



## Análisis de sentimientos: comparativa de las series “Big Bang Theory” y “Silicon Valley”

En este trabajo realizaré un análisis de los sentimientos relacionados con dos series de televisión: “Big Bang Theory” y una serie comparativa en cuanto a tipo y público “Silicon Valley”.

La motivación de este análisis es la noticia sobre los millonarios cachés de los protagonistas de la serie “Big Bang Theory”. Analizaré si la serie realmente está gustando mucho más que sus comparables y si vale la pena pagar tanto a los actores.

El análisis se realiza en base a los últimos 1000 tweets descargados de tweeter sobre ambas series.

### Big Bang Theory

```
In [108]: tweet_list

Out[108]: [{"New 'Big Bang Theory' episode 11,season 10 Delayed Tonight, Set to air on Dec 15 Instead https://t.co/h9x7iOX0QX
'BigBangTheory'
'RT @EinsteinsofGod: The God Model Project can neither confirm nor deny that Kip Thorne is a member. \n#rawscienc
efilmfestival #bigbangtheo\xe2\x80\xa6',
'RT @E4tweets: We all have that one friend who loves Christmas songs a little too much\xe2\x80\xa6 #BigBangTheory
https://t.co/kYFygM0vsR',
'In 20 short minutes get ready for some belly aching laughs. The #BigBangTheory at 3:55PM PST on E4. Stream here:
https://t.co/v7uQ2HFP',
'Don\xe2\x80\x99t be crazy! Watch another episode of #BigBangTheory tonight at 7pm on #TheCWProvidence! https://t.co/40fogyKnID
https://t.co/AacFnItxu6',
'i'm struggling.... i want promo pics!! NOW NOW!! #BigBangTheory 10X11 #Shamy things coitus, amy's b-day, hallie h
https://t.co/6pz3lAnnuf',
'RT @bigbangtheory: Howard and Bernadette's baby is on the way! Make sure you watch the latest #BigBangTheory befo
re it comes next week: htt\xe2\x80\xa6',
'RT @bigbangtheory: Howard and Bernadette's baby is on the way! Make sure you watch the latest #BigBangTheory befo
re it comes next week: htt\xe2\x80\xa6',
'RT @bigbangtheory: Howard and Bernadette's baby is on the way! Make sure you watch the latest #BigBangTheory befo
re it comes next week: htt\xe2\x80\xa6',
'I bet #sheldoncooper his favorite #Christmasmovie is #PolarExpress #trains @bigbangtheory @jimpersonsnet',
100 #sheldoncooper Did you see Christmas 3Ded on #BigBangTheory? Watch his response now! http://t.co/8pQGM
```

### Silicon Valley

```
In [95]: tweet_list

Out[95]: [{"Monica: Peter is going to be furious when he finds out...\nHe's going to make that noise.\n #SiliconValley #Silic
onHBO #SiliconValleyHBO",
'Richard: Yeah, we are the... finally the, the belles of the balls. #SiliconValley #SiliconHBO #SiliconValleyHBO',
'Naveen: I invited you to my wedding, Gavin. \nGavin: Why?\n #SiliconValley #SiliconHBO #SiliconValleyHBO',
'Erich: Jesus! This is a good apple. Where did you get this?\nBig Head: Oh, I have a fruit guy. #SiliconValley #S
iliconHBO #SiliconValleyHBO',
'RT @SiliconHBO: Congratulations to the writers of #SiliconValleyHBO on the #WGA nomination for best writing for a
comedy series. https://t.co/xe2\x80\xa6',
'RT @SiliconHBO: Who doesn't want a Pied Piper t-shirt for the holidays? Get 30% off #SiliconValleyHBO products wi
th code: CYBERMONDAYSV htt\xe2\x80\xa6',
'Gavin: In case you didn't know, here at Hooli, we make the impossible possible. #SiliconValley #SiliconHBO #Silic
onValleyHBO',
'Erich: You tore double-asshole a third asshole! Triple-A! #SiliconValley #SiliconHBO #SiliconValleyHBO',
'Jared: I have a fragile posterior. My aunt used to call me "glasshole." #SiliconValley #SiliconHBO #SiliconValley
HBO',
'Yeah HBO\xe2\x80\x99s Silicon Valley is confusing: One Month\xe2\x80\x99s Learning Library can clear that up for
you. #SiliconValleyHBO https://t.co/N4o58zWANO',
'Erich: Kind of a Mentor-mentee sort of thing\nGilfoyle: Coming from a manatee sort of thing. #SiliconValley #Sil
iconHBO #SiliconValleyHBO',
'Erich: You look like a force that came on earth 6 months ago. #SiliconValley #SiliconHBO #SiliconValleyHBO'}
```

Se procesan los tweets, extrayendo el texto, desglosando en palabras y quitando los caracteres innecesarios y espacios.

```
In [109]: #Con esta función se van a formatear los tweets

def processTweet(tweet):
    tweet = tweet.lower()
    # www.* or https://.* to URL
    tweet = re.sub('([www\.|https://])', '', tweet)
    #Username to AT_USER
    tweet = re.sub('@([\w]+)', 'AT_USER', tweet)
    #Remove additional white spaces
    tweet = re.sub('([ ]+)', ' ', tweet)
    #Replace $word with word
    tweet = re.sub(r'(\$([\w]+))', r'\1', tweet)
    #trim
    tweet = tweet.strip('\'\"')
    return tweet

stopWords = []

#quitando caracteres repetidos
def replaceTwoOrMore(s):
    pattern = re.compile(r'(.+)\1+', re.DOTALL)
    return pattern.sub(r'\1', s)

In [110]: # Dividimos los tweets en palabras
def getFeatureVector(tweet):
    featureVector = []
    #split tweet into words
    words = tweet.split()
    for w in words:
        #replace two or more with two occurrences
```

Los tweets se clasificarán con la metodología de machine learning, usando el algoritmo Naive Bayes. El entrenamiento realizo con unas listas de palabras positivas y negativas extraídas de Loughran Mc Donald Master Dictionary 2014, con ciertas modificaciones.

```
In [113]: # palabras positivas y negativas que voy a usar para entrenar, podria entrenar con un corpus de tweets etiquetados pero tengo lista de palabras
positivos = read_corpus('./positiveMcDonalds.csv', 'POS')
no_positivos = read_corpus('./negativeMcDonalds.csv', 'NO_POS')
```

Una vez entrenado el algoritmo se clasifican los tweets entre positivos y no positivos.

## Big Bang Theory

```
# A continuación clasifico los tweets sobre la serie "Big Bang Theory".
for tweet in tweet_list:
    processedTestTweet = processTweet(tweet)
    print tweet + ' ----> ' + NBClassifier.classify(extract_features(getFeatureVector(processedTestTweet))) + '\n'

Vissie just watched the Big Bang Theory 6x1 "The Monster Isolation" https://t.co/nVwlmZuk #BigBangTheory #trak
t ----> NO_POS

The Big Bang Theory 10x10 "The Property Division Collision" *****☆ https://t.co/2dytKVdrr1 #BigBangTheory #
trak ----> NO_POS

#SigBangTheory happy days ----> POS

RT @E4Tweets: Sheldon: a big bang expert in more ways than one. #SigBangTheory https://t.co/0ZviwHbi8c ----> NO_
POS

The Big Bang Theory 6x16 "The Tangible Affection Proof" *****☆ https://t.co/YFsmIuXWf #BigBangTheory #trak
t ----> NO_POS

Vissie just watched The Big Bang Theory 6x16 "The Tangible Affection Proof" https://t.co/YFsmIuXWf #BigBangTheory
#trak ----> NO_POS

RT @richnairn: Loving the latest #BigBangTheory so good! @billprady @missmayim @kunalnayyar @simonhelberg @MelissaR
auch not posted this in_ ----> NO_POS
```

## Silicon Valley

```
In [96]: # Clasificación de los tweets de la serie "Silicon Valley"
for tweet in tweet_list:
    processedTestTweet = processTweet(tweet)
    print tweet + ' ----> ' + NBClassifier.classify(extract_features(getFeatureVector(processedTestTweet))) + '\n'

Monica: Peter is going to be furious when he finds out...
He's going to make that noise.
#SiliconValley #SiliconHBO #SiliconValleyHBO ----> NO_POS

Richard: Yeah, we are the... finally the, the belles of the balls. #siliconValley #siliconHBO #siliconValleyHBO
----> NO_POS

Naveen: I invited you to my wedding, Gavin.
Gavin: Why?
#SiliconValley #SiliconHBO #SiliconValleyHBO ----> NO_POS

Erich: Jesus! This is a good apple. Where did you get this?
Big Head: Oh, I have a fruit guy. #SiliconValley #SiliconHBO #SiliconValleyHBO ----> POS

RT @SiliconHBO: Congratulations to the writers of #SiliconValleyHBO on the #WGA nomination for best writing for a c
omedy series. https://t... ----> POS

RT @SiliconHBO: Who doesn't want a Pied Piper t-shirt for the holidays? Get 30% off #SiliconValleyHBO products with
code: CYBERMONDAYSV htt... ----> NO_POS
```

Finalmente cuento los tweets positivos de ambas series.

```
201
NO_POS

201
NO_POS

201
POS

202
```

202 tweets sobre la serie "Big Bang Theory" han sido positivos.

```
---

126
NO_POS

126
NO_POS

126
POS

127
```

127 tweets sobre la serie "Silicon Valley" han sido positivos.

Efectivamente la serie “Big Bang Theory” tiene bastantes más tweets positivos.