Sentiment Analysis using Modified Naive Bayes Classifier

Sentiment Analysis

"Sentiment analysis is the task of identifying positive and negative opinions, emotions and evaluations in text"

The weather is great today :)

I think its almost 30 degrees today

Nooo, it is very humid :(

Opinion

Fact

Opinion

Introduction

- Sentiment analysis (opinion mining) is a Natural Language Processing and information extraction task.

- Identify opinion expressed in text as positive, negative and neutral.

Need

- Mainly built for commercial use.

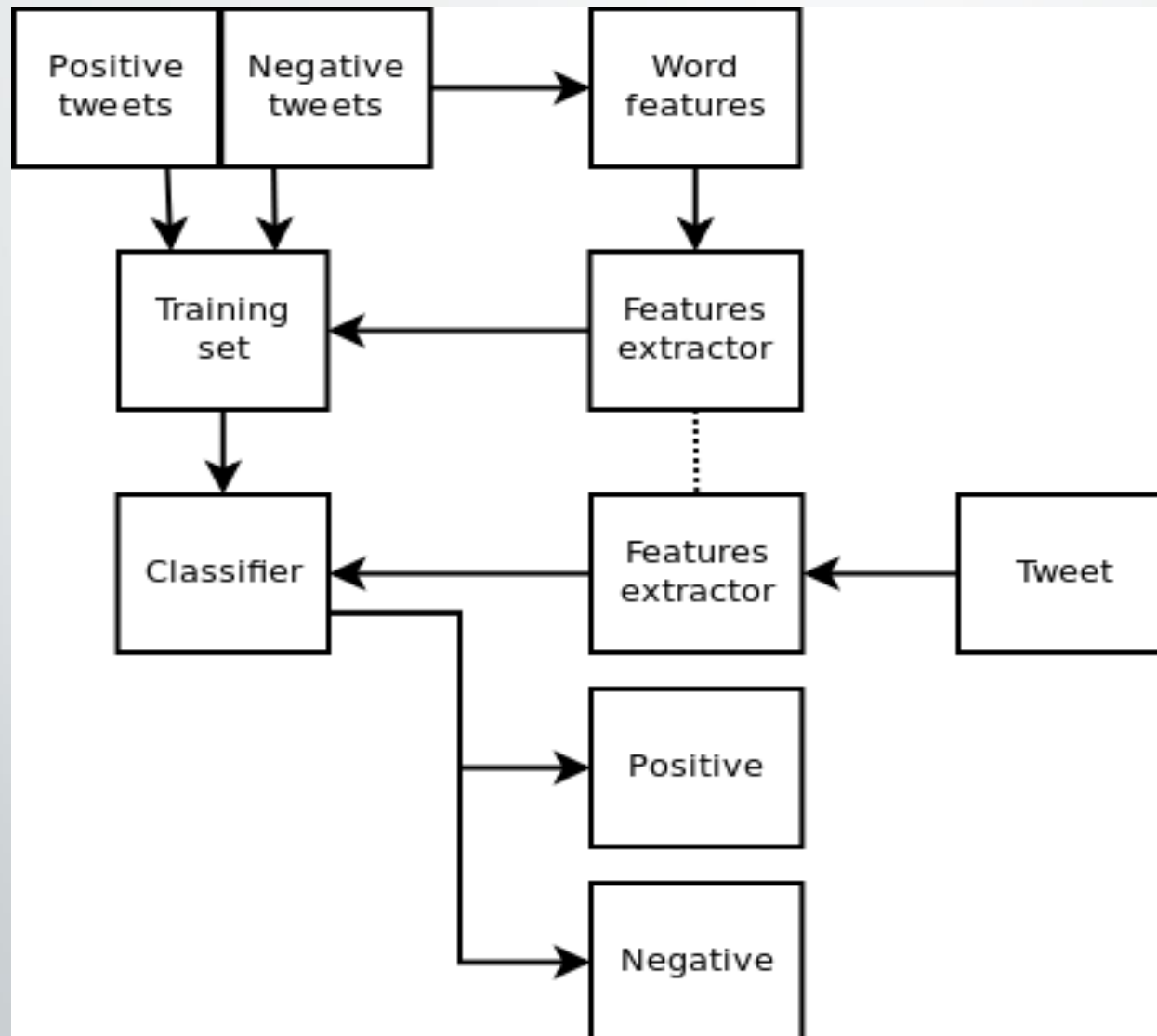"Sentiment Analysis using Modified Naïve Bayes Classifier"

Aim

To analyze the sentiments about a particular movie of user's interest which is discussed on twitter.

Here is a summary of what we did

Implementation Details

EXAMPLE

Good: 506
Bad: 507
Goodness: 506/(506+507) = **0.5**
Badness: 507/(506+507)= **0.5**

Good: 10
Bad: 14
Goodness: 10/(14+10) = **0.41**
Badness: 14/(14+10)= **0.59**

"it's rather like a lifetime special -- pleasant , sweet and forgettable . "

Good: 15
Bad: 6
Goodness: 15/(6+15) = **0.71**
Badness: 6/(6+15)= **0.29**

Good: 46
Bad: 22
Goodness: 46/(46+22) = **0.68**
Badness: 22/(46+22)= **0.32**

"it's rather like a lifetime special -- pleasant , sweet and forgettable . "

	#GOOD	#BAD	GOODNESS	BADNESS
it's	506	507	0.5	0.5
rather	42	63	0.4	0.6
like	242	396	0.61	0.39
a	3446	3112	0.53	0.47
lifetime	3	5	0.38	0.62
special	29	40	0.42	0.58
pleasant	15	6	0.71	0.29
sweet	46	22	0.68	0.32
and	3198	2371	0.57	0.43
forgettable	10	14	0.42	0.58

"BAG OF WORDS" model

it's sweet special rather pleasant

SUM: 5.22 4.8

So we should classify this as a POSITIVE review!

PLAN

- Unigram tokenizer (look at one word at a time)
- Tokens were stemmed (with the Porter stemmer algorithm; this converts "greater", "greatest", and "greatly" all to "great", thereby giving us more data points for the concept "great" and reducing our entropy)
- Flag any word to the left and right of a negation term (i.e., put an exclamation point in front of any word next to "not" or "never" or similar; this turns that word into a different, negatively-connotated word as far as Bayes is concerned. "great" becomes distinct from "!great", which implies "not great".)
- Deal with rare tokens by pulling them towards 50% "wordicity" (a term I use in the Bayes article) with a weight of 3 (i.e., words we've seen fewer times are forced closer to 50%, because we don't trust them as much).
- Only label a document if we determine the probability to be 75% or greater.
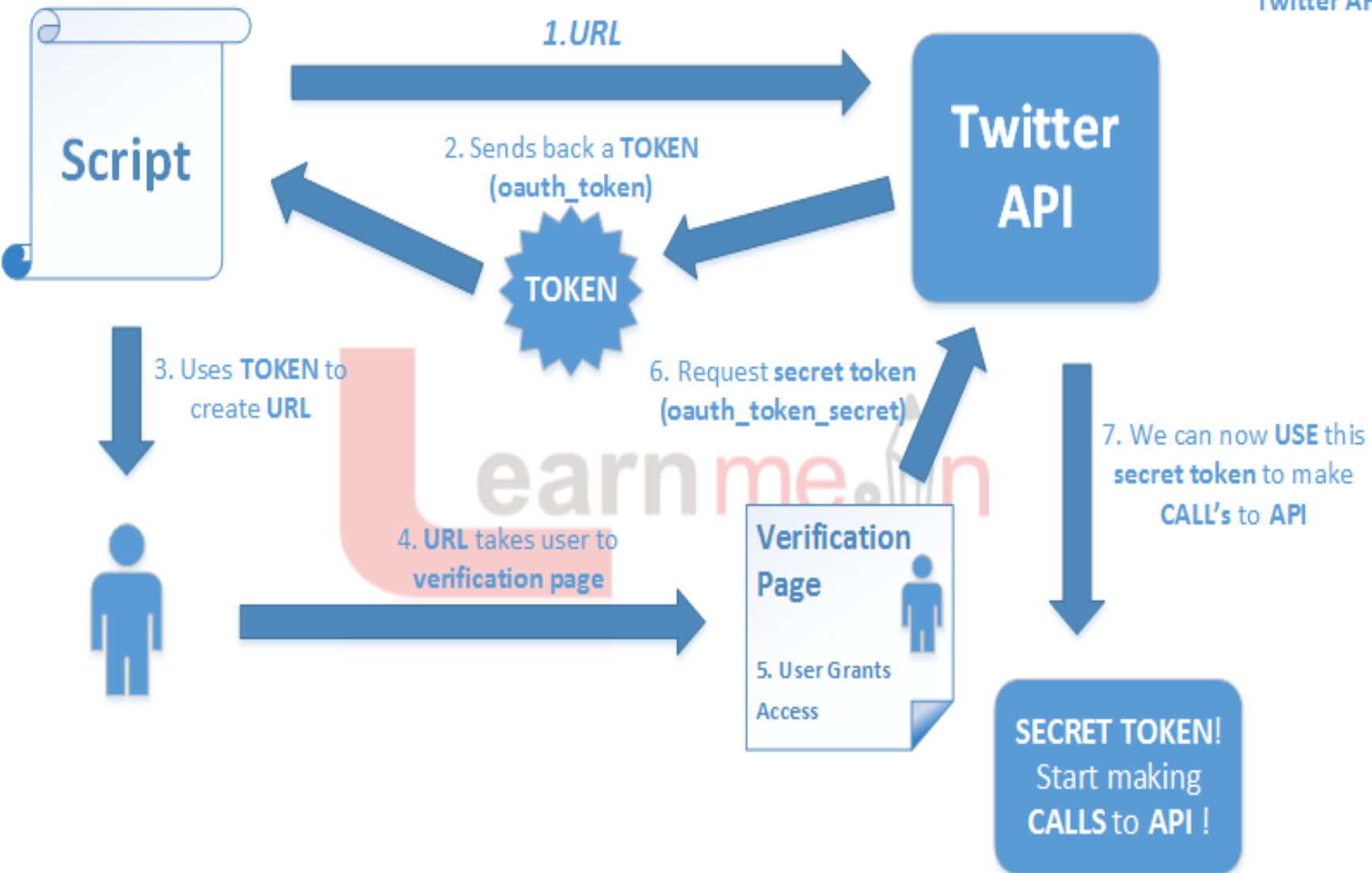
Dataset

Positive and **Negative** classified dataset of Movie Reviews is obtained from website of **University of Cornell**

Neutral dataset of Movie Reviews is **manually** collected and classified from twitter.

TWITTER API + OAUTH2

Author: Sachin Rajput
Twitter API Workflow

Create an application

Application Details

Name: *

BresponZive Pro

Your application name. This is used to attribute the source of a tweet and in user-facing authorization screens. 32 characters max.

Description: *

BresponZive Pro Social Counter Widget

Your application description, which will be shown in user-facing authorization screens. Between 10 and 200 characters max.

Website: *

http://themepacific.com/

Your application's publicly accessible home page, where users can go to download, make use of, or find out more information about your application. This fully-qualified URL is used in the source attribution for tweets created by your application and will be shown in user-facing authorization screens.

(If you don't have a URL yet, just put a placeholder here but remember to change it later.)

Callback URL:

Where should we return after successfully authenticating? For @Anywhere applications, only the domain specified in the callback will be used. OAuth 1.0a applications should explicitly specify their oauth_callback URL on the request token step, regardless of the value given here. To restrict your application from using callbacks, leave this field blank.

Technical site
http://www.ciphergig.com

Organization

Information about the organization or company associated with your application. This information is optional.

Organization	None
Organization website	None

OAuth settings

Your application's OAuth settings. Keep the "Consumer secret" a secret. This key should never be human-readable in your application.

Access level	Read-only
	About the application permission model
Consumer key	zHb1YHPPGJprLAMGb9sSw
Consumer secret	Ekwc4LcMUSPH9wkaeiufR4DOvglYFIm5yQTxPOyEVd0
Request token URL	https://api.twitter.com/oauth/request_token
Authorize URL	https://api.twitter.com/oauth/authorize
Access token URL	https://api.twitter.com/oauth/access_token
Callback URL	http://www.ciphergig.com/
Sign in with Twitter	No

OTHER NEEDED PYTHON SUPPORTING LIBRARIES. . .

Setuptools- 0.6c11

web.py- 0.37

What is a Pickle file?

Use of Google Charts.

Number of Tweets by Time and Sentiment

⦿ Today ○ Last 7 days
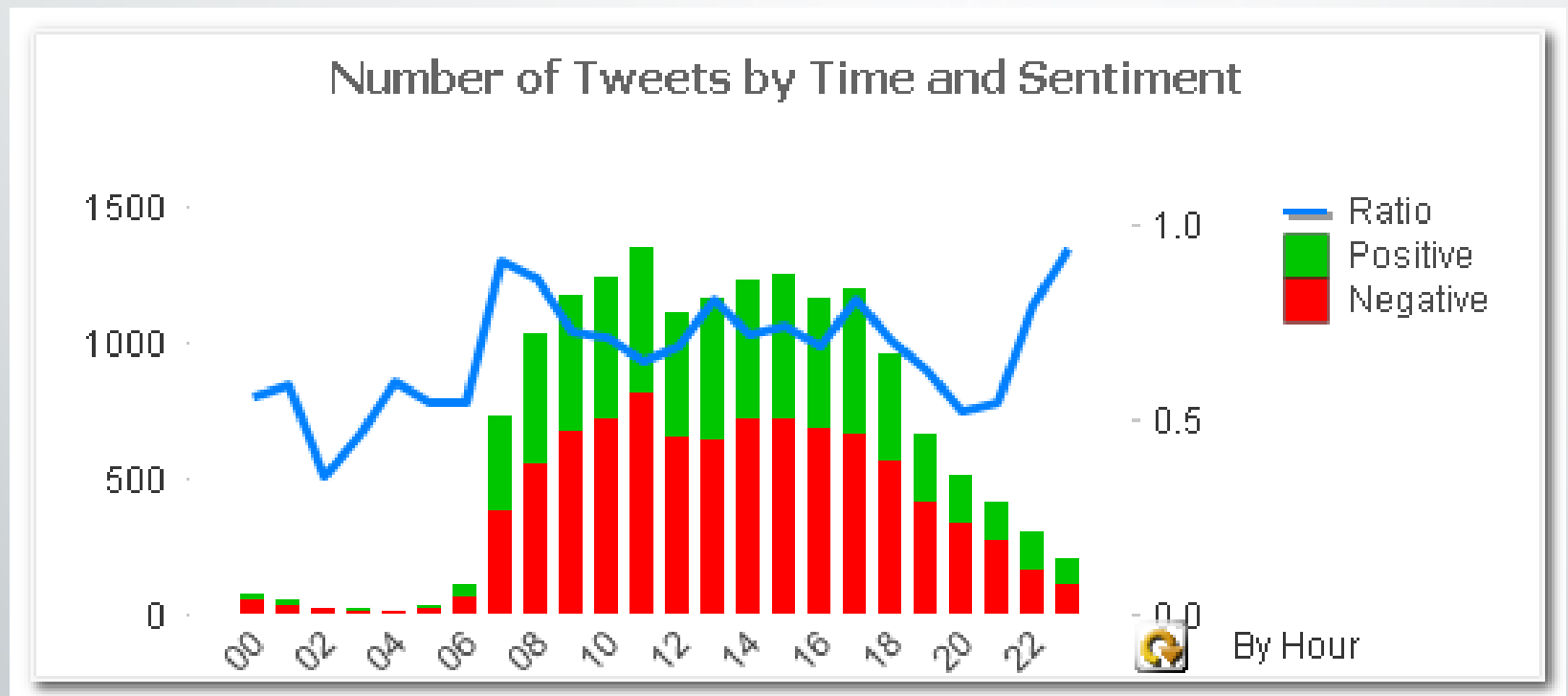
○ Baseline ⦿ Naive Bayes

Sentiment Classification

Positive

"@girlposts: These Disney paintings ðŸ˜³ http://t.co/3Z8esVxAkW" I have the Cinderella one as a puzzle! I need the rest! ðŸ˜™

"Cinderella" was enchanting! I especially loved the "brave & kind" theme. Incredibly empowering words girls need to hear. Bravo!

RT @shoppriceless: ðŸ˜›ðŸ˜‘ Modern Day Cinderella Dress: http://t.co/6HjqjdVy6X http://t.co/xKlgacvY3G

fun fact: anna cried 10 minutes into Cinderella

LOVED ðŸ˜˜ðŸ˜˜ðŸ˜˜ every minute of this movie!!! #Cinderella #Disney Lily James really brought the character & story to life http://t.co/iYkTJxwarL

I'm seeing Cinderella xox

Cinderella was absolute perfection â¤ï¸â¤

Neutral

When you ask for a ticket to see Cinderella and the attendant hands you 50 Shades of Gray...humm very different movies bud #fail

Last friday night :) @idktricia @nhikky14 @Jesseeeeeeeeeei @elleluminario @deniseesison #cielacortez #Cinderella ðŸ˜‘ðŸ˜‘Œ�ðŸ˜‘ðŸ˜‘‹ http://t.co/ZlGBowrMHK

RT @iUpdateMixers: Gonna go watch Cinderella later ðŸ˜™Š Can't wait to see Frozen Fever too â˜ºðŸ˜™– -S #KCA #VoteLittleMixUK

Seeing #Cinderella tonight with @IsabellaAlves96

How everybody watched/is watching CinderellaðŸ˜¡....... #k

RT @Scholastic: Seeing @CinderellaMovie this weekend? Author @sarahmlynowski talks #Cinderella & #fairytales: http://t.co/OpwwqLdl4D http:/â€¦

Negative

RT @Jojo_51213: Cinderella was obviously drunk. Why else would she lose her shoe and not even run back for it? Shoes ain't cheap. Especialâ€¦

I don't care what anyone says Cinderella was a great movie

RT @Kel1324: dairy queen and cinderella ðŸ˜»

We saw #Cinderella today ðŸ˜˜ now I'm just waiting for #Pan to come out and then I'll be ready forâ€¦ https://t.co/bWLKIvenpB

GO SEE CINDERELLA THE PRINCE IS SO FIRLDND CUTE OH POOPY THIS MAN AND HIS ACCENT

The iDOLM@STER Cinderella Girls TV anime will air in split seasons http://t.co/05rAzzd0GD

RT @shoppriceless: ðŸ˜› Modern Day Cinderella Dress #ShopPriceless

RT @DisneyStudiosCA: #LilyJames hit the red carpet in Toronto this week. #Cinderella is now playing! Get tix here: http://t.co/VlUsdAyxb4 hâ€¦

Absolutely loved it!! Beautiful telling of a classic! #courage #kindness â€" watching Cinderella at Santikos... http://t.co/qI1RjDlzB4

I will be 80 years old and still going to see the latest Disney film #Cinderella #disneykid4life

RT @ellabigg: Romantic movies with baeðŸŽ€ðŸ'• @EmmaSlotabec #Cinderella

I wanted to go see the Cinderella movie today

â¤ï¸ the new #Cinderella @drafthouse with the whole family! #HaveCourageBeKind http://t.co/z7XYwQXJrL

RT @annaujwala_26: Just saw Cinderella it was Amazing ðŸ¨™

Obsessed with Cinderella rn. I cried like at every sad part. BEST MOVIE I'VE SEEN YET.

Cinderella: A Night at the Ball - $1.44 - http://t.co/3SDSQrUgab - #Cinderella2015 #Wow http://t.co/wNk4rALPb1

Saw the new #Cinderella last night, and LOVED IT. Like the cartoon come to life, with added layers. And gorgeous visuals.

Cinderella was so magical and perfectâ˜ºï¸

I need to see Cinderella ðŸ˜ðŸ˜

Cinderella was so good. I'm tore up.

About to start watching Last Cinderella

RT @Gr8DisneyQuotes: Have faith in your dreams and someday your rainbow will come smiling through. â€"Cinderella

Watched CinderellaðŸ'™ðŸ'ŽðŸ'™

Will someone go see Cinderella with me? Pls

Omg Cinderella was soooo good!! ðŸ'™

RT @StarplexCinemas: Who are you seeing #Cinderella w/ this weekend? TAG them in a tweet w/ @StarplexCinemas to ENTER TO WIN a prize pack! â€¦

RT @CinderellaMovie: Watch what happens when the live-action #Cinderella trailer meets the animated version. [via @OhMyDisney] https://t.coâ€¦

RT @jessicapak: I wanna see Cinderella

RT @MacyCalder: Cinderella is officially my favorite disney movie...ðŸˆ¸

@twolfhes i thought it was amazing but ive always been obsessed with cinderella soo

RT @KingBach: Yayyyy there was 1 black guy in the Cinderella movie. We've come a long way

the Cinderella movie was alright lol

@lindseyduquette I wanna see Cinderella again ðŸ˜©

Watching Cinderella! ðŸŽ¥ @BataFish94

i thought we were watching cinderella ðŸ˜

Cinderella Dress #ShopPriceless http://t.co/RZRpMLeY1x

Does the shoe fit? #Cinderella #Disney #GC3 http://t.co/kKNcguTOWh

HARDWARE & SOFTWARE REQUIREMENTS

Hardware requirements:

RAM : 8 GB

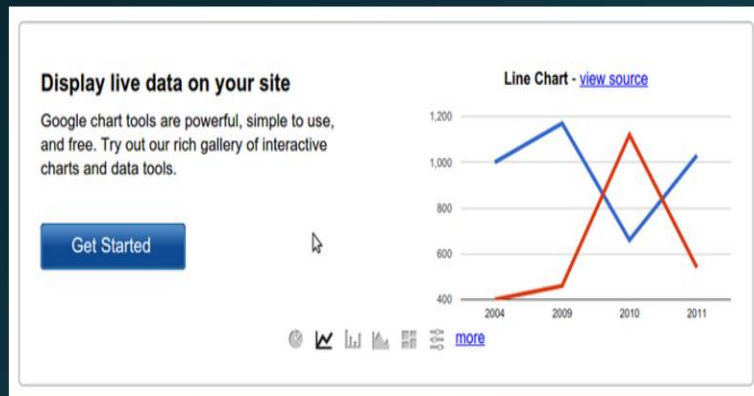**Processor : Intel i5**

Software requirements:

Ubuntu 12.04   Python 2.7

NLTK 3.0

Natural Language Processing with NLTK

Twitter API 1.1

Developers

Google Charts

Display live data on your site

Google chart tools are powerful, simple to use, and free. Try out our rich gallery of interactive charts and data tools.

Get Started

Line Chart - view source

OAuth 2.0

OAUTH

2

Testing Methods

Testing The Code
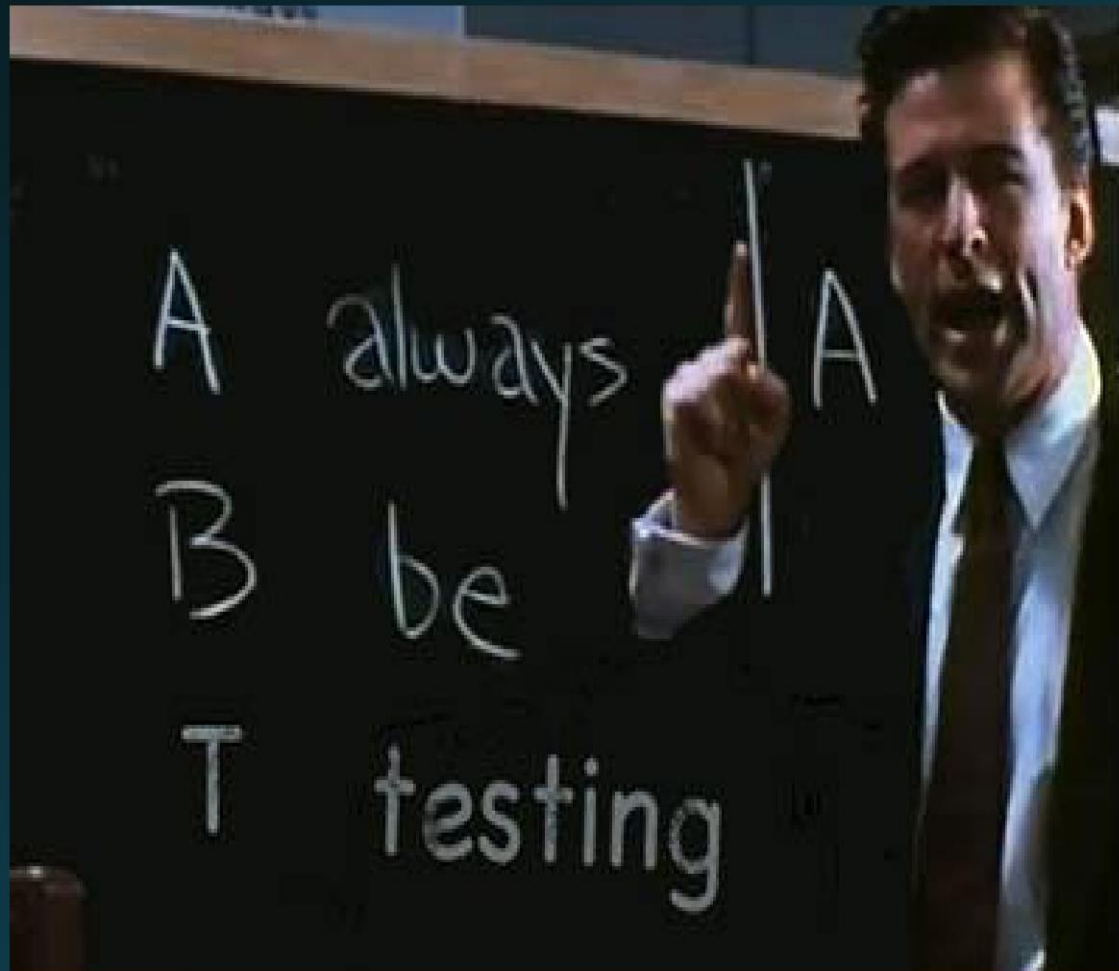
Testing Algorithm Accuracy

★ BASELINE ALGORITHM

★ NAIVE BAYES ALGORITHM

★ MAXIMUM ENTROPY ALGORITHM

Applications

Movie Reviews

Understanding Potential Customers

Automated Feedback

Predicting Poll Results

Conclusion

→ Very High Scope for Further Improvements

→ Can be Combined with Emoticons to Improve Efficiency

→ Emotions can be added like angry, happy, etc

→ Applications Can Further Be Extended to Include Artificial Intelligence, Speech Recognition, etc

THANK YOU !!