees	
Men's Playwear	

TABLE VIII
CUSTOMER INTENT CATEGORY-EXAMPLE 5

Existing Category Descriptions	Customer Intent Category
Basic T-shirts	
Screen T-shirts	
Event T-shirts	T-shirts
T-shirts(AK)	
Screen T-shirts(High Misc.)	
T-shirts(D23)	

TABLE IX
CUSTOMER INTENT CATEGORY-EXAMPLE 6

Existing Category Descriptions	Customer Intent Category
Keyrings	
Jewellery Diamond	
Jewellery Gifts	
Jewellery Gemstones-CZ	Jewellery
Metal Showcase	
Jewellery Sunglasses	
Jewellery Bracelets	
Necklace	

3) Classifying a new item description to an intent category using deep learning model: Clusters obtained from the hierarchical clustering are used as the intent classes. These intent classes act as classes for classification task. The classification task helps in automatically categorizing the item descriptions to a customer intent category.

TABLE X
CUSTOMER INTENT CATEGORIES FROM HIERARCHICALCLUSTERING

Customer Intent Category Cluster Names		
Jewellery	Daywear	Shoes
Socks	Shorts	Maternity
Accessories	Swimwear	Fleece
Undergarments	Slippers	Licensed
Activewear	Gifts	T-shirts
Knit	Woven	Outerwear
Sports	Baby	Holiday Collection
Puerto Rico	Dress	School
Casual	Tops	Kids
Seasonal	Capris	Rain
Alaska	Missy	Others
Hawaii	Jeans	Tights
Thermal	Beachwear	Skirts
Pants	Workwear	Promotions
Sweaters	Jackets	Dancewear

A deep learning model [4] is fitted with these classes as response variables and the item descriptions as the predictors. So, each time a new item description comes up, using the deep learning model we trained, it is easy to predict customer intent category for the item description.

Once the categorization is done for all the item descriptions to proper categories, the trained model helps retailer a lot as the same model could be leveraged to increase efficiency in identifying items when a customer searches for an item. The model built is cross validated and accuracy of the model turned out to be very good. Table XI presents specifications of deep learning model we built. Table XII presents details of our deep neural network. Table XIII presents epoch number and cross validation accuracy at the end of the epoch.

TABLE XI
SPECIFICATIONS OF DEEP LEARNING NEURAL NETWORK MODEL

Model Parameters & Hyperparameters	Values
Input Dimension	2000
Output Dimension	49
No. of hidden layers	5
Activation Function in hidden layers	Rectified Linear Units
Activation Function in output layer	Soft-Max
Regularization Type	Dropout
Optimization Type	Stochastic gradient descent
Loss Function	Cross entropy

TABLE XII
DEEP NEURAL NETWORK STRUCTURE FOR APPARELCLASSIFICATION

Layer(type)	Output Shape	Parameters
Dense_1	(Nrows,1000)	2001000
Dropout_1	(Nrows,1000)	0
Dense_2	(Nrows,800)	800800
Dropout_2	(Nrows,800)	0
Dense_3	(Nrows,400)	320400
Dropout_3	(Nrows,400)	0
Dense_4	(Nrows,200)	80200
Dropout_4	(Nrows,200)	0
Dense_5	(Nrows,100)	20100
Dense_6	(Nrows,49)	4949

TABLE XIII
NO. OF EPOCHS AND CROSS VALIDATION ACCURACY

Epochs	Cross Validation Accuracy	Time
1	0.7263	10sec
2	0.8016	9.7 sec
3	0.8299	9.8 sec
4	0.8521	8.9 sec
5	0.8628	8.8 sec
6	0.8715	8.7 sec
7	0.8788	8.76 sec
8	0.8797	8.6 sec
9	0.8861	8.6 sec
10	0.8917	8 sec
11	0.8916	8 sec

Figure 4 shows training and validation loss with increase in epochs and Figure 5 shows training and validation accuracies with increase in epochs.

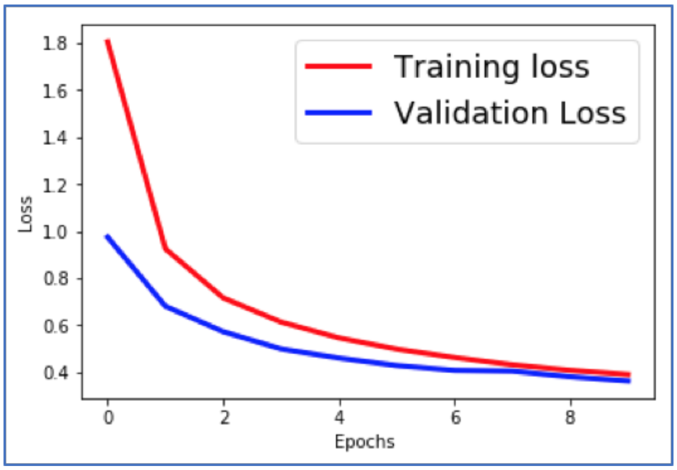

Fig. 4. Training and Validation Loss for Deep Learning Classification Model

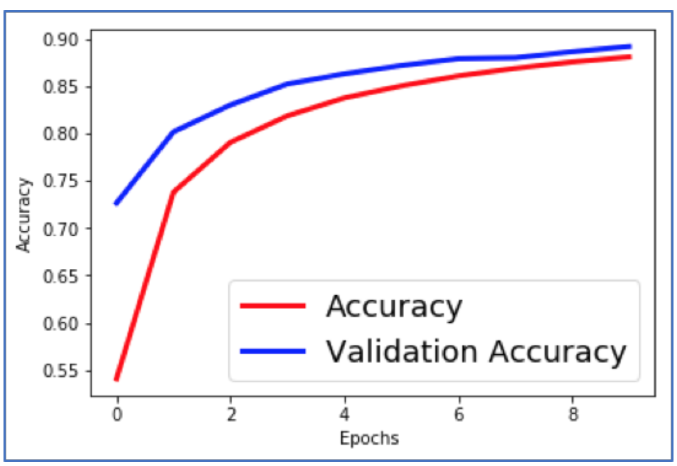

Fig. 5. Training and Validation Accuracy for Deep Learning Classification Model

B. Mapping voice of customer to intent categories

1) Extraction and collection of voice of customer: We collect the voice of customer in many forms such as social comments, customer feedback, posts, reviews etc. But for this case, we have taken social activities of customers on Twitter.

Twitter comments of an individual in the past 2 to 3 weeks are collected using full contact API and fuzzy logic. Table XIV shows sample of tweets we have collected after removing hyperlinks and user related information.

TABLE XIV
SAMPLE TWEETS

Tweet
Im not crying, its just emotion
I still yell at the fact that I missed avengers and captain america . THats called A FRIEND
RT @user: It's perfectly OK to not love your spiderman ,that's fine@@
RT @user: s/o to the girls that hate love stories,
I just really do miss Hermoine from harry potter
the "y'all girl" Im still laughing
RT @user she really might lol but I always wanted Mona to be supergirl
RACHEL IS SERIOUSLY THERE GIVING ME ALL MY
Im crying laughing hahahahah RT @user
okay RT @user just let them don't go
Imagine thinking I do not want whole freaking fanfiction about this
hahaha thanks for the idea RT
if I see @user cry because she misses her wife-to-be and daughters, I will not be okay
RT @user: the father, the son, and the holy spirit
still shooketh Demetria is top
RT @user: if it at least helped you feel better and not the opposite
RT @user: We love facts
RT @user: please
RT @user: I love you justin
RT @user: Remember when 13 Reasons king queen stories sucked
RT @user RT @user: Was there any goss about it?
RT RT @user: Woman after my own heart!
RT @user: Want more @EverythingSuxTV on @netflix? Watch Captain America, it's amazing
RT @user: I read harry potter and cry@user #TMYLMTour #YoureMYHero
RT @user: the "i don't understand how i can miss avengers part " starter pack
RT @user: movies like superman motivates
RT @user:Can we take a minute to discuss the Friday box office? The top 5 films were: a female led action movie, a POC led superhero movie, a gay teenage coming of age story, a film with one of the most diverse casts I've ever seen. 1 of the 5 is about a straight white dude. This is epic
RT @user:Love these girls so much! @IISuperwomanII @user @user
RT @user:why is drunk sasha the funniest, hottest and most adorable thing at once? i can't cope
RT @user:We are very excited to announce that we have decided to organize the very first Pretty Little Liars Convention
RT @user: I still watch cartoons
RT @user: I love comfortable tshirts
RT @user:Ok I am ADDICTED to @FortniteGame I've dedicated my entire weekend to playing pokemon go

2) Pre-processing and sentiment analysis on tweets: We collect twitter data of an individual and perform few steps of pre-processing such as removing stop-words, filtering unnecessary tweets based on rules. We classify the pre-processed tweets based on sentiments into positive and negative tweets.

The positive tweets are vectorized using previously built TF-IDF Vectorizer. So, the words which are relevant to apparel category are being extracted from the tweets using the previous built TF-IDF vectorizer. Negative tweets are used as a filter ensure we do not recommend the items which the user may not like [1] .

3) Identifying potential categories of interest for an individual: We consider the tweet TF-IDF vector obtained from TF-

IDF vectorizer as a test data point to the deep learning model we trained. We consider intent categories predicted with the highest probabilities and in this case we have considered top three intent categories.

Top 3 classes predicted for the tweets shown in Table XIV are:

- T-shirts
- Licensed
- Accessories

The customer intent categories that are identified are quite relevant in this case. From the sample tweets shown in Table XIV we can observe that person is mainly interested about super heroes. The top 3 recommendations justify that items related to super heroes and items of similar context from apparel category needs to be recommended for the individual.

C. Recommending items to an individual

The final item recommendation to Customers from the top 3 Customer Intents for that customer have been done in two ways:

1) Recommendation Using Word-Word Match and Cosine Similarity [5]: In this approach, we have primarily focused on the word matches with the tweet and the item descriptions for the top 3 customer intent categories.

So firstly, a term-frequency based model is formed with all the item descriptions corresponding to the previously obtained top 3 categories.

Then the tweet is being vectorized with the current model and then a cosine similarity matrix is being formed with all the item vectors and the tweet vector. The top 10 similar items are selected and recommended for the user.

To ensure that same item (just varying on size or gender) doesn't get selected many times, we have introduced a sub-clustering and recommended from each sub clusters.

2) Recommendation Using Context Match [6] and LSTM Deep Learning Model: In this approach, we have primarily focused on the context of the word and hence used word embeddings to represent the item vectors. Word embeddings represents the word in some abstract dimensions which explains the inner context of the word and document.

LSTM deep learning neural network model is used to determine the embeddings for each of the item description. Similarly, we have determined the vector for the tweet and computed similarity-based recommendations for the tweet.

Here, the semantic similarity of the tweet with the item gets captured helping us for an advanced recommendation.

3) Traditional Recommendation if no relevant information is present [7]: If the tweets are not really guiding us with any product recommendation with significant confidence/high probabilities, then we will go for standard recommendations based on user-user/item-item similarity.

We can use the low probability selected items from our model to leverage the process.

Table XV shows final recommendations made to customer by our system.

The top customer intents and corresponding item recommendations based on the tweet example mentioned above have been presented from the Customer Intent basket.

TABLE XV
FINAL RECOMMENDATIONS MADE TO CUSTOMER

Reccommendations Of Items	Social Content
MV FACE OFF HOCKEY MV FACE OFF HOCKEY DC COMICS JUNIORS "CAPTAIN AMERICA CIVIL DISNEY - DISNEY AP	Tshirts
MARVEL MEN'S CAPTAIN ONLINE ONLY MARVEL MEN'S CAPTAIN AMERICA AND SPIDERM ONLINE	Tshirts
CAPTAIN SHIELD PANEL CAPTAIN SHIELD PANEL JUNIORS TANK TOP CAPTAIN AMERICA AP	Tshirts
FIRE WITHIN SPIDERMAN TEST 2017 MENS TEE LICENSE	Tshirts
BATMAN LOGO SOLID GREY BATMAN BOYS TEE BATMAN AP	Licensed
SUPERMAN CTNIC TEE SUPERMAN POLY TEE SUPERMAN BOYS TEE SUPERMAN AP	Licensed
SUPERMAN LOGO SOLID NAVY SUPERMAN BOYS TEE SUPERMAN AP	Licensed
CAPTAIN POLY POLY HOOD AVENGERS HOODIE AVENGERS AP	Licensed
SPRGRL PSSPRT XBODY PASSPORT XBODY BAG SUPERGIRL PASSPORT CROSSBODY BAG SUPERMAN AP	Accessories
KIDS' BACKPACKS ONLINE ONLY $9.88 SUPERMAN MINI BACKPACK (P9CM27-W) WARNER BROTHERS	Accessories
SUP4050WM SUPERMAN L APPAREL SUP4050WM SUPERMAN LCD POP FLASH SUPERMAN	Accessories
CAPTAIN AMERICA CAPTAIN AMERICA ADULT LIC WATCH DISNEY - DISNEY	Accessories

IV. CONCLUSION

The above method has proven to provide logical and meaningful results in the context of item recommendation for various test cases.

The cross-validation accuracy of deep-learning model was above 91%.

The time complexity for the entire process is less and it has been proven to be computationally efficient.

The methodology has been designed keeping in mind scalability and its application across various categories of any retailer.

Overall the algorithm can extract customer interest categories from customers voice in any form and recommend items which is as per needs of the customer.

REFERENCES

[1] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," *CS224N Project Report, Stanford*, vol. 1, no. 12, 2009.

[2] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl, "Item-based collaborative filtering recommendation algorithms," China, 2001, p. 295.

[3] S. Aranganayagi and K. Thangavel, "Clustering categorical data using silhouette coefficient as a relocating measure," in *International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007)*. IEEE.

[4] H. Liang, X. Sun, Y. Sun, and Y. Gao, "Text feature extraction based on deep learning," *EURASIP*, no. 2017, 2017.

[5] A. Huang, "Similarity measures for text document clustering." [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.332.4480&rep=rep1&type=pdf

[6] J. Pennington and R. Socher, "Glove: Global vectors for word representation." [Online]. Available: https://nlp.stanford.edu/pubs/glove.pdf

[7] R. Zhang and Q. dong Liu, "Collaborative filtering for recommender systems," in *2014 Second International Conference on Advanced Cloud and Big Data*. IEEE.