# DALHOUSIE UNIVERSITY

# Sentiment Analysis Report

# CSCI5408 – Data Warehousing & Analytics

**Submitted by:**

**Shivam Gupta**

**B00810723**

**Shivam.Gupta@dal.ca**

## What is Sentiment Analysis?

Sentiment Analysis is a part of NLP (Natural Language Processing) and is an analysis on a text to identify and extract opinions to create a system. This system extracts attributes of the expression, such as:

- Polarity: Positive or a negative opinion.
- Subject: The topic of discussion in the text.
- Opinion Holder: The person or entity that expression an opinion.

Sentiment Analysis is done on the unstructured data presented over the Internet. Internet has many text information available on review sites, forums, blogs and social media, using which, we can identify the opinions on a product. The unstructured data could be transformed into a structured information of public opinion about product, services, brands, politics or any topic that people can express opinions about. This data can be very useful for commercial applications like marketing analysis, public relations, product reviews, net promoter scoring, product feedback, and customer service.

Opinions are subjective expression that describe people's feeling toward a subject. In Sentiment Analysis, two sub-problems must be resolved: Subjectivity classification (classifying a sentence as a subjective or objective) and Polarity Classification (classifying a sentence as a positive, negative or a neutral opinion. Opinions can be Direct, or Comparative based upon the opinion holder.

## Types of Sentiment Analysis

- Fine-grained Sentiment Analysis: precise level of polarity is considered, which include very positive, positive, neutral, negative, very negative e.g. Rating system.
- Emotion Detection: emotions like happiness and frustration are detected by using lexicons. E.g. words like *shit* or *kill* can describe emotions.
- Aspect-based Sentiment Analysis: particular features of the product are analysed. E.g. *the battery life of camera is too short*, which express about battery life, a feature of a camera.
- Intent Analysis: analyses what people want to do. e.g. *can you help me fill out the form.*
- Multilingual Sentiment Analysis: different languages are analysed for opinions.

## Why Sentiment Analysis is important?

Most of the data in this world is collected after 2010 and 80% data is unstructured. Sentiment analysis helps companies to make sense of this data to make business decisions. Advantages of Sentiment analysis are: Scalability (large amount of data processed in minutes); Real-time Analysis (real time processing); Consistent criteria (remove errors and improve data consistency).

## How does Sentiment Analysis work?

There are three type of system that implement sentiment analysis: Rule-based system; Machine learning system; Hybrid system.

Rule-based approach uses scripting language to perform classic NLP techniques like stemming, tokenization, parts of speech tagging and parsing. A list of words known as lexicons are also used to determine the subjectivity or polarity of an opinion.

- Determine a list for positive words and negative words.
- Count the number of positive or negative words in the text.
- If the number of positive words is more than number of negative words, then print a positive polarity and vice versa.

Automatic approach relies on machine learning techniques, where a classifier is fed with a text and returns the polarity of the text.

- Training and prediction process, where model learns to associate words with sample input and the feature are extracted into feature vectors, which are fed to machine learning model.
- Feature extraction, where machine learning classifier transform the text into numerical vector. This is known as feature extraction, where bag of words is collected which occurs frequently, which generates positive, negative or neutral words.
- There are many algorithms for that: Naïve Bayes; Linear Regression; Support Vector Machines; Deep Learning

Hybrid approach is an intuitive method, where both methods are used to generate the best possible way. In this way, we can improve accuracy and precision.

## Sentiment Analysis Use Cases & Applications?

Sentiment analysis is used in many places and impact businesses, cities, and society. The real word applications are:

- Social media monitoring
- Brand monitoring
- Voice of customer (VoC)
- Customer service
- Workforce analytics and voice of employee
- Product analytics
- Market research and analysis

## References:

[1] https://monkeylearn.com/sentiment-analysis/