

A close-up, low-angle shot of a massive, curling ocean wave. The water is a vibrant turquoise color, with white spray and foam at the breaking point. The wave's surface is textured with fine ripples and streaks of light, suggesting sunlight filtering through the water. The overall mood is dynamic and powerful.

Sentiment Analysis Using Machine Learning

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.



Paper Published in International Journal of Trends and Technology (IJETT) - Vol:15 No:6 - Sept, 2014

“A Survey of Various Machine Learning Techniques for Text Classification”

Literature Survey

Sr. No	Author and Year	Description
1.	Pang Lee et al. (2002, 2004)	Compared the performance of classifiers on different features like considering only unigrams, bigrams, incorporate position information, taking only adjectives, small feature space etc.
2.	Bikel et al. (2007)	Implemented a Subsequence Kernel based Voted Perceptron and compared its performance with SVM. It is observed that as the number of positives increase, the increase in the number of negatives is much less.
3.	Wang et al. (2008)	Focus on concept expansion using Wikipedia where they expand the feature vector constructed from a movie review with related concepts from the Wikipedia.

Sr. No	Author and Year	Description
4.	Songho Tan (2008)	Presented empirical study of sentimental categorization on chinese documents.
5.	Alec et al. (2009)	Describes a distant supervision-based approach for sentiment classification.
6.	Barbosa et al. (2010)	Cited noisy data as one of the biggest hurdles in analyzing text.
7.	Cornell Movie Review (2010) and Stanford Sentiment Analysis NLP	Classified tweets as negative and positive only.
8.	Long-Sheng chen (2011)	Proposed neural network based approach.

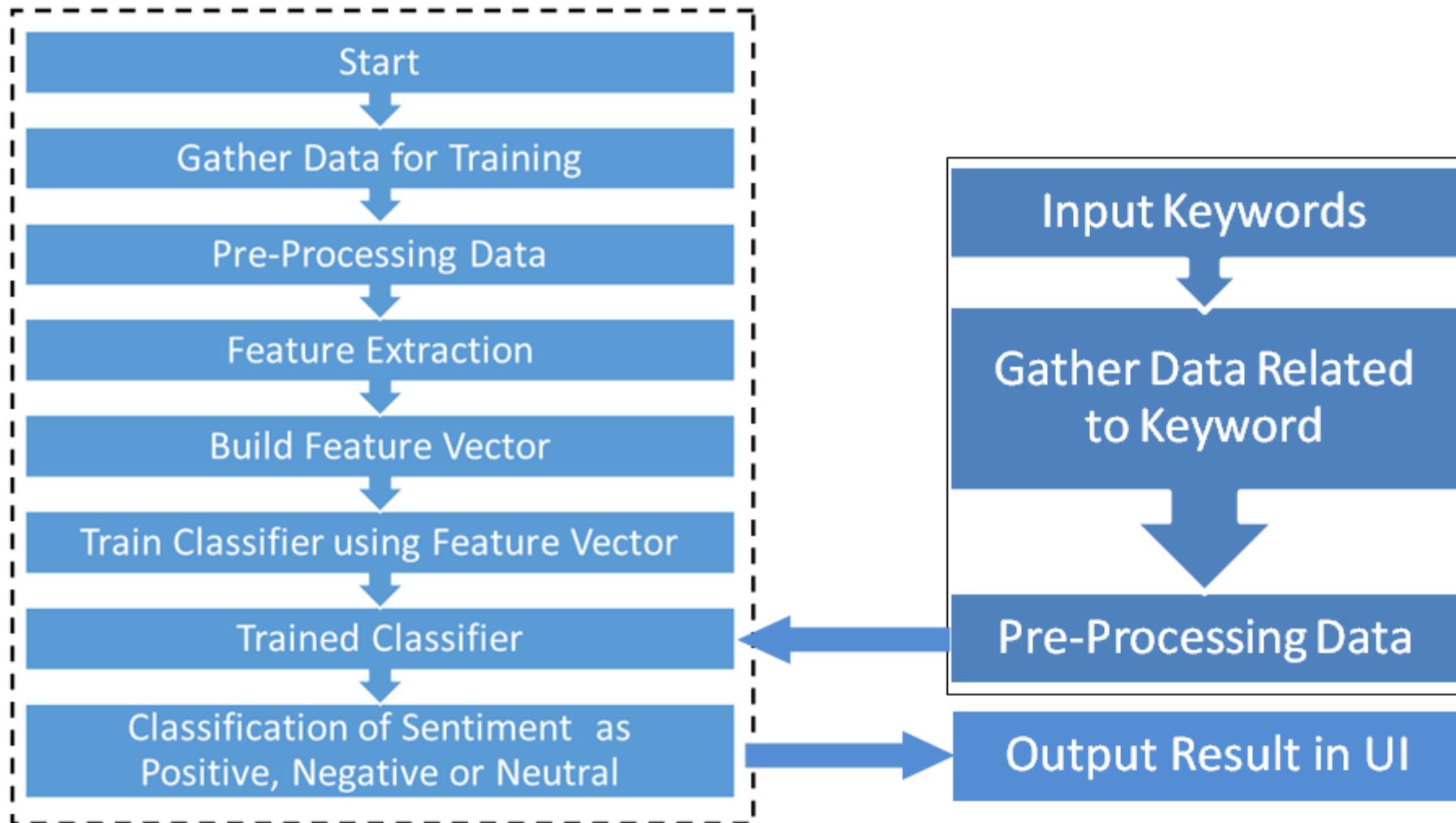
Aim

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





Block Diagram





Implementation Details

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.

Hardware requirements:

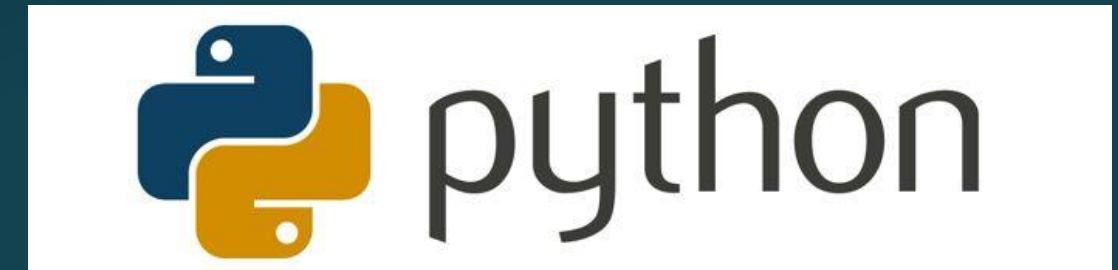
RAM : 8 GB

Processor : Intel i5

Software requirements:

Ubuntu 12.04

Python 2.7



NLTK 3.0

Natural Language Processing with NLTK

Google Charts



Twitter API 1.1



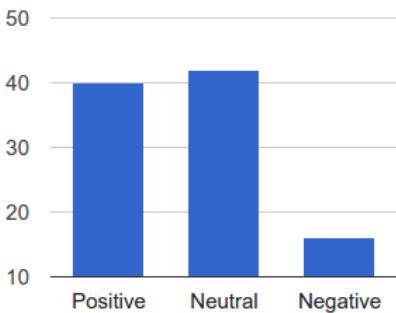
Developers

OAuth 2.0





Results

Today Last 7 days Baseline Naive Bayes**Sentiment Classification****Positive**

"@girlposts: These Disney paintings ôÝ³ http://t.co/3Z8esVxAkW" I have the Cinderella one as a puzzle! I need the rest! ôÝ'™

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

RT @shoppriceless: ôÝ'>ôÝ“ Modern Day Cinderella Dress: http://t.co/6HjqjdVy6X http://t.co/xKlgacvY3G

fun fact: anna cried 10 minutes into Cinderella

LOVED ôÝ'>ôÝ“ 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¤â¤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 @Jesseeeeeeee @elleluminario @deniseesison #cielacortez #Cinderella ôÝ'>ôÝ“#ôÝ“ http://t.co/ZlGBowrMHK

RT @iUpdateMixers: Gonna go watch Cinderella later ôÝ'> Can't wait to see Frozen Fever too â~ôÝ‘ -S #KCA #VoteLittleMixUK

Seeing #Cinderella tonight with @IsabellaAlves96

How everybody watched/is watching CinderellaôÝ‘..... #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 ôÝ“

We saw #Cinderella today ôÝ“ 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: ôÝ'> 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/VIUsdAyxb4>

Absolutely loved it!! Beautiful telling of a classic! #courage #kindness etc watching Cinderella at Santikos... <http://t.co/qIIRjDlzB4>

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

RT @ellabigg: Romantic movies with baedZeoY• @EmmaSlotabec #Cinderella

I wanted to go see the Cinderella movie today

at the new #Cinderella @drafthouse with the whole family! #HaveCourageBeKind <http://t.co/z7XYwQXJrL>

RT @annaujwala_26: Just saw Cinderella it was Amazing ☺

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

Cinderella: A Night at the Ball - \$1.44 - <http://t.co/3SDSQuRugab> - #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 perfect~!

I need to see Cinderella ☺

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. ~Cinderella

Watched Cinderella™ðŸ'ŽðŸ™

Will someone go see Cinderella with me? Pls

Omg Cinderella was soooo good!! ☺

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

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

RT @jessicapak: I wanna see Cinderella

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

@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 ☺

Watching Cinderella! ☺ @BataFish94

i thought we were watching cinderella ☺

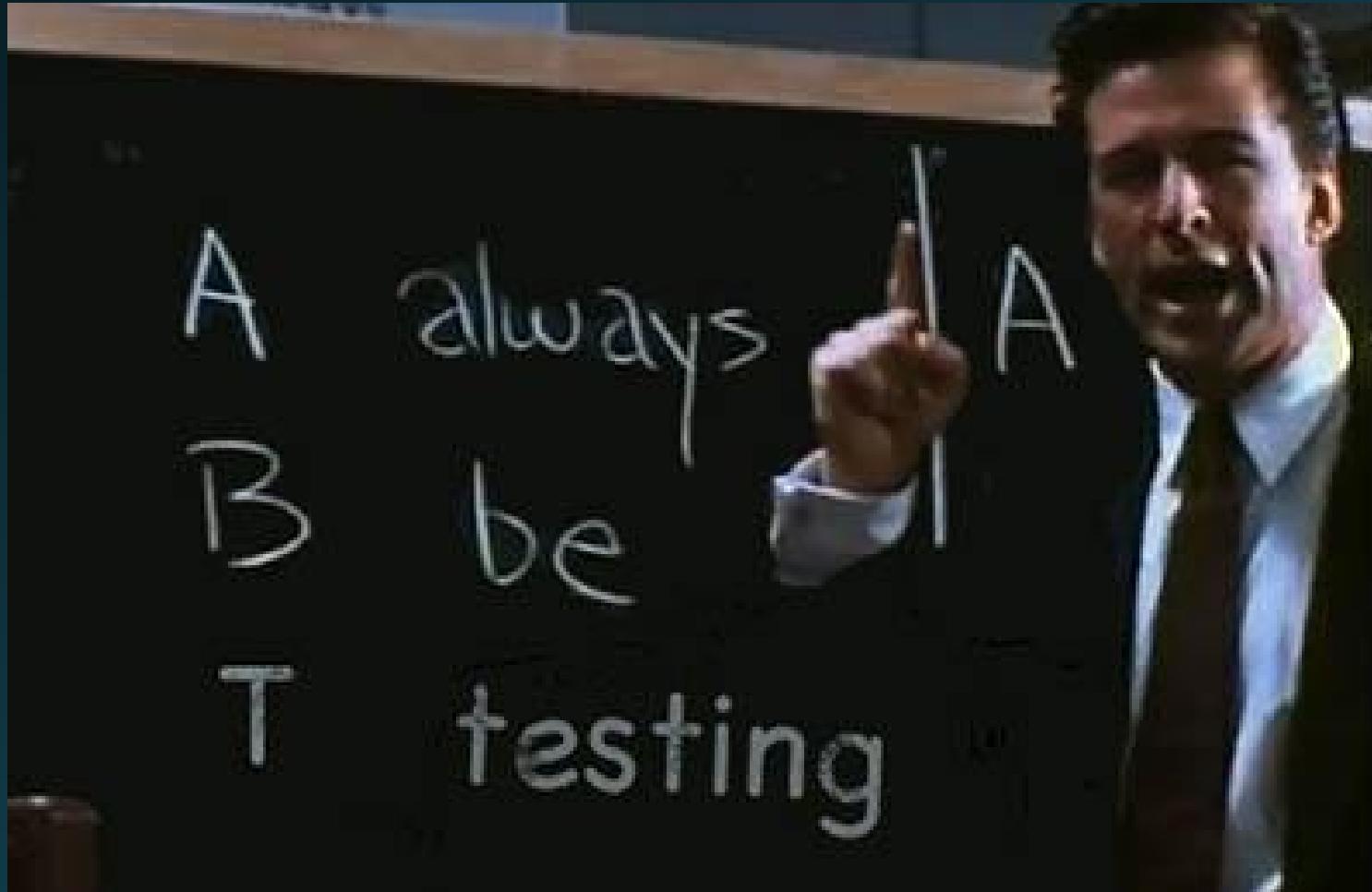
Cinderella Dress #ShopPriceless <http://t.co/RZRpMLEY1x>

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



Testing Methods

Testing The Code



Testing Algorithm Accuracy

★ BASELINE ALGORITHM

★ NAIVE BAYES ALGORITHM

★ MAXIMUM ENTROPY ALGORITHM

Most Informative Features

```
~/Project/project/simpleDemo.py - Sublime Text
  simpleDemo.py x  twitter_sentiment.py x
105 # training_set = nltk.classify.naivebayes.extract_features(extract_features, tweets,
107 #trainset=training_set[1:]
108 # Train the Naive Bayes classifier
109 #NBClassifier = nltk.NaiveBayesClassifier.train(training_set)

Most Informative Features
contains(seventeen) = True      "neutr : positi" = 3106.1 : 1.0
    contains(hates) = True       "neutr : positi" = 3106.1 : 1.0
    contains(hunger) = True      "neutr : negati" = 2961.7 : 1.0
    contains(escaped) = True     "neutr : positi" = 1863.7 : 1.0
        contains(am) = True      "neutr : negati" = 1269.3 : 1.0
contains(district) = True       "neutr : neutra" = 782.0 : 1.0
    contains(name) = True       "neutr : positi" = 716.8 : 1.0
    contains(games) = True      "neutr : negati" = 683.5 : 1.0
    contains(fifty) = True       "neutra : positi" = 226.4 : 1.0
contains(review) = True         "neutra : positi" = 188.3 : 1.0
    contains(tub) = True        "neutra : negati" = 158.3 : 1.0
    contains(home) = True       "neutr : positi" = 152.8 : 1.0
    contains(grey) = True        "neutra : negati" = 127.7 : 1.0
contains(shades) = True         "neutra : positi" = 99.8 : 1.0
contains(knight) = True         "neutra : positi" = 89.8 : 1.0
contains(lately) = True         "neutra : positi" = 80.2 : 1.0
    contains(alice) = True      "neutra : negati" = 61.4 : 1.0
contains(performance) = True   "positi : negati" = 60.7 : 1.0
    contains(machine) = True    "neutra : positi" = 55.3 : 1.0

Line 122, Column 31
Spaces: 4 | Python
```

Testing Different Movies

The
Interview

The Imitation Game

The Dark Knight

The Theory of Everything

Badlapur

American Sniper

The Grand Budapest Hotel

Fury

Birdman

Foxcatcher

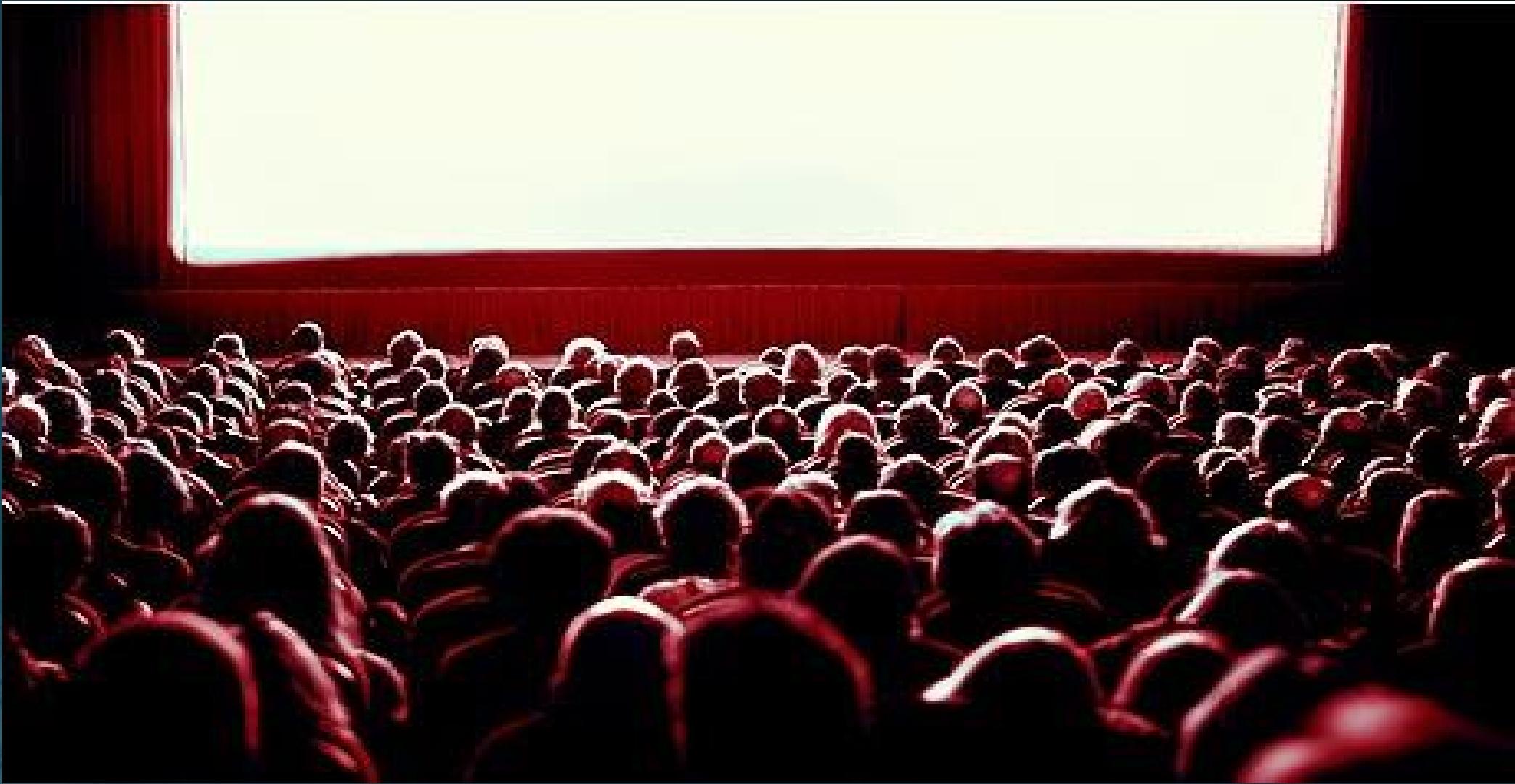
Inception

Interstellar



Applications

Movie Reviews



Understanding Potential Customers



Automated Feedback



© Int'l St Clair/Blend Images/Corbis

Predicting Poll Results



A large, powerful ocean wave is shown crashing, with white foam at the top and deep blue-green water below.

Limitations

- Sentiment Analyzer Cannot Understand Bi-Grams
- Sentiment Analyzer Cannot Understand Sarcastic,
Ironic Textual Data



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

PARIKSHIT HEGDE

SAGAR MANJARE

GAURAV CHAVAN