



Sentiment Extraction-Analysis on Twitter

Under the Supervision of Prof. Arun Kumar

SUBMITTED BY:

Tenzin Choedon ENG19CS0337

Tseten Tashi Bhutia ENG19CS0339

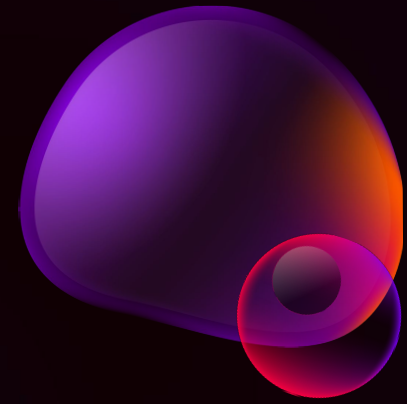
Ayush Kumar ENG19CS0055

Harsha P ENG19CS0113



Table of **contents**

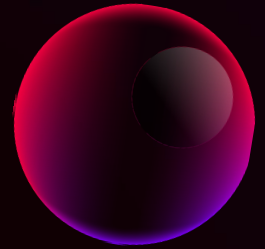
1. Abstract
2. Problem Statement
3. Introduction
4. Literature survey
5. Design
6. Methodology
7. References



ABSTRACT

- This project addresses the problem of sentiment analysis in twitter, that is classifying tweets according to the sentiment expressed in them: positive, negative or neutral.
- Due to this large amount of usage we hope to achieve a reflection of public sentiment by analysing the sentiments expressed in the tweets.
- The aim of this project is to develop a functional classifier for accurate and automatic sentiment classification of an unknown tweet stream

PROBLEM STATEMENT



- To classify polarity of a given text at the document, sentence or a features/aspect level.
- Whether the given document, sentence or a entity a features/aspect is positive, negative or neutral.





INTRODUCTION

- This project of analyzing sentiments of tweets comes under the domain of “Pattern Classification” and “Data Mining”.
- The project would heavily rely on techniques of “Natural Language Processing” in extracting significant patterns and features from the large data set of tweets and on “Machine Learning” techniques for accurately classifying individual unlabelled data samples (tweets) according to whichever pattern model best describes them.

Literature Survey

Author's Name/ Paper Title	Conference/ Journal Name and year	Technology/ Design	Results shared by author	What you infer
Jean-Francois Petiot Bernard Yannou	Measuring consumer perceptions for a better comprehension, specification and assessment of product semantics	comprehension, specification and assessment of product semantics	emerging product concepts was directly assessed with regards to the requirements in a simple qualitative and comparative way.	Pattern Classification and Data Mining

Design

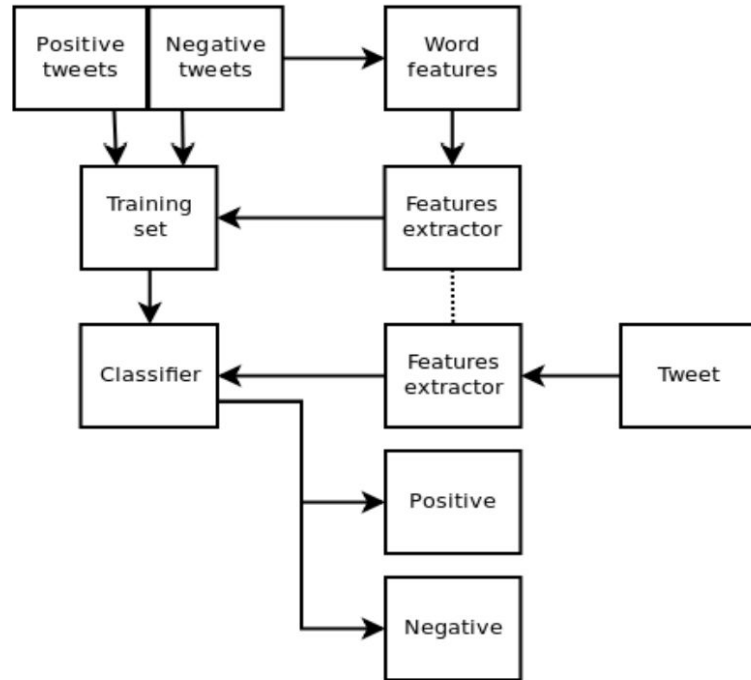


Fig.1. Sentiment Analysis Architecture

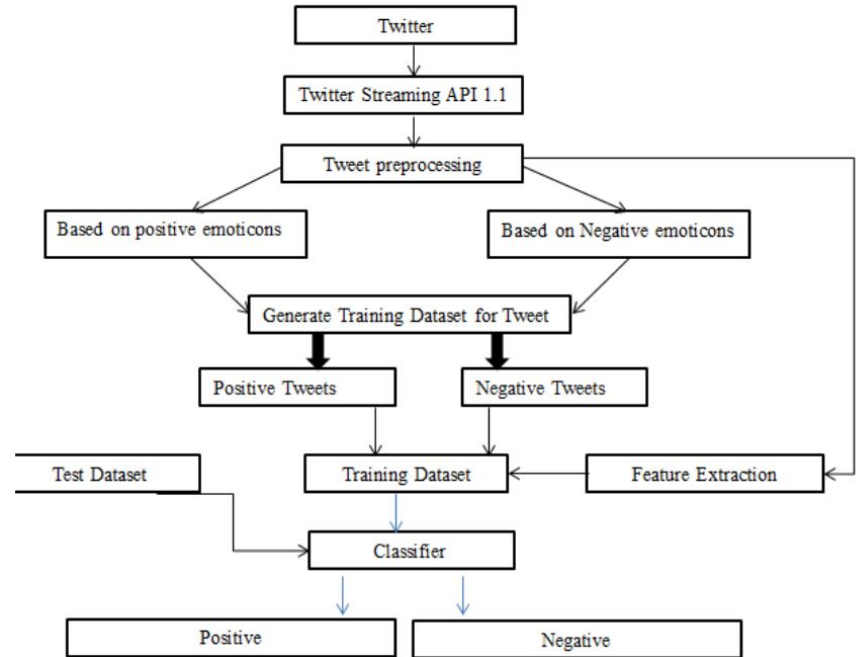


Fig.2 Sentiment Classification Based On Emoticons

Methodology



This project has been divided into 2 phases. First, literature study is conducted, followed by system development. Literature study involves conducting studies on various sentiment analysis techniques and method that currently in used.

In phase 2, application requirements and functionalities are defined prior to its development. Also, architecture and interface design of the program and how it will interact are also identified.

Reference:

- Ian H. Witten, Eibe Frank & Mark A. Hall. Data Mining – Practical Machine learning Tools and Techniques.
- Richard O. Duda, Peter E. Hart & David G. Stork: Pattern Classification.
- Steven Bird, Ewan Klein & Edward Loper. Natural Language Processing with Python.

Thank you

