

Mini Project 3 - TripAdvisor Restaurant Reviews (NLP)

Marta Huhoja,
January 2023

01

TripAdvisor Data

02

Text EDA

03

Modelling

04

Results and Next Steps

About the Data

- part of [six new TripAdvisor datasets](#) from the restaurants of six different cities:
London, New York, New Delhi, Paris, **Barcelona** and Madrid
- insights for forecasting, proposing tasks related to recommender systems, sentiment analysis, text-based personalisation or text summarisation
- used the data to predict if review is positive or negative

| | | | | |
|--------|---------|---|-------|---------|
| Rows | Columns | → | Rows | Columns |
| 416356 | 13 | | 10000 | 10 |

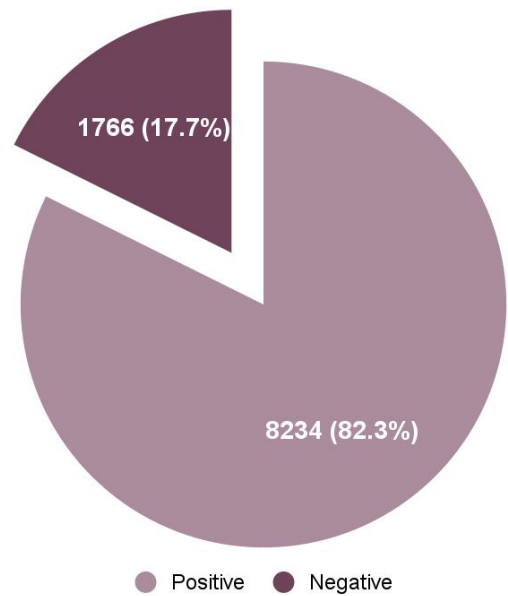
- 2352 unique restaurants (Cerveceria Catalana and Ciudad Condal 120+)
- 9959 unique reviewers

| | restaurant_name | rating_review | sample | review_id | title_review | review_preview | review_full | date | city | author_id |
|--------|-----------------|---------------|----------|------------------|------------------------------|---|---|------------|---------------------|------------|
| 399811 | Ciudad_Condal | 5 | Positive | review_303189870 | best tapas in town | always a must for an early afternoon lunch.ver... | always a must for an early afternoon lunch. ve... | 2015-08-25 | Barcelona_Catalonia | UID_115016 |
| 2496 | Rao | 5 | Positive | review_637072072 | Excellent diner experience ! | Nice decoration, nice service, great Tapas and... | Nice decoration, nice service, great Tapas and... | 2018-12-02 | Barcelona_Catalonia | UID_2413 |

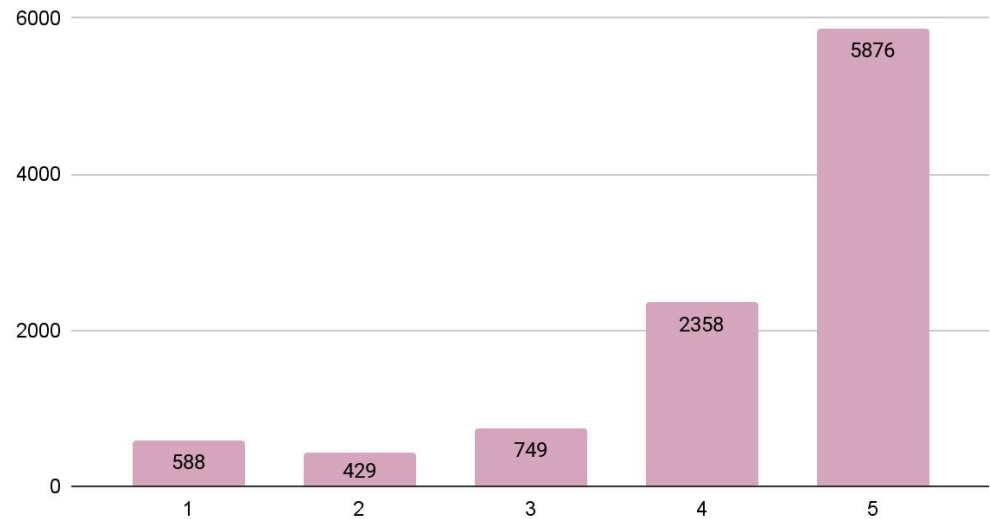
- **restaurant_name:** categorical (string), name of the restaurant matching the review
- **rating_review:** numerical (integer), review score in the range 1-5
- **sample:** categorical (string), indicating “positive” sample for scores [4-5] and “negative” for scores [1-3]
- **review_id:** categorical (string), univocal and internal identifier of the review (review_XXXXXXXXXX)
- **title_review:** text, review title
- **review_preview:** text, preview of the review, truncated in the website when the text is very long
- **review_full:** text, complete review
- **date:** timestamp, publication date of the review in the format (day, month, year)
- **city:** categorical (string), city of the restaurant which the review was written for
- **author_id:** categorical (string), univocal, incremental and anonymous identifier of the user (UID_XXXXXXXXXX)

Reviews

Reviews



Review rating



Working with Text

Text based features:

- char_count - number of characters in a review
 - sent_count - number of sentences in a review
 - word_count - number of words in a review
-
- Maximum number of characters in a review is 6180 and minimum 13 with majority being around 380 characters.
 - Longest review has 61 sentence, shortest just 1 and mean is around 5 sentences per review.
 - Most reviews have around 69 words; minimum number of words in a review is 2 and maximum 1109.

Longest (positive) comment

The meal began when three waiters placed in front of us on large golden plates four single fork sized pieces of food arranged in an eating order so as to maximize the play of flavors on our tongue. The first portion was a truffle, which brought an earthy, woody taste to my palette that was stronger and more fleeting than usual as the truffles had been ground and compacted together so as to make it quickly disappear upon hitting my tongue. The next tasting portion made me feel if the chef was inviting us to not only pay attention to flavors of the delicious foods we were eating but to the order as well for following the truffle was the very animal trained to find it - pig. Smiling at this culinary pun, I sipped the delicious cava my fiancé had chose and then put the marble sized Iberian ham croquette in my mouth. I was not able to bite into the croquette as upon putting it into my mouth and putting the slightest amount of pressure on the fried brown breading, the ham quite literally melted and washed over my tongue bringing with it the nutty taste of the Iberian ham. Following this was a thin slice of candied watermelon rind. After the sweetness of the rind, we each opened individually wrapped coffee beans to cleanse the palette before our next course. Gazpacho was then served. I was reluctant at first as my previous encounters with the quintessential Spanish dish less had been either too salty or too watery. However this gazpacho had neither of these excesses and also had in it some fresh peach juice, which gave it an exceptionally refreshing flavor. I believe that it also had a few drops of rose water, but whether this came from that or the edible flower in the center of the soup was impossible for me to tell. Following these preparatory dishes, our ordered appetizers started to arrive. There are few words that I can use to describe the visually and olfactory beauty of the fresh leaves salad that I had. Its appearance made me think that it was made by the same mad and magical mind that designed Neuschwanstein. However rather than making a magical castle for a king and his bride, this dish was a magical park for their children to play in. The fresh flowers were in a semi-gelatinous base that accented their colors and had a light taste. Mixed in all of this were small pieces of lobster. I ate with some reluctance as destroying such beauty seemed criminal, but the taste was so good that such feelings quickly subsided. A friend ordered the roasted, chopped scallops topped with Iranian caviar and a small side of artichokes, raw and creamy celery salad puree. I managed to obtain a small bite of such a delicacy, then watched it disappear with the only interruption being several sips of cava. The slices of Iberian ham cured for 36 months another friend ordered covered a large plate and was too much and too heavy for him to eat alone so all three of us enjoyed it and paired it with fresh bread the wait staff offered. At this point a second bottle of cava came to the table. We had several minutes to discuss the delicious food and our upcoming wedding plans when our entrée's arrived. My male friend, like myself, orders food based on what is typical to the region or generally unavailable in the United States. As such, he ordered the breast of pigeon with durum gravy, mashed potatoes accented with truffles and a surprise addition of baby carrots. The meat was dark in color from flash roasting that kept in all the succulent juices. It was without any of the gaminess that we thought it would have and a topic for discussion as to the general silliness of food discrimination based upon inherited notions of what is appropriate. My female friend, lover of steaks that she is, ordered a filet, potato and bacon mille-feuille with foie gras sauce. I find it hard to go into detail when describing a steak that is cooked to perfection and accompanied with perfectly paired foods as I find that only the taste itself is a sufficient indicator of superb quality. Saying that the center was reddish pink just as she ordered is no substitute for a bite of it. All I shall say is that after three months of our living in Barcelona, she finally gave me the notice that she had a steak that was worthy of the name. I ordered pigs trotters stuffed with Catalan black pudding and a side of mushroom and cheese toast. I thought it a little pricey considering the cut of meat, but I wanted to try a traditional Catalan dish and soon discovered that my adventurous choice well rewarded. At this point in my life there are very few things that I can eat and say the taste, texture or flavor are completely novel to me. My parents served me a wide variety of international cuisine since a baby and I've kept this habit since leaving their nest when cooking at home as well as when traveling to the two dozen or so countries that I've been. That said, the flavor and taste of these trotters was a completely new to me and something that I enjoyed immensely. Inside the pork trotters was the "pudding", a thick gelatinous sauce of a consistency slightly thicker than marrow and filled with pieces of pork. It was somewhat thick, being basically a flavored gelatin shell filled with metal filled gelatin, but the smallest dab of the sauce garnishing the plate cut through it without overwhelming it. It was precisely what I had hoped to get from going to the restaurant, a modern take on classic Catalan cuisine, and I was richly rewarded for my choice. By this time each of us were so full by the appetizers, the cava and main courses that we foreswore dessert. Yet sure enough shortly after the amazingly prompt and courteous wait staff cleared our plates they placed a tray down in the center of our table that was made from silver forks arranged in such a way as to hold small shot glasses of four different types of deserts. At the sight of this presentation and hearing their description our initial reluctance and treated ourselves to yet another round of delicious indulgences. Stuffed and awed after such a spectacular meal, we all left feeling as if we had finally had some delicious Spanish cuisine that was worthy of writing about

Shortest (positive) comment

VS

Nice food, beautiful terrace,
but a bore

Longest (negative) comment

My husband and I came here as it was across the road from our hotel (h10 port vell). It was his 30th birthday and as the restaurant we had originally planned to visit was closed we decided to try some good old fashioned Italian food. On entry to the almost empty restaurant, we were pointed to a table by a young waiter who then took our drinks order and ignored our request to see an English version of the menu. We could already sense from the complete lack of atmosphere that we may be in for disappointment. When we managed to catch the attention of another waiter we asked again for an English menu and were presented with one. My husband decided on steak and I chose a taglierine with garlic, chilli and prawns. First off, I asked if I could have a side order of salad and was told no, fair enough, this wasn't on the menu but in this day and age you would think they could manage it. The waiter then walked away without asking how my husband would like his steak cooked. We called him back and tried to ask if it was perhaps cooked a certain way in explanation to why he hadn't asked and were fobbed off with the old 'I don't understand, let me get someone else'. After a few minutes we were approached by another waiter and asked the question to which he said 'well how do you want it done?' So my husband asked for medium rare. 15 minutes or so later our food arrived where things went from bad to worse. My husbands food was cold and the steak was cooked well done. My 'taglierine' - which I will admit I had never heard of nor eaten before - was almost exactly like super noodles (most English should understand) with shrimp - not prawns. It was completely tasteless, stodgy and not very warm either. After one unsatisfactory bite we called the waiter back over and said 'I'm sorry but the food is really bad' to which was met with an absolutely disgusted face and 'what is wrong with it?'. We explained that it was neither warm nor appetising and after being bombarded with pure rudeness I was forced to say that it was not satisfactory and that we would not be eating or paying for it but that he could bring the bill for our drinks. He said he would have to consult his manager and left our table with the food in front of us to do so. When he returned he wanted further explanation as to why we were unhappy with the meals. We explained again - the menu said prawns 'yes, shrimp, prawns'...those are two very different things! Cold temperature/ poor taste etc and after many screwed up facial expressions he picked up my husbands food and said 'this I can understand but yours - that's just how it comes, there is nothing wrong with it, that's just pasta'. Obviously at this point I was quite frustrated/ insulted and explained that I had eaten pasta once or twice before in my life and had never experienced this lack of standard. Again, he screwed up his face and told me how wrong I was whilst clearing the table and muttering under his breath. My husband, who's birthday was looking thoroughly bleak at this point, was so stressed out by the situation that I had to send him outside whilst I waited to pay for the drinks. I am absolutely outraged by the horrendous service. If the food was the same but the complaint had been better received I may not have been compelled to leave such a scathing review but unfortunately, this is now my only outlet for expressing my disappointment with such a terrible experience. If I could think of one positive point I would be more than happy to express it but unfortunately, there were no redeeming moments to my visit - other than the favour they did us in a respect because we went on to have a fantastic meal just down the road.

VS

Shortest (negative) comment

Great paella!

Processing Text Data

Reviews has been:

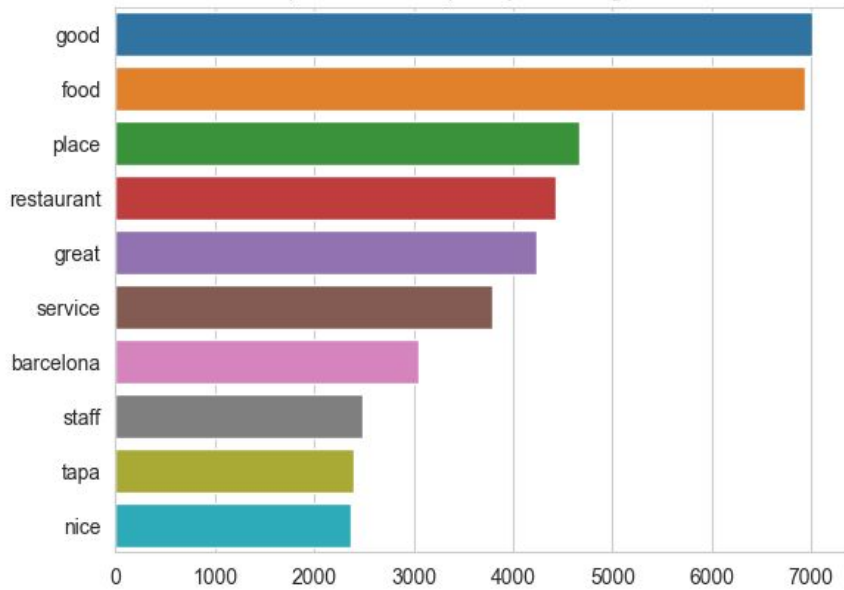
- tokenized
- removed stop words
- removed punctuations
- lemmatized



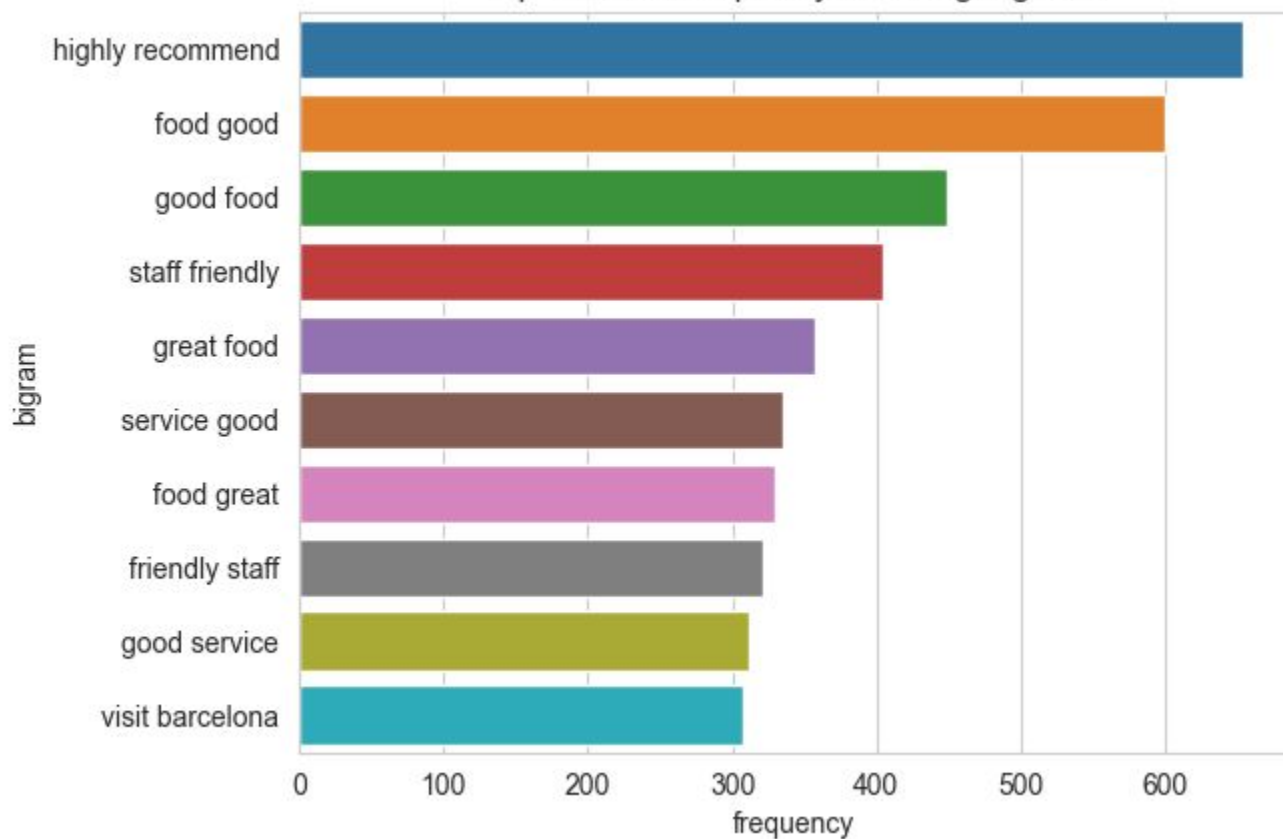
Processed version has been used
for machine learning models.

british plan week gastronomic visit barcelona eat michelin star restaurant trip choose 1 angle
2 disfrutar 3 abac order week book month advance line think good way review post
comparison choose taste menu wine pairing price person angle 141 euros disfrutar 220 euros
abac 283 euro angle 1 modern setting greeting arrival warm friendly show seat lounge offer
drink immediately invite bar sample small amuse bouche create serve friendly skilled chef
guide upstairs restaurant maître d. room intimate beautifully decorate staff warm friendly
waiter sommelier introduce staff work ensure timely rate service make service feel pressurise
total course serve glass wine dish come complete prepare table angle menu line highlight
include pork fritter oloroso broth veal cap pita rice oyster duck foie gra mole sauce wine
range sherry saké lovely bellmont red spain dish wine superb staff excellent thoroughly
fantastic faultless evening delicious imaginative food great value highly recommend disfrutar
2 publish menu advance dish surprise menu food wine give end souvenir course wine staff
incredibly friendly efficient introduce personally chef speak arrival show directly table pass
kitchen en route restaurant brightly light sense busyness 20 dish come impeccable timing
include serve wine pair competent sommelier dish serve complete prepare table fabulous
taste visual aspect level service superb wine perfectly match food range sherries saké white
red finally desert wine dish playful surprising look like savory walnut candy gazpacho
sandwich deconstruct ceviche squab foie overall fantastic star experience highly recommend
abac 3 restaurant set lovely garden modern design feeling menu divide part terrace kitchen
main restaurant arrival immediately greet show table terrace leave waiter deliver consecutive
amuse bouche offer drink waiter walk forth introduce feel service cold detached table ask
waiter smoke say yes restaurant allow smoking nowadays luckily table free move food good
take kitchen amuse bouche serve time glass cava stand counter feel service distant engage
kitchen staff serve counter acknowledge eye contact food good stage restaurant contrast
staff warm friendly maître d sommelier waiter food fantastic innovative great variety
ingredient concentrated taste wine pair perfect wide variety wine sherry final desert course
delicious end beautiful display petit invite terrace choose stay table smoking issue overall
great experience dampen terrace kitchen service issue souvenir menu printer break promise
email copy day arrive recommend abac main restaurant food service definitely star initial
terrace kitchen experience need improve overall deserve star rating angle exceptionally good
value disfrutar fantastic service unusual dish abac resolve issue ultimate gastronomic
experience

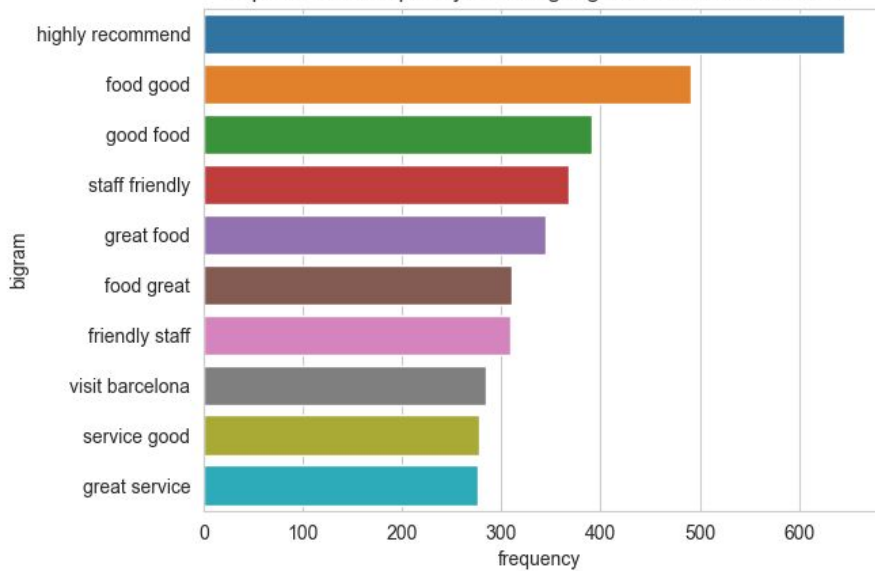
Top 10 Most Frequently Occuring Words



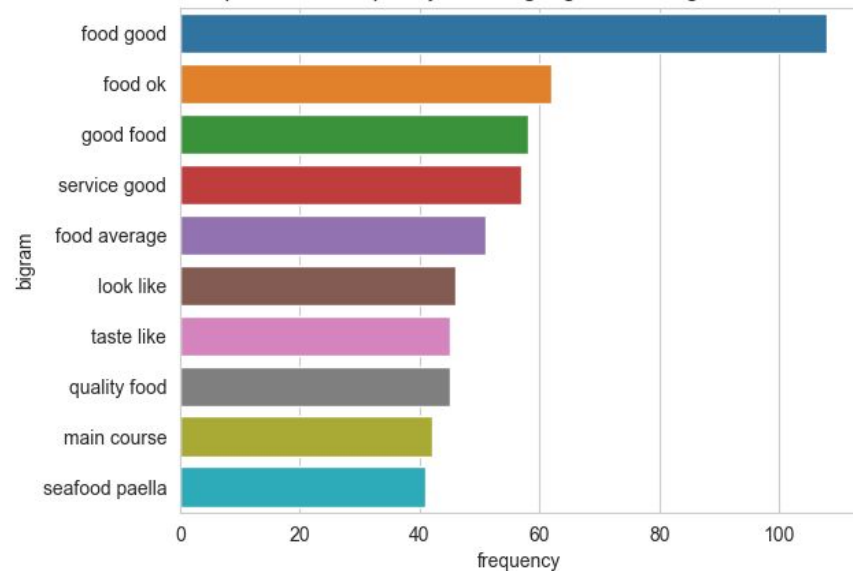
Top 10 Most Frequently Occuring Bigrams



Top 10 Most Frequently Occuring Bigrams in Positive Reviews



Top 10 Most Frequently Occuring Bigrams in Negative Reviews



Calculating polarity of the text with TextBlob

**Sangria was so frutty,
cream catalana delicious
and steak was perfect!
Service is absolutely
perfect! Gracias!!!!**

Polarity = 1

**We went to this
restaurant following
some businessmen that
we were also looking
where to lunch. Lunch
menu is €11.50, I had
empedrado de bacalao
and canelones de
vegetales, lunch price
includes drinks, bread
and dessert.**

Polarity = 0

**Terrible service,
terrible food, dishes
are charged twice as
'they come as 2
servings per plate'.
AVOID!!**

Polarity = -1

Machine Learning Models

- Used **TF-IDF** and **Count Vectorizer** methods
- CountVectorizer simply counts the number of times a word appears in a document (using a bag-of-words approach), while TF-IDF Vectorizer takes into account not only how many times a word appears in a document but also how important that word is to the whole text
- Both approaches were tried with SVM and Logistic Regression

CountVectorizer LR

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.84 | 0.68 | 0.75 | 391 |
| 1 | 0.93 | 0.97 | 0.95 | 1609 |
| accuracy | | | 0.91 | 2000 |
| macro avg | 0.88 | 0.82 | 0.85 | 2000 |
| weighted avg | 0.91 | 0.91 | 0.91 | 2000 |

TF-IDF LR

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.91 | 0.52 | 0.66 | 391 |
| 1 | 0.89 | 0.99 | 0.94 | 1609 |
| accuracy | | | 0.90 | 2000 |
| macro avg | 0.90 | 0.75 | 0.80 | 2000 |
| weighted avg | 0.90 | 0.90 | 0.89 | 2000 |

CountVectorizer SVC

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.76 | 0.70 | 0.73 | 391 |
| 1 | 0.93 | 0.95 | 0.94 | 1609 |
| accuracy | | | 0.90 | 2000 |
| macro avg | 0.85 | 0.82 | 0.83 | 2000 |
| weighted avg | 0.90 | 0.90 | 0.90 | 2000 |

TF-IDF SVC

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.86 | 0.65 | 0.75 | 391 |
| 1 | 0.92 | 0.98 | 0.95 | 1609 |
| accuracy | | | 0.91 | 2000 |
| macro avg | 0.89 | 0.81 | 0.85 | 2000 |
| weighted avg | 0.91 | 0.91 | 0.91 | 2000 |

Conclusion and next steps

- Logistic Regression method where CountVectorizer has been applied and SVC where TF-IDF Vectorizer has been applied both has scored **accuracy 91%** on test data
- we can use this model to detect (negative) reviews business want to respond to
- if trained on data for specific restaurant, it can be used to detect areas that have to be improved
- further actions could include analysing comments of specific users to see if they are more prone to leave positive/negative comments (polarity) or to detect if they always write very long/short reviews
- more sophisticated model could be build to use reviews for recommendations to specific users

¡Gracias!