Sentiment Analysis
Polarity & Subjectivity

Fellow -Jithin

Sentiment Analysis

Sentiment analysis, also called **opinion mining**, is the field of study that **analyzes people's opinions**, sentiments, evaluations, appraisals, attitudes, and emotions **towards entities** such as **products, services, organizations, individuals, issues, events, topics, and their attributes.**

The inception and the **rapid growth** of sentiment analysis coincide with those of the **social media.**

Applications

- Automated Social Media Sentiment Analysis systems.
  - Twitter
  - Facebook
  - Amazon
- Intent Analysis systems.
- Fake News Detection.

Sentiment Analysis : Levels

- Document Level
  - given a product review, the system determines whether the review expresses an **overall positive or negative opinion** about the product.
- Sentence Level
  - goes to the sentences and determines whether **each sentence** expressed a positive, negative, or neutral opinion.
  - related to **subjectivity classification**.distinguishes sentences **(objective)** that express factual information from sentences that express subjective views and opinions **(Subjective)**.
- Entity and Aspect Level
  - document level and the sentence level analyses do not discover what exactly people liked and did not like.
  - It is based on the idea that an **opinion** consists of a **sentiment** (positive or negative) and a **target** (of opinion)

Sentiment Lexicons

- **Sentiment words**, also called opinion words. These are words that are commonly used to express **positive or negative sentiment**.
  - For example, *good, wonderful and amazing* are **positive sentiment** words,
  - *bad, poor, and terrible* are **negative sentiment** words.
- A list of such words and phrases is called a **sentiment lexicon** (or **opinion lexicon**)
- Sentiment lexicon is necessary but not sufficient for sentiment analysis

Issues of Sentiment Lexicon

- *'This camera sucks'* & *'This vacuum cleaner really sucks'*.
  - Meaning of certain words changes as per context.

- *'Can you tell me if this camera is good?'* & *'Does anyone know how to repair this terrible printer'*.
  - presence of a sentiment word doesn't imply it conveys that sentiment.

- *'What a great car! It stopped working in two days'*
  - Sarcasm

- *'This washer uses a lot of water'*
  - No sentiment word yet it is negative.

Sentence Subjectivity & Polarity(Emotions)

- An **objective sentence** presents some **factual information** about the world, while a **subjective sentence** expresses some **personal feelings, views, or beliefs.**

- **Polarity - emotions expressed in a sentence.**

- **Emotions** are our **subjective** feelings and thoughts. and are related to sentiments. **strength of a sentiment** or opinion is typically **linked** to the **intensity of certain emotions.**
  - Opinions that we study in sentiment analysis are mostly **evaluations.**
    - **Rational evaluations.** eg-*This car is worth the price , I am happy with this car.*
    - **Emotional evaluation.** eg-*I love iPhone , I am so angry with their service people.*

Sentence Subjectivity & Emotions

- To make use of these two types of evaluations in practice, we can design 5 sentiment ratings,
  - ***emotional negative (-2)***
  - ***rational negative (-1)***
  - ***neutral (0)***
  - ***rational positive (+1)***
  - ***emotional positive (+2)***

Drawback -

- An opinion / sentiment can have multiple perspectives
  - *The housing price has gone down, which is bad for the economy.*

Sentiment Analysis - Supervised Approach

- two-class classification problem, **positive** and **negative.** eg- Product Reviews.

  Important features -
  - Terms and their frequency(unigram) ,
    - *word and count of word*
  - Part of speech,
    - *adjectives*
  - Sentiment words and phrases,
    - *sentiment lexicon - good, bad, wonderful.*
  - Rules of opinions,
    - *After my wife and I slept on the mattress for two weeks, I saw a mountain in the middle*
  - Sentiment shifters,
    - *negation words (not only.. but also), (fails to deliver / impress).*

Sentiment Classification -Unsupervised Approach

It performs classification based on some fixed syntactic patterns that are likely to be used to express opinions. POS tags forms an integral part of this approach.

Algorithm

- Two consecutive words are extracted if their POS tags conform to a fixed pattern.
  - eg-Pattern [two consecutive words are extracted if the first word is an adverb, the second word is an adjective, and the third word is not a noun]. example sentence - ***This piano produces beautiful sounds***. - *beautiful sounds is extracted.*

Sentiment Classification -Unsupervised Approach

- It estimates the sentiment orientation (SO) of the extracted phrases using the pointwise mutual information (PMI) measure
  - PMI measures the degree of statistical dependence between two terms

$$PMI(term_1, term_2) = \log_2\left(\frac{\Pr(term_1 \wedge term_2)}{\Pr(term_1)\Pr(term_2)}\right)$$

- The sentiment orientation (SO) of a phrase is computed based on its association with the positive reference word “excellent” and the negative reference word “poor”:

- the algorithm co $SO(phrase) = \log_2\left(\frac{hits(\text{phrase } NEAR \text{ "excellent"})hits(\text{"poor"})}{hits(\text{phrase } NEAR \text{ "poor"})hits(\text{"excellent"})}\right)$ s in the review and classifies the review as positive if the average SO is positive and negative otherwise.

Thank you..!

Reference - *Sentiment Analysis and Opinion Mining by Bing Liu*

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.244.9480&rep=rep1&type=pdf

Open for Questions..!