SMART INTERNZ

PROJECT REPORT

Movie Review Analysis based on sentiment using Text-Processing API with IBM Cloud

ABSTRACT

Movie reviews help users decide if the movie is worth their time. A summary of all reviews for a movie can help users make this decision by not wasting their time reading all reviews. Movie-rating websites are often used by critics to post comments and rate movies which help viewers decide if the movie is worth watching. Sentiment analysis can determine the attitude of critics depending on their reviews. Sentiment analysis of a movie review can rate how positive or negative a movie review is and hence the overall rating for a movie. Therefore, the process of understanding if a review is positive or negative can be automated as the machine learns through training and testing the data. This project aims to rate reviews using two classifiers and compare which gives better and more accurate results. Classification is a data mining methodology that assigns classes to a collection of data in order to help in more accurate predictions and analysis. Naïve Bayes and decision tree classifications will be used and the results of sentiment analysis compared.

INTRODUCTION

The main aim of this project is to identify the underlying sentiment of a movie review on the basis of its textual information. In this project, we try to classify whether a person liked the movie or not based on the review they give for the movie. This is particularly useful in cases when the creator of a movie wants to measure its overall performance using reviews that critics and viewers are providing for the movie. The outcome of this project can also be used to create a recommender by providing recommendation of movies to viewers on the basis of their previous reviews. Another application of this project would be to find a group of viewers with similar movie tastes (likes or dislikes). As a part of this project, we aim to study several feature extraction techniques used in text mining e.g. keyword spotting, lexical affinity and statistical methods, and understand their relevance to our problem. In addition to feature extraction, we also look into different classification techniques and explore how well they perform for different kinds of feature representations. We finally draw a conclusion regarding which combination of feature representations and classification techniques are most accurate for the current predictive task.

MOVIE REVIEWS

Movie reviews are an important way to gauge the performance of a movie. While providing a Numerical/stars rating to a movie tells us about the success or failure of a movie quantitatively, a collection of movie reviews is what gives us a deeper qualitative insight on different aspects of the movie. A textual movie review tells us about the strong and weak points of the movie and deeper analysis of a movie review can tell us if the movie in general meets the expectations of the reviewer.

SENTIMENT ANALYSIS

Sentiment Analysis is a major subject in machine learning which aims to extract subjective information from the textual reviews. The field of sentiment of analysis is closely tied to natural language processing and text mining. It can be used to determine the attitude of the reviewer with respect to various topics or the overall polarity of review. Using sentiment analysis, we can find the state of mind of the reviewer while providing the review and understand if the person was “Positive”, “Negative”,” Neutral” and so on.

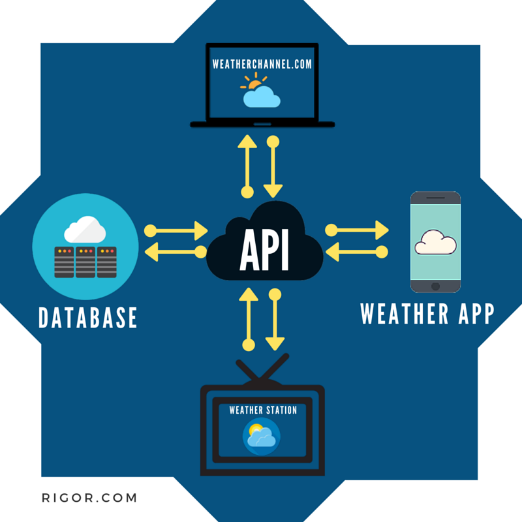
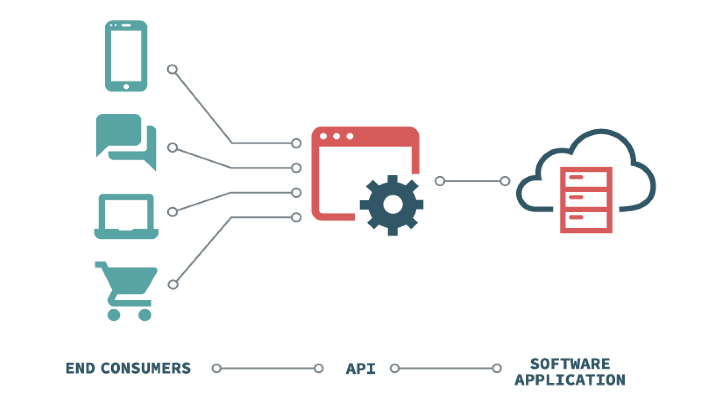
Sentiment analysis is also known as opinion mining or emotion AI. It refers to the use of natural language processing text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information. Sentiment analysis is widely applied to voice of the customer materials such as reviews and survey responses, online and social media, and healthcare materials for applications that range from marketing to customer service to clinical medicine.In this project we aim to use sentiment analysis on a set of movie reviews given by reviewers and try to understand what their overall reaction to the movie was, i.e. if they liked the movie or they hated It.We aim to utilize the relationships of the words in the review to predict the overall polarity of the review.



Application programming interface

In computing, an application programming interface (API) is an interface that defines interactions between multiple software applications or mixed hardware-software intermediaries

In building applications, an API (application programming interface) simplifies programming by abstracting the underlying implementation and only exposing objects or actions the developer needs.

**TEXT PROCESSING API** is a simple JSON over HTTP web service for text mining and natural language processing. It is currently free and open for public use without authentication, though that may change in the future.

**FLASK**

Flask is a micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions. However, Flask supports extensions that can add application features as if they were implemented in Flask itself. Extensions exist for object-relational mappers, form validation, upload handling, various open authentication technologies and several common framework related tools.

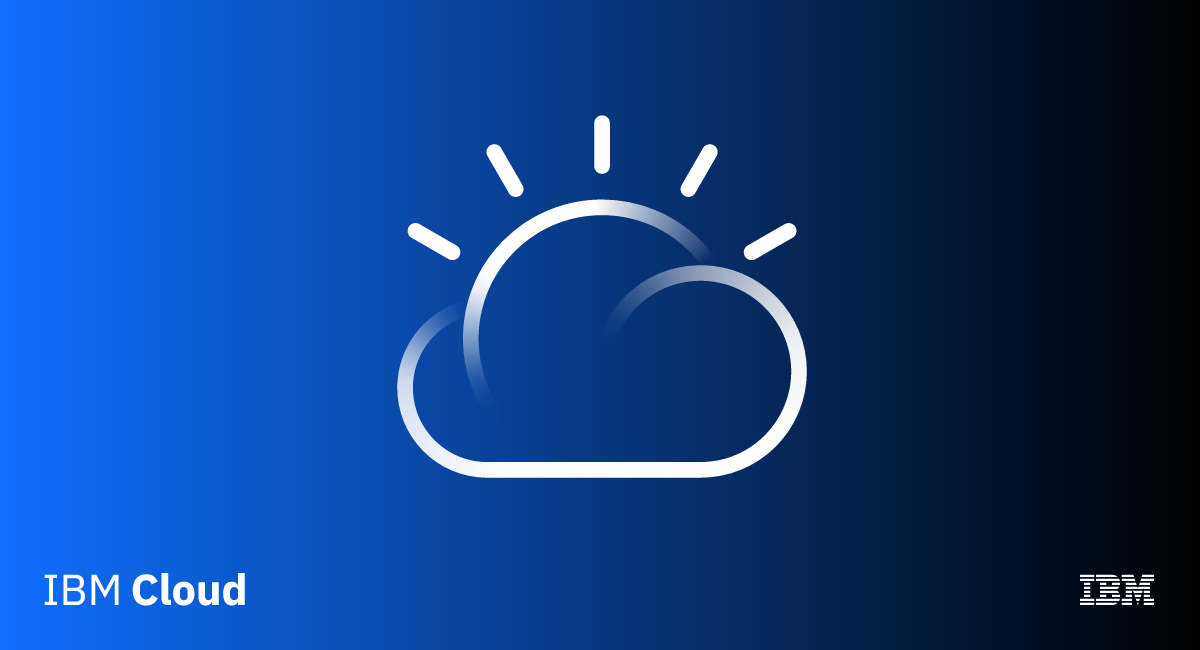


**HTML**

Web pages can be created and modified by using professional HTML editorsHTML allows users to create and structure sections, headings, links, paragraphs, and more, on a website using various tags and elements. Almost everything you want to create on a web page can be done using a specific HTML code.



**IBM**

IBM Cloud provides a full-stack, public cloud platform with various products in the catalog, including options for compute, storage, networking, end-to-end developer solutions for app development, testing and deployment, security management services, traditional and open source databases, and cloud-native services. Scalability. Cloud infrastructure scales on demand to support fluctuating workloads . Storage options. Users can choose public, private or hybrid storage offerings, depending on security needs and other considerations. Control choices. ... Tool selection. ... Security features.

**CONCLUSION**

This process of text processing, sentiment analysis is not only used for movie reviews but for other purposes like in taxi service apps like ola, uber, rapido, etc after completion of the ride we will be asked for a review of the driver, ride and the app. In online shopping sites we will be asked about the product. Now a days this process is entering our daily lives in every aspect.

From this project we have learnt about:

API

IBM CLOUD

FLASK APPLICATION

SENTIMENT ANALYSIS.

TEAM MEMBERS

20R15A0221 - THADEM HEMANTH KUMAR

20R15A0223 - THOUTA SHYAM SUNDER

20R15A0224 - AKIRI SAI SWETHA

20R15A0226 - BIJJULA PREM KUMAR REDDY

20R15A0227 - MODALI YAGNA LAKSHMI MADHURI