### Customer Review Score (CRS) as an alternative to Net Promoter Score (NPS) – A Sentiment Analysis Based Approach

**Why you shouldn't rely on NPS alone?** In the services industry Net promoter Score (NPS) is a widely used metric for measuring the customer satisfaction based on whether he/she is willing to recommend the store/service to anyone or not. NPS is based on a direct question: *How likely is it that you would recommend our company/product/service to a friend or colleague?* The response to this question is a number from 1 to 10, with a score of 9-10 is considered as **Loyal Customers** and a score of 0-6 is considered as **Unhappy Customers**. NPS has its own criticisms. Russel Hayes in 2008 recommended that the *likelihood to recommend* is not a measure of business growth or customer loyalty, and it does not provide any more information on what is good and what is wrong, why exactly the customer is happy and why exactly the customers is not happy.

**Using customer generated data for measuring customer feedback**

With the availability of so many platforms across the internet for customers to share, connect and interact with the brands and with individuals, there is a plethora of information getting piled up on a lot of services-review sites. **Yelp** is a very good example of restaurant reviews and ratings. A lot of reviews of Hotels, Restaurant and resorts can be found on **Tripadvisor**. There are specific sites for people to write a review on specific services, and different platforms are popular in different parts of the world. For example, in USA **Yelp** is fairly popular for restaurant reviews, while in India this role is played by **Zomato**. Also, **Tripadvisor** has a bias built in for travelers as it is used very widely by the travelling community. All these sites have hundreds of reviews for a huge number of businesses.

So are these reviews helpful? Yes, already a lot of businesses are tracking these reviews to understand feedback from customers, and a lot of customers are using these reviews to select the services or the product. But, there are cases where thousands of customers are writing the reviews about a single restaurant, and there is no way the business owner can go through all of them to find about the overall good things and bad things about his/her business without investing a lot of time and resources. The **NPS** only uses one metric, but there is a huge information already stored in these reviews which are richer in content, and can tell you exactly why the customer is happy and why the customer is unhappy - *Something the NPS can't tell but the CRS can.*

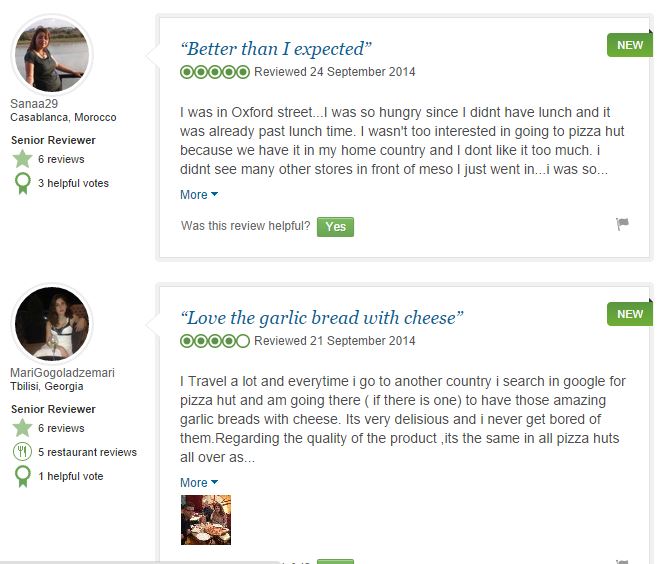
***So why is CRS not used widely?***  
The answer is because of the difficulty to go through thousands of text-based reviews. NPS will give you just one number at the end, which you can get out on a calculator. The problem with the reviews is the difficulty in analyzing them and **extracting** information, understanding customer *emotion* and the *polarity* of the review (positive/negative). This is something which only a human mind can do, and is not a job for machines.  
With the advancement in data sciences and Natural Language Processing, this is now possible to some extent. Of course the computer cannot find out the difference between a happy "I am Fine" and a sarcastic "I am Fine", but it can be **75%-80% accurate** in predicting the sentiment in a sentence or a paragraph as compared to Human Brain which does it at **100%** accuracy. This is good enough for a predictor model, and skimming through thousands of documents/reviews in a matter of seconds.

### An Example of using CRS - Analyzing Reviews of Pizza Hut - Covent Garden, London, UK on Tripadvisor.com

**Overview**

Data Source: [Tripadvisor.com](http://www.tripadvisor.in/Restaurant_Review-g186338-d1010274-Reviews-Pizza_Hut-London_England.html)  
Number of Reviews: 410 as on 23rd September,2014  
Reviews Analyzed : 185 Reviews written in English (Rest of the reviews are in other languages)

Disclaimer : All information (reviews, customer name & location) is already present in public domain over the internet on review sites, no personal data or data from any other source has been used. We do not claim the accuracy of Name/Location information. The reviews are personal opinion of the writers and not of Tripadvisor LLC or Capillary Technologies Pvt. Ltd.

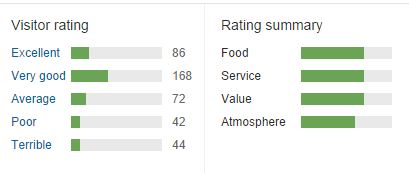
***How the Data looks like before it is collected***  


**Data Collection**

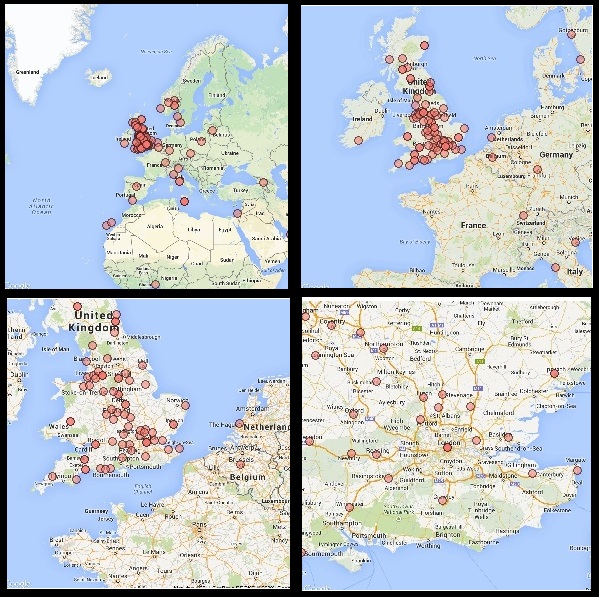
The analysis starts with user-generated data collection, which is obtained by scraping the reviews of the particular store from the website [Tripadvisor.com](http://www.tripadvisor.in/Restaurant_Review-g186338-d1010274-Reviews-Pizza_Hut-London_England.html). The reviews are collected along with the information of reviewers (Name, location and reviewer ranking according to Tripadvisor). A lot of other information about the users can be obtained by scraping through their profile pages but we are skipping this for simplicity.

**Analysis**

The website provides an overall rating and ranking based on how the customers fill them and quantifies them according to number of reviewers



***Where are the reviewers located***  
The customers are from all parts of the world (this is a bias built in since the website is a Treveller-focussed website). An analysis of locations can show where the reviewers reside.



Most of the customers are from the UK, and within UK they are from all over the country. We do not have their address or postal codes which can be used to identify which are the customer who live within a close distance to the store. All the reviewers do not share their location, hence this chart only marks the reviewers who mention their city in their profile.

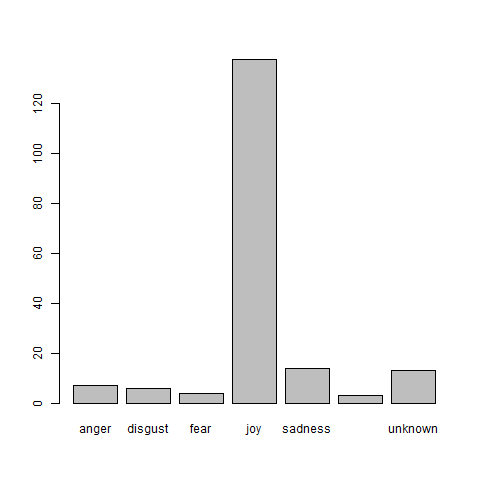
***What are the customers talking about***  
An initial analysis of all the reviews throw up the following Wordcloud. The customers are talking about a lot of things, but from this Wordcloud we can't say if the customer who is talking about 'service' has a positive feedback or a negative feedback.



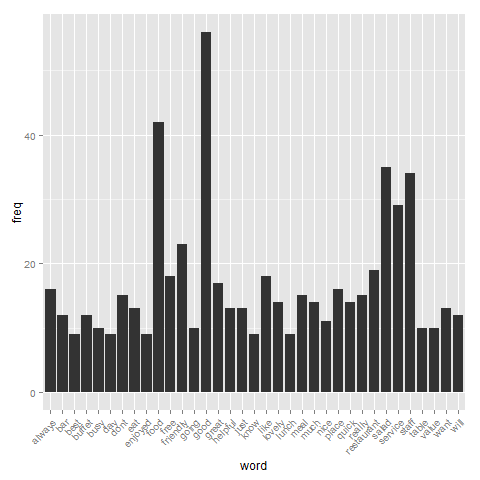
Most of the customers are talking about 'Food' and 'Staff'. A deeper analysis of the reviews throw up much insightful results than the overall picture.

### Classification of Emotions and Polarity of review

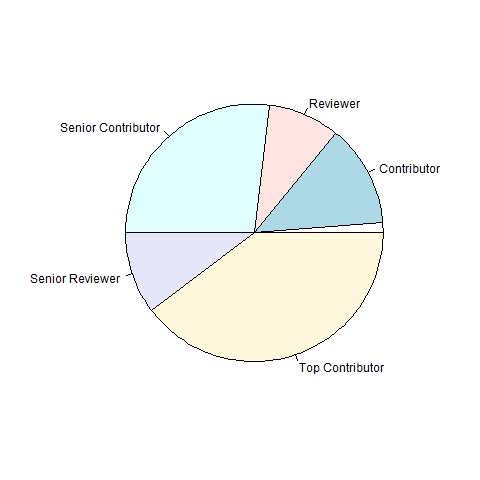
The emotions and polarity are calculated based on the **Naive Bayes Text Classification Algorithm** which classifies the emotions as "joy", "anger", "sadness", "disgust", "fear" etc. The polarity of the text (positive/negative/neutral) is assessed based on the degree of emotion by assigning a numeric to it (ranging from -3 to +3) and then summing them up over the sentence to account for contradicting statements (example: "The food is excellent but the service is pathetic").

***Emotions and polarity classification from the reviews***  
 

The charts indicate that the emotions are predominantly of 'Joy' and most of the reviews are 'Positive'. However, it'll be good to understand what the customers are talking about most when they are giving a negative review or the negative review

***Positive Reviews***  
 

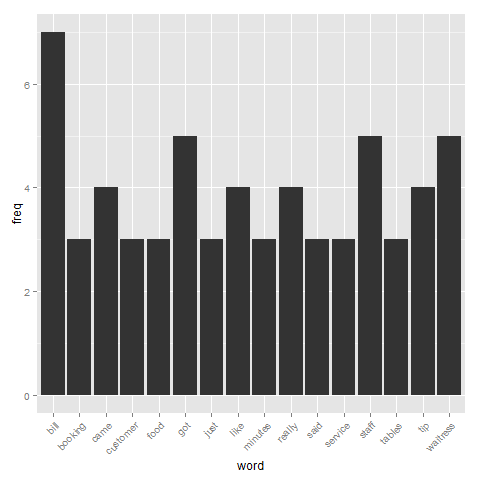
This analysis shows that the customers are most happy about food, staff, salad and services and have written a positive review about this store mentioning these.

  
The chart shows that 'Top Contributors' and 'Senior Contributors' form a bulk of these positive reviewers which is a good thing, as people rely on the reviews given by these people and they have more influence than an one-off rating provider.

Positive Sentiments – Example :

*“Enjoyed our lunch here, was actually plesently suprised as being in London we wasn't expecting much with the service but to our suprise it was very quick and good! The staff were welcoming and food was very fresh! would recommend this if your after some food in London and dint want long waits.”*

*“Whilst visiting London recently we were looking for somewhere to eat and relax in between sightseeing and shopping.. Pizza Hut Piccadilly was perfectly located, just off Piccadilly Circus. Upon our arrival at the restaurant we were greeted by the host who was baying back the crowds at the front door, and we waited a very short time before being shown to our table on the upper floor of the hotel. We decided to have a real feast and get bottomless drinks and the lunch buffet and all under £20 for two people! There was a number of different pizzas n the buffet as well as two types of pasta and a wide variety of salad stuff! Would definitely recommend and return again.”*

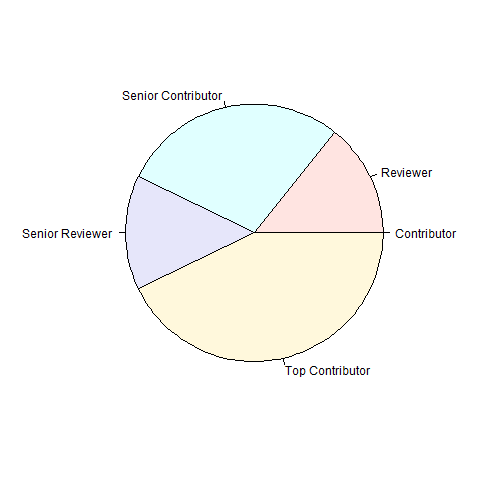
***Negative Reviews***  
 

The word "Bill" occurs most frequently in the negative reviews indicating that the negative reviewers are most unhappy by the "Bill". The store needs to look into this issue, maybe by improving turnaround time for providing Bill to the customers, or checking the accuracy of Bills. This is followed by "Staff" and "Waitress", indicating customers have issues with the staff of the store.

A further drill down is possible which makes use of **Word Correlation** which identifies how frequently the words come together in a text. For example, "Bill" could be having a strong correlation with "Late" or "Inaccurate", which can be used to identify the exact issue.

This analysis does not include words correlation since the data size is a bit small, and a larger data set will be used to identify words correlation correctly. Since the data is small, the Store can directly look into these negative reviews to see why these customers are unhappy.

*Reviwers*

  
Most of these negative reviews are written by the 'Top Contributors' and the 'Senior Contributors' which is a cause of worry, as these categories of Reviewers are most influential. The store can answer to these reviewers on the website itself so that their concerns are addressed

*Negative Sentiments : Examples*

*“Went in told to sit down not shown to table or welcomed in , 15 minutes before any service , food took over 45 minutes, complained to manager she really didn't care she just said "what do you want?? " pizzas were cold didn't pay for pizzas got bill wrong !! another 10 minutes later asked for mangers name as the rest of the family had to leave again her attitude was awful !! Ruined the evening DONT GO!!”*

*“Pizza was horrible , tasted vile and stodgy. I complained and got knowhere. I find pizza Huts attitude towards a complaint from a customer very arrogant. Pizza Hut need to realise that they're not the main player in pizza anymore and cannot afford to be so arrogant. Pizza express, papa johns and even Dominos are better pizzas.”*

**Accuracy of Results**  
A major question about Automated Text Mining is the accuracy of results. As mentioned earlier, algorithm based text classification does not claim to be (and cannot be) 100% accurate as compared to Human Interpretation which is a bit subjective. However, these algorithms provide an accuracy level of 70-80% which is fairly acceptable for an automated analysis.

In this case, since the data is pretty small, it was possible to go through the classification manually, and compare the algorithmic sentiment classification with the manually interpreted classification by reading the comments. The negative reviews are **81% accurate** and the positive reviews are **88% accurate**. The overall accuracy will go down a bit since ambiguous reviews ("Pizza is good but Garlic bread is not good") will be flagged as neutral. Still a good amount of accuracy for an Automated Model for positive/negative classification!

### *What else can be done with Customer Review Score*

The CRS has its own positive and negative sides. You need to have a good amount of data to analyze, and it should be generalized (not collected from a single source). A very good way to enrich the data is to include the feedback from other social media channels like Twitter. If the reviews are available on any medium, **CRS** is a much more analytically advanced way to measure the customer satisfaction score rather than the single-metric **NPS**. The results obtained from CRS can be further broken down by store, by country, by user age group (if we have the information), or any other data available.

This has huge potential not only for Restaurants, but also for businesses where customer satisfaction is a very important part of business. Hotels and Resorts (for which a lot of customers write reviews) can understand more about customer feedback and take necessary action based on insights. Airlines can understand the key issues that the customers face based on different locations/airports or type of aircraft. Wherever there is user generated data, the CRS can be a good tool to understand your customers.