Using AI to understand how customers feel

Predict Net Promoter Scores and identify whether your customer is potentially a promoter, neutral, or detractor. Take corrective actions timely to improve customer service.

Credits | Prashanth Suresh Babu   Sumit Thakur

Measuring CSAT is Crucial for DSPs to improve Customer Service Operating Procedures

However, only **15-20%** of customers respond to CSAT survey

> Digital Service Providers (DSP) must ensure that they provide **right customer service** every time and across all touchpoints.
> **A dissatisfied & frustrated customer quickly opts to switch the service provider.**

It is vital for DSPs to get insights into customer satisfaction levels, in order to bring continuous improvements in customer service operating procedures.

However, a research shows that, **only 15-20% of customers respond to the CSAT survey** after their interactions with the customer support service. Thus, it is crucial for DSPs to **predict the remaining 80% of CSAT responses.**

Among different touchpoints, **Virtual Agents(VA)** play vital role in shaping customer experience. According to Gartner, by the year 2020, <u>**85% of customer interactions**</u> will be managed without a human.

25% of customers **will switch providers** after **1** bad experience

45% **will abandon a purchase** if required to make **2** or more separate interactions

19% **will not trust a company** after **1** bad experience

Source: TM Forum

| **A rule based system is not effective**	**Less Reusability** Reusing across business units is difficult	**Not Scalable** Cannot be easily scaled, if volume of chat transcripts increases exponentially	**Tedious and Complex task** • Requires categorizing and defining rules for each customer utterances • Rules are difficult to trace back in the event of any exceptions

Thus most innovative DSPs are trying to address this problem with **Machine Learning** approach.
This insight details on how a DSP can implement **ML based CSAT prediction model** for customer service center.

Prodapt

Building a robust ML Model can address problem of CSAT Prediction and bring Continuous Improvements
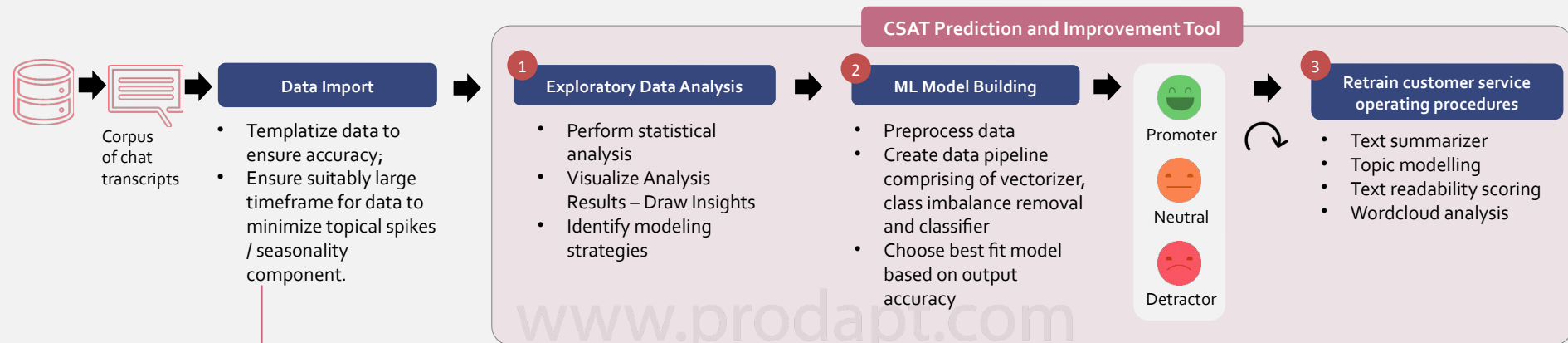
Corpus of chat transcripts

Data Import
- Templatize data to ensure accuracy;
- Ensure suitably large timeframe for data to minimize topical spikes / seasonality component.

CSAT Prediction and Improvement Tool

1 **Exploratory Data Analysis**
- Perform statistical analysis
- Visualize Analysis Results – Draw Insights
- Identify modeling strategies

2 **ML Model Building**
- Preprocess data
- Create data pipeline comprising of vectorizer, class imbalance removal and classifier
- Choose best fit model based on output accuracy

Promoter

Neutral

Detractor

3 **Retrain customer service operating procedures**
- Text summarizer
- Topic modelling
- Text readability scoring
- Wordcloud analysis

Fig: Sample image of templatized data import

	StartTime	EngagementId	Transferred	Escalated	CSAT Score	AgentText	CustText	Text
79	2019-06-03T00:10:33-0600	7333445206937329992	T	F	NaN	[a:script] got it. please tell us your first n...	[c:] i would like to cancel my directv now ser...	[c:] i would like to cancel my directv now ser...
746	2019-06-03T03:05:32-0600	7358777796172233984	T	F	NaN	[a:script] so i can assist you further please ...	[c:] hello i am trying to find my account numb...	[c:] hello i am trying to find my account numb...
752	2019-06-03T03:06:38-0600	7358777796176558976	T	F	NaN	[a:script] are you asking about:can i review o...	[c:] primary number[c:] primary number for acc...	[c:] primary number[a:script] are you asking a...

Building a robust ML model can predict CSAT score with an overall **F1 score** (combining precision and recall) of **82%**

DSPs need a Machine Learning solution to

Predict CSAT response for remaining 80% of customer base who don't respond to survey. This information would have been lost otherwise.

Analyse and find gaps in agent's response to the customer.

Retrain the service operating procedure to bring continuous improvement

Further slides bring out more details on the key steps towards building the ML model

Prodapt

Exploratory Data Analysis (EDA) - Formulate an Effective Model Building Strategy

Discover patterns & anomalies in the obtained **15-20%** of CSAT response

A	**Perform Statistical Analysis**	Identify properties such as delimiters, chat timestamps, and the average length of a given interaction; perform this for both agent and customer components of text as required. Discover significant patterns & spot anomalies.
B	**Visualize Analysis Results – Draw Insights**	Visualize distribution of the variable under survey; in particular, there is bound to be class imbalance (especially for data such as CSAT). Use appropriate choices of visual representation to obtain insights such as data distribution, structure, and skewness.
C	**Identify Modeling Strategies**	Leverage properties of data and insights to identify modeling strategies such as hypothesis formulation, approach to be used and data cleanup steps to be performed.

Fig: Sample outputs of EDA analysis

Obtain Insights from Summarized Feature Details

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 95875 entries, 79 to 164541
Data columns (total 8 columns):
StartTime        95875
EngagementId     95875
Transferred      95875
Escalated        95875 non-null object
CSAT Score       20916 non-null float64
AgentText        95875 non-null object
CustText         95875 non-null object
Text             95868 non-null object
dtypes: float64(1), object(7)
memory usage: 6.6+ MB
```

Only about a fifth of customers submit CSAT rating

The column details suffices to find that only a fraction of customers **(20% entries of all the records)** submit a CSAT survey.

Check Distribution of avg customer word count per CSAT score

Customers clearly tend to prefer a speedy resolution to their issue
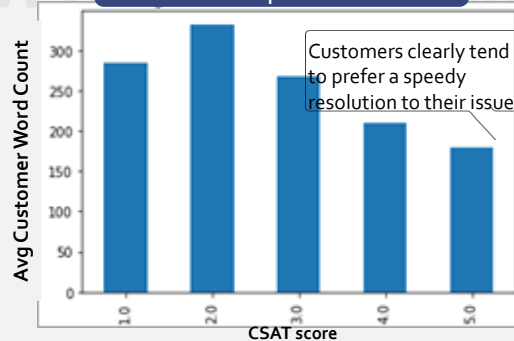
In general, **longer the conversation, higher the propensity for lower CSAT score**..
*Analysis of CSAT with score 1 showed that some customers closed chat earlier due to frustration

Check Distribution of CSAT Scores

No. Of Customers per CSAT Score

Here Neutral seems to fall in minority class

Detractor
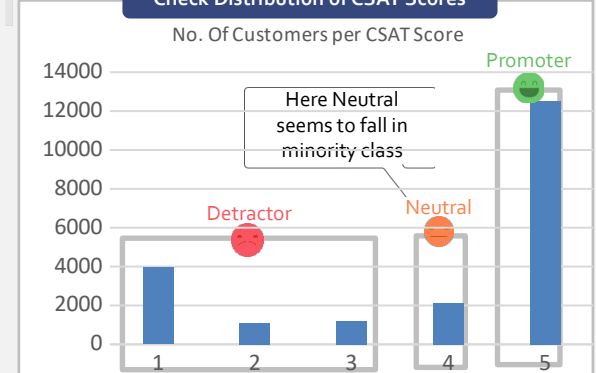
Neutral

Promoter

Here we see **skewness in data,** as the number of detractors (1+2+3) and the number of promoters (5) are significantly higher than neutrals (4).

Prodapt.

ML Model Building - Vectorize Chat Transcripts and Remove Class Imbalance Before Building Classifier Models

A — **Use TFIDF for Vectorizing Corpus of Chat-transcripts**

Term Frequency – Inverse Document Frequency (TFIDF) is recommended to create word vectors as it takes into account the frequency and relevance of words in the context of the text corpus.

B — **Correct Class Imbalance**

Data imbalance is typically present for CSAT scenarios. Ensure that, promoters, neutrals & detractors are represented equally while fitting the model. Below **class imbalance removal techniques** can be used.
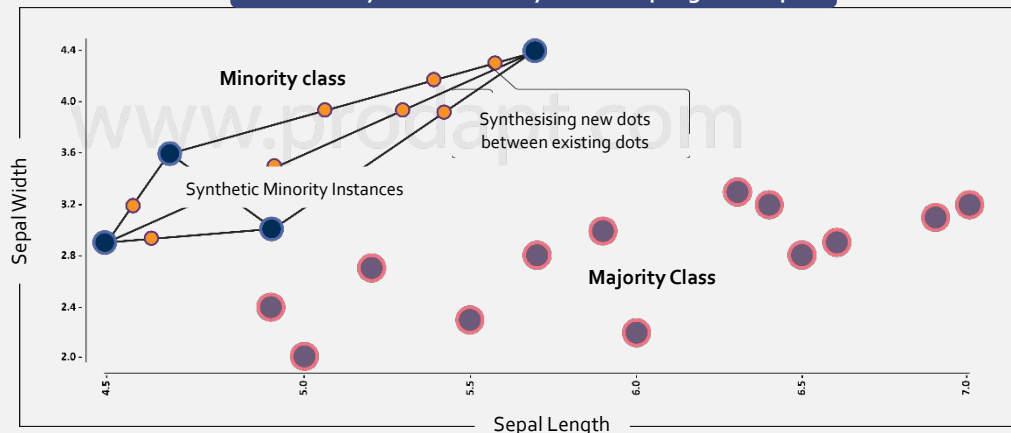
Undersampling the data

Deliberately taking fewer samples from the majority class – i.e., the class having the highest proportion of values – to 'balance out' the overall class distribution.

Oversampling the data

Taking multiple samples of the minority classes to increase their representation in the overall distribution.

SMOTE - Synthetic Minority Over-sampling Technique

Minority class

Synthesising new dots between existing dots

Synthetic Minority Instances

Majority Class

Sepal Width / Sepal Length

Fig: Sample image of SMOTE technique used to correct class imbalance

- Leverage SMOTE technique for increasing the number of cases in the dataset in a balanced way.
- The technique works by generating synthetic samples from the minority class that can be used as an input to the classifier model.

C — **Build Classifier Model**

Once the imbalances are rectified, the obtained CSAT scores can be used as a training data set to build a classifier model.
Fit different statistical model that predicts the CSAT levels (i.e. Promoter, Neutral or Detractor) for the remaining 80% of records.

Prodapt

ML Model Building-Identify Best Fit Model Using Confusion Matrix and F1 score

Predict if remaining 80% of customers will be a promoter, neutral or detractor based on the chat utterances

Original dataset of chat transcripts with known CSAT scores
[Total Records – 20,916]

Training Set
[16,732 records] – **80%**

Validation Set
[4,183 records] – **20%**

Train the model for different classification algorithms

XGB   DMM   SVM

Random Forests   Neural Networks

Fit model over list and choose best classifier based on output accuracy

Best Fit Model

Enhance model for target metric

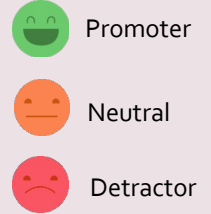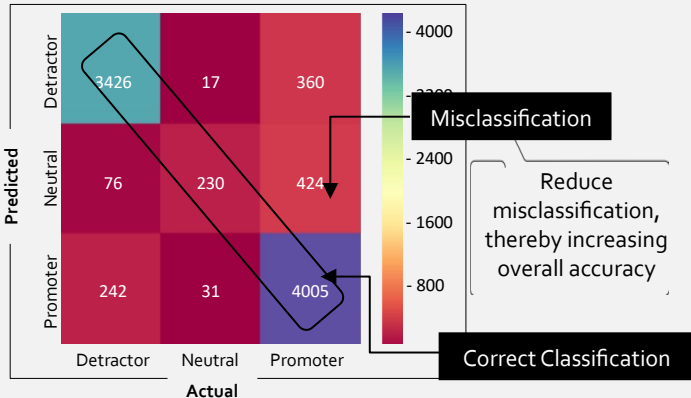Fine tune **Hyper Parameter** based on Confusion Matrix and F1 Score

CSAT Prediction

Promoter

Neutral

Detractor

Fig: Sample image of confusion matrix and Key Accuracy Metrics

Misclassification

Reduce misclassification, thereby increasing overall accuracy

Correct Classification

	Precision	Recall	F1 Score
Detractor	0.82	0.78	0.80
Neutral	0.83	0.32	0.46
Promoter	0.84	0.94	0.88
Avg/Total	0.83	0.83	0.82

Keep finetuning the hyper parameters until you achieve desired precision, recall and f1-score.

## ML Model's output - Distribution of CSAT score for entire customer base
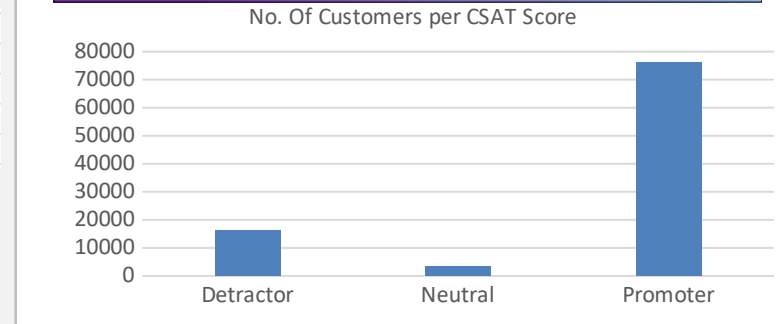
No. Of Customers per CSAT Score

Prodapt

Promoter

[c:] reschedule service call for today[c:] upcoming[c:] repair[c:] installation[c:] xxxxxxxx[c:] who reschedules appts[c:] sheila[c:] i can't be here at 12 today because of work. would like to **reschedule for wednesday**.[c:] xxxxxxxxx [c:] 12-4pm[c:] **i also want** to verify that an additional 24 month agreement is not required for the work being completed. i was told last night it wasn't but on the confirmation email it was indicated at the bottom.[c:] ok **that sounds great**. i **appreciate** you changing the appt. for me.[c:] i will take care of that. **thank you**.[c:] you have been **very helpful**. that is all for today. **have a great day adam**.
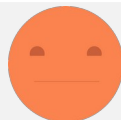
Neutral

[c:] hi there my internet hasnt been working for 4 days[c:] it said that **youre unable to troubleshoot at this time**...[c:] **please connect me to a human**[c:] nina[c:] hi myra[c:] my internet hasnt been working all weekend[c:] i've unplugged and plugged it in i've restarted it[c:] nothing[c:] its **frustrating** because i'm paying for it of course[c:] than kyou[[c:] i'm not home anymore but when i left this morning[c:] the modem had the power light on green[c:] and then the rest was red[c:] okay![c:] **will do thank you very much**[c:] **when will my tech appointment be**?[c:] okay **thank you very much**[c:] that would be all thank you

Detractor
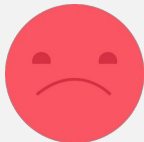
[c:] i need to contact a manager about an issue [c:] billing[c:]. i have an issue that has not been resolved despite many calls[c:] i signed up with internet xxx via a representative making sales in our neighborhood. he told me the installation fee would be waived as apart of my offer. that did not occur. i was charged $X. my due date was delayed for them to do an investigation. i never heard from them. i went ahead and paid the bill because i did not want any negative information on my payment history. i am **beyond disappointed** that this **fee has not been waived**. i have the paper the rep gave me.[c:] there should be some sort of note on my account . is there not?[[c:] no worries. thanks for checking[c:] i have already paid the bill. [c:] i **need to talk with a manager** is there a number i can reach[c:] **i don't understand** why i can not send you the paper i was given. if a **representative stated the fee would be waived even if he were mistaken that is not my fault and u should stand by what their representatives say**. otherwise it looks like a bait and switch and very underhanded tactics. [c:] i am **so very disappointed**. [c:] that is the same number **i have called 5 + times**[c:] do you mean lower future bills[c:] ok

Customer Text Summarizer	ML-generated summary of the customer utterances in the conversation **brings out the key essence of long chat transcripts** • **Gensim based text extractive model** is recommended to build this functionality

Agent Text readability scoring	Analysis of the agent utterances, by finding the readability index helps in retraining the operating procedures of agents/chatbots • Denotes the grade-level at which a person can understand the text • Readability tests such **as Coleman – Liau index** can be leveraged using Text stats-based model

AgentText	Customer Text Summary	Readability
[a:script] good afternoon![a:script] got it. please tell us your first name.[a:script] thank you. you will now be connected to a specialist for shopping assistance.[a:script] hi! my name is ana. i'm happy to help! let's get started.[a:script] to better assist you i'd like to ask you a few questions.[a:] where did she purchase the device?[a:script] **thanks for the info!**[a:script] **sounds great.**[a:] in this case yes **you can get that insurance with us the one of $X.XX or you also have the option of apple care**[a:] if you want the one with us so that is done at any att store[a:] i can make an appointment for you to go and get it[a:] i'm afraid that can be done at the stores only :([a:] **if you want i can book an appointment for you**[a:script] you're welcome.	hello.i am trying to add device protection to a phone my wife just purchased. daniel.id like to add device protection to the phone withx.741.xxx number. the plan that is $X.XX a month. apple store.it is a brand new phone.she just activated it today. id like thex.x option through x.can it be done online okay.ill break the news to her.all good. thanks for the help.	5.83
[a:script] got it. please tell us your first name.[a:script] thank you. you will now be connected to a specialist for account support.[a:script] hi. my name is rachel. to get started can i please get your name?[a:] hi delanie![a:] how may i help you with your uverse bill?[a:] i understand. thank you for letting me know about it.[a:] **is this the access program account?**[a:]while pulling up the account are xxx-xxx-xxxx andxxx-xxx-xxxx good contact numbers so we can reach you by phone or text with information about your services?[a:] thank you for confirming it to me.[a:] upon checking here on the account i can see that **you are in the access account** delanie.[a:] i just want to let your expectation that they are available through phone.[a:] thank you for that information.[a:] we have limited access on your account since it is an access account but if you want i can call you right now to connect you over to them?[a:] sure it's okay. please allow me a moment.[a:] here's the number delanie "xxx-xxx-xxxx".	need help with my bill reverse $XX if I put a card on my account late fees should be taken off could you give me number to call them thankyou	8.41
[a:script] i apologize for the inconvenience. i can help you with that.[a:script] ill look into this for you right away.[a:script] to pull up your account may i know the 9-digit u-verse account number or the telephone number registered on the account?[a:script] thank you.[a:script] i'm **remotely testing** the internet connectivity to find out the reason for slow internet bandwidth.[a:script] i have checked and it seems that the **network is functioning to it's optimum level** which suggest no **network fault** or line issues.[a:script] this seems to be an issue with the modem **wireless settings** and i would make changes with the **wireless frequency** channel.[a:script] in order to clarify changing the channel would decrease the chances of **signal frequency clashes.**[a:script] i would change it to channel 9 and mostly household equipment emits **wireless frequency ranging between 1-11 except 9.**[a:script] may i have your permission to access your modem settings?[a:] are we connected?[a:] .	technical support.this cora. my internet is down. ok i am having trouble with my internet.yes xxxx. yes this is the number on my account. thanks again sheldon. chat window restored thanks for your help. chat window restored.	12.33

> Higher readability index indicates agent's response is complex.

Fig: Sample examples of Text Summarizer and Agent readability scoring

Performing **"Cross Correlation Analysis"** of predicted CSAT score with Text summary & Readability Index helps to identify gaps and retrain operating procedures.

Prodapt

Retrain Customer Service Operating Procedures - Perform Topic Modelling & Correlate Topics with Customer's Positive and Negative Sentiments

Fig: Sample flow for topic modelling and topic categorization

```
Topic 0: day,sure,good,help,great,just,thanks,ok,yes,thank
Topic 1: moving,modify,customer,services,directv,like,need,want,service,cancel
Topic 2: verse,wireless,article,esupport,html,www,com,att,url,https
Topic 3: want,cell,sim,iphone,need,insurance,new,number,mobile,phone
Topic 4: hello,rep,question,credit,plans,usage,features,service,customer,billing
Topic 5: cancelled,charge,removed,want,charged,remove,package,sunday,nfl,ticket
Topic 6: yw,thx,want,high,billing,hi,ty,phone,hello,ok
Topic 7: want,like,international,plans,hotspot,cost,change,unlimited,data,plan
Topic 8: phone,uverse,connection,slow,cable,home,speed,tv,working,internet
Topic 9: activation,questions,charged,late,question,like,contact,explanation,charge,fee
Topic 10: hi,payments,chat,talk,need,hello,billing,speak,live,agent
Topic 11: directv,account,app,verse,channel,watch,package,channels,direct,tv
Topic 12: access,remove,cancelled,xxxxxxxxx,want,email,need,wireless,credit,account
Topic 13: paying,paid,months,charged,00,im,told,credit,pay,month
Topic 14: amber,ashley,spoke,hello,payments,mobile,talk,billing,speak,representative
Topic 15: need,payments,set,paid,change,date,make,pay,arrangement,payment
Topic 16: chat,wifi,person,working,speak,password,tech,need,technical,support
Topic 17: device,added,watch,like,new,need,lines,remove,add,line
Topic 18: oh,pay,thats,im,hello,welcome,yes,account,thanks,okay
Topic 19: dollars,code,change,address,yes,thanks,need,number,xxxx,xxx
```

Topic Modelling

Categorization of these topics (one-time activity) makes topic-wise indexing of all current and future chats possible
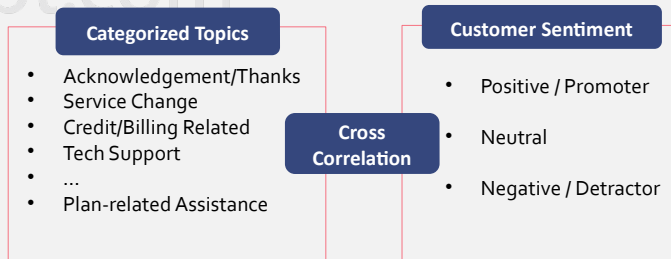
```
topic_dict = {0: "Acknowledgment/Thanks", 1:"Service Change", 2:"Product eSupport/Self-service", 3:"CTN/Sim/Mobile Queries",
4:"Credit Issues", 5:"Cancellation/Package Removal", 6:"Billing/ Bill Amount", 7:"Plan/Line Details",
8: "UVerse/IPTV Related", 9:"Charges/Fee Queries", 10:"Introduction/Request for Live Agent Transfer",
11:"DTV/UVerse Package/Contract related", 12:"Remove/Cancel Service", 13:"Credit/Billing Related",
14:"Bill/Payments/LiveAgent Request", 15:"Payment Arrangement", 16:"Tech Support", 17:"Add/Remove Line",
18:"Conversation Utterances", 19:"Plan-related Assistance"}
```

Topic Categorization

Perform Cross Correlation between "Topics" and "CSAT Scores" → **Fine-tune the operating procedures**

Key Considerations

- Find the optimal level of topics based on the existing corpus of words.

- Identify key topics discussed by customers (for each CSAT level) – ML techniques such as **Non-negative Matrix Factorization(NNMF) or Latent Dirichlet Allocation (LDA)** can be used of this exercise.

- Cross correlate identified topics with CSAT scores to find which items are strongly correlated with customer's positive sentiments and which are with the negative sentiments.

Categorized Topics
- Acknowledgement/Thanks
- Service Change
- Credit/Billing Related
- Tech Support
- ...
- Plan-related Assistance

Cross Correlation

Customer Sentiment
- Positive / Promoter
- Neutral
- Negative / Detractor

Example: If results of correlation says that "Credit/Billing Related" service is highly correlated with customers negative sentiment, then the operating procedure of that service needs to be fine-tuned.

Prodapt

Retrain Customer Service Operating Procedures - Word Clouds of Promoters, Neutrals and Detractors can be compared for Predicted Vs Actual or Against Each Other

1 2 3

Improve Customer Engagement & Overall CSAT Ratings

- Predicted promoters are set of people satisfied with service but don't care to give rating because of **lower customer engagement**.

- Identify such topics that lead to lower engagement by comparing word cloud for Actual vs Predicted promoters.

Sample Wordcloud for Actual Promoters

Sample Wordcloud for Predicted Promoters
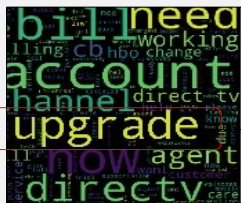
Customers are satisfied with **"upgrade"** related service, but they don't submit a rating (lack of time, lack of interest)

Enhance the customer engagement by changing certain operating procedure for upgrade service.
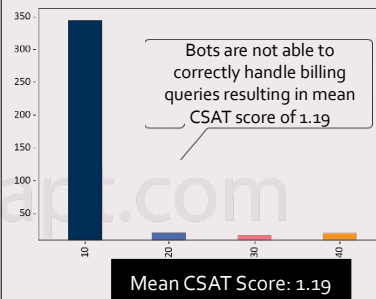E.g.: Attractive UI Design, Simple & short feedback question etc.

- Similarly, compare and draw inferences from other word clouds and make necessary changes required to promote detractors to neutrals and neutrals to promoters.

Analysis across intents and agent-groups, for both humans & VA (bots)

Intents Handled by VA (bots)

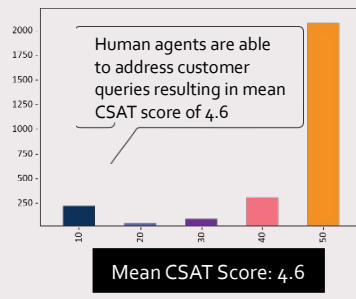Sample CSAT Distribution for intent: Billing_Related_Query

Bots are not able to correctly handle billing queries resulting in mean CSAT score of 1.19

Mean CSAT Score: 1.19

Sample Word Cloud for: Billing_Related_Query

Agent Group Handled by Human

Sample CSAT Distribution for Group: Technical_Support_Group

Human agents are able to address customer queries resulting in mean CSAT score of 4.6
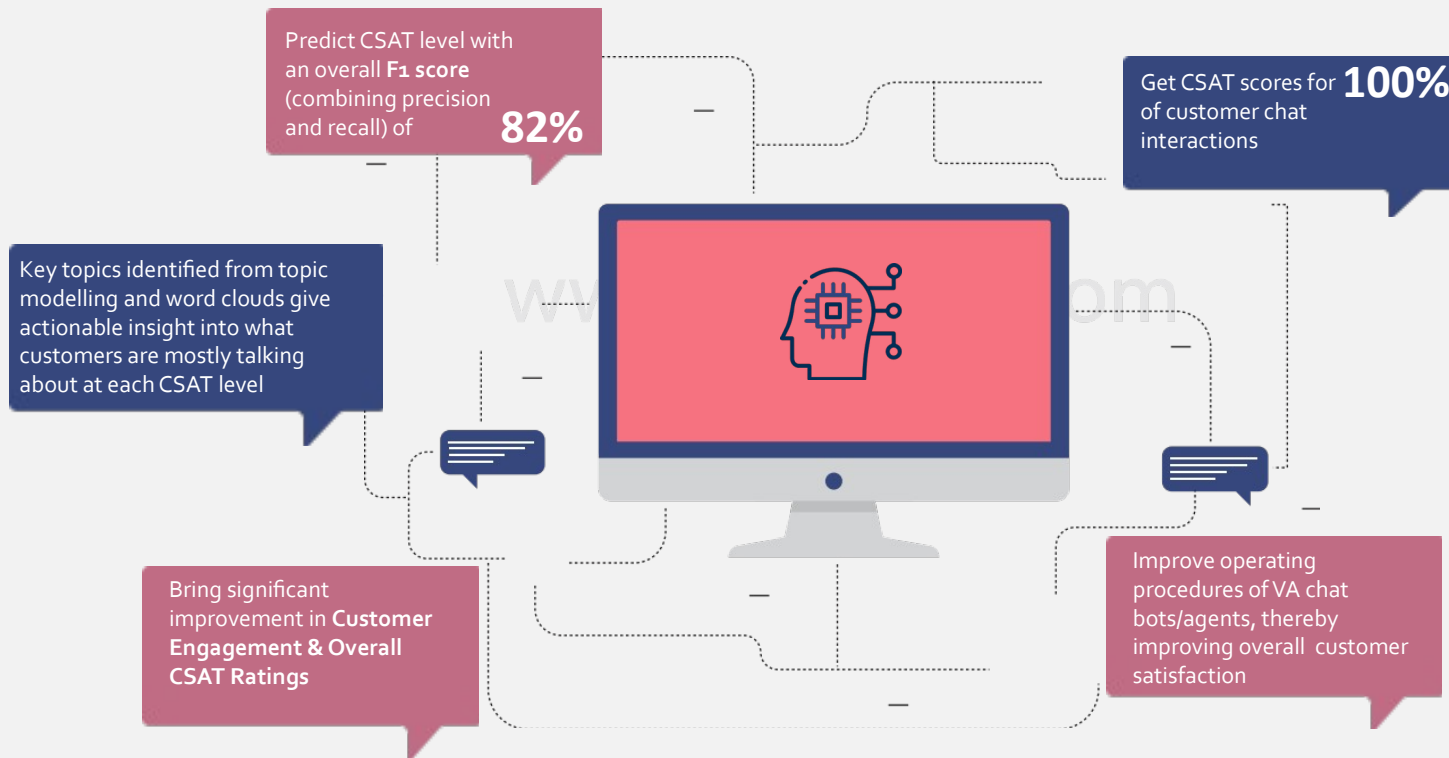
Mean CSAT Score: 4.6

Sample Word Cloud for: Technical_Support_Group

- Identify topics where humans outperform bots and vice versa.
- Identify key characteristics of high-performing groups and intents.

Key Takeaways

Leveraging a machine learning-based approach for CSAT prediction can have following benefits

Predict CSAT level with an overall **F1 score** (combining precision and recall) of **82%**

Get CSAT scores for **100%** of customer chat interactions

Key topics identified from topic modelling and word clouds give actionable insight into what customers are mostly talking about at each CSAT level

Bring significant improvement in **Customer Engagement & Overall CSAT Ratings**

Improve operating procedures of VA chat bots/agents, thereby improving overall customer satisfaction

Prodapt

Get in touch

USA

Prodapt North America
Tualatin: 7565 SW Mohawk St.,
Phone: +1 503 636 3737

Dallas: 1333, Corporate Dr., Suite 101, Irving
Phone: +1 972 201 9009

New York: 1 Bridge Street, Irvington
Phone: +1 646 403 8161

CANADA

Prodapt Canada Inc.
Vancouver: 777, Hornby Street,
Suite 600, BC V6Z 1S4
Phone: +1 503 210 0107

UK

Prodapt (UK) Limited
Reading: Davidson House,
The Forbury, RG1 3EU
Phone: +44 (0) 11 8900 1068

EUROPE

Prodapt Solutions Europe &
Prodapt Consulting BV
Rijswijk: De Bruyn Kopsstraat 14
Phone: +31 (0) 70 4140722

Prodapt Germany GmbH
Münich: Brienner Straße, 80333
Phone: +31 (0) 70 4140722

Prodapt Switzerland GmbH
Zürich: Mühlebachstrasse 54,
8008 Zürich

SOUTH AFRICA

Prodapt SA (Pty) Ltd.
Johannesburg: No. 3,
3rd Avenue, Rivonia
Phone: +27 (0) 11 259 4000

INDIA

Prodapt Solutions Pvt. Ltd.
Chennai: Prince Infocity II, OMR
Phone: +91 44 4903 3000

"Chennai One" SEZ, Thoraipakkam
Phone: +91 44 4230 2300

IIT Madras Research Park II,
3rd floor, Kanagam Road, Taramani
Phone: +91 44 4903 3020

Bangalore: "CareerNet Campus"
2nd floor, No. 53, Devarabisana Halli,
Phone: +91 80 4655 7008

THANK YOU!

Prodapt powering global telecom