COT-6930 Natural Language Processing, Spring 2019
Assignment 2 - Sentiment analysis using TextBlob
Christian Garbin

Assignment

In this assignment we were asked to "*... use Textblob to label sentiment (positive or negative) for the given dataset...The dataset contains 40 sentences and each line is one sentence.*"

Report structure

This report is structured as follows:

1. The first section below explains the TextBlob features used for the report.
2. The next section after that shows the results and the analysis.
3. The last section shows the source code.

TextBlob features for sentiment analysis

TextBlob has two sentiment analyzers:

- `PatternAnalyzer`: based on the pattern library from the University of Antwerpen. It returns a polarity value between +1 (very positive) and -1 (very negative), and a subjectivity value between 0 (very objective) and 1 (very subjective).
- `NaiveBayesAnalyzer`: a Naive Bayes analyzer from NLTK, trained on a movie review corpus. It returns the classification as a string (positive or negative), and the probability of the positive and negative classes (the negative probability is simply "1 - positive probability").

For this report we used both analyzers, comparing their results.

Results and analysis

The table below shows the result from the `PatternAnalyzer` and the `NaiveBayesAnalyzer` for each sentence.

The NaiveBayesAnalyzer misclassified 13 of the sentences (33%), while the PatternAnalyzer misclassified 8 (20%). To be fair to the NaiveBayesAnalyzer, it was trained on movie reviews, but it was still a low accuracy.

Ground Truth	Naive Bayes	p_pos	p_neg	Pattern Analyzer	polarity	subj.	Sentence
pos	neg	+0.36	-0.64	pos	+0.35	0.30	That's right....the red velvet cake.....ohhh this stuff is so good.
neg	pos	+0.79	-0.21	pos	+0.00	0.00	They never brought a salad we asked for.
pos	pos	+0.59	-0.41	pos	+0.39	0.42	This hole in the wall has great Mexican street tacos, and friendly staff.
neg	pos	+0.67	-0.33	pos	+0.20	0.78	Took an hour to get our food only 4 tables in restaurant my food was Luke warm, Our sever was running around like he was totally overwhelmed.
neg	neg	+0.40	-0.60	neg	-1.00	1.00	The worst was the salmon sashimi.
pos	neg	+0.32	-0.68	pos	+0.17	0.67	Also there are combos like a burger, fries, and beer for 23 which is a decent deal.
neg	neg	+0.40	-0.60	pos	+0.00	1.00	This was like the final blow!
pos	pos	+0.69	-0.31	pos	+0.00	0.00	I found this place by accident and I could not be happier.
neg	pos	+0.65	-0.35	pos	+0.47	0.53	seems like a good quick place to grab a bite of some familiar pub food, but do yourself a favor and look elsewhere.
pos	pos	+0.72	-0.28	pos	+0.00	0.00	Overall, I like this place a lot.
neg	neg	+0.33	-0.67	pos	+0.23	0.60	The only redeeming quality of the restaurant was that it was very inexpensive.
pos	pos	+0.65	-0.35	pos	+0.70	0.60	Ample portions and good prices.
neg	neg	+0.31	-0.69	neg	-0.60	0.80	Poor service, the waiter made me feel like I was stupid every time he came to the table.
pos	pos	+0.82	-0.18	pos	+0.31	0.33	My first visit to Hiro was a delight!
neg	neg	+0.09	-0.91	neg	-0.30	0.30	Service sucks.
pos	pos	+0.53	-0.47	pos	+0.00	0.00	The shrimp tender and moist.
neg	neg	+0.42	-0.58	pos	+0.20	0.39	There is not a deal good enough that would drag me into that establishment again.
neg	neg	+0.44	-0.56	neg	-0.10	0.67	Hard to judge whether these sides were good because we were grossed out by the melted styrofoam and didn't want to eat it for fear of getting
pos	pos	+0.96	-0.04	pos	+0.52	0.77	On a positive note, our server was very attentive and provided great service.
neg	neg	+0.10	-0.90	neg	-0.70	0.85	Frozen pucks of disgust, with some of the worst people behind the register.
neg	pos	+0.83	-0.17	pos	+0.00	1.00	The only thing I did like was the prime rib and dessert section.
neg	neg	+0.14	-0.86	neg	-0.35	0.33	It's too bad the food is so damn generic.
pos	neg	+0.18	-0.82	pos	+0.49	0.57	The burger is good beef, cooked just right.
pos	neg	+0.14	-0.86	pos	+0.00	0.00	If you want a sandwich just go to any Firehouse!!!!!
pos	pos	+0.99	-0.01	pos	+0.22	0.33	My side Greek salad with the Greek dressing was so tasty, and the pita and hummus was very refreshing.

pos	pos	+0.99	-0.01	pos	+0.20	0.56	We ordered the duck rare and it was pink and tender on the inside with a nice char on the outside.
pos	pos	+0.77	-0.23	pos	+0.00	0.00	He came running after us when he realized my husband had left his sunglasses on the table.
pos	neg	+0.48	-0.52	pos	+0.88	0.60	Their chow mein is so good!
neg	pos	+0.88	-0.12	neg	-0.25	0.60	They have horrible attitudes towards customers, and talk down to each one when customers don't enjoy their food.
pos	neg	+0.49	-0.51	pos	+0.50	0.90	The portion was huge!
pos	pos	+0.97	-0.03	pos	+0.78	0.81	Loved it...friendly servers, great food, wonderful and imaginative menu.
neg	neg	+0.19	-0.81	pos	+0.08	0.48	The Heart Attack Grill in downtown Vegas is an absolutely flat-lined excuse for a restaurant.
neg	pos	+0.78	-0.22	neg	-0.10	0.20	Not much seafood and like 5 strings of pasta at the bottom.
pos	pos	+0.98	-0.02	pos	+0.64	0.77	The salad had just the right amount of sauce to not over power the scallop, which was perfectly cooked.
neg	neg	+0.13	-0.87	neg	-0.30	0.95	The ripped banana was not only ripped, but petrified and tasteless.
neg	pos	+0.76	-0.24	neg	-0.30	0.40	At least think to refill my water before I struggle to wave you over for 10 minutes.
pos	pos	+0.82	-0.18	pos	+0.00	0.00	This place receives stars for their APPETIZERS!!!
pos	pos	+0.88	-0.12	pos	+1.00	1.00	The cocktails are all handmade and delicious.
pos	pos	+0.64	-0.36	pos	+0.00	0.00	We'd definitely go back here again.
pos	pos	+0.61	-0.39	pos	+0.50	1.00	We are so glad we found this place.

Source code

```
'''COT6930 Natural Language Processing Spring 2019
Christian Garbin

Assignment 3
Sentiment analysis with TextBlob
'''

from textblob import TextBlob
from textblob import Blobber
from textblob.sentiments import NaiveBayesAnalyzer
from textblob.sentiments import PatternAnalyzer

import textblob as tb
print(tb.__version__)

# Read all sentences in the file into one string.
# The `replace(..)` is used to combine sentences that are split into two
# lines in the file back into one line.
# TextBlob will parse into sentences, using . ! ? etc. to separate them.
with open('./assignment3/homework 3 dataset.txt', 'r') as file:
    text = file.read().replace('\n', ' ')

blob = TextBlob(text)
```

```
print('Naive Bayes Analyzer')
tb = Blobber(analyzer=NaiveBayesAnalyzer())
for sentence in blob.sentences:
    sentiment = tb(str(sentence)).sentiment
    print('{} +{:.2f} -{:.2f} {}'.format(sentiment.classification,
                                        sentiment.p_pos, sentiment.p_neg,
                                        sentence))

print('\n\nPattern Analyzer')
tb = Blobber(analyzer=PatternAnalyzer())
for sentence in blob.sentences:
    sentiment = tb(str(sentence)).sentiment
    classification = 'pos' if sentiment.polarity >= 0 else 'neg'
    print('{} {:+.2f} {:.2f} {}'.format(classification,
                                        sentiment.polarity,
                                        sentiment.subjectivity,
                                        sentence))
```

The development environment used