

# **“Sentiment Analysis - Detect the sentiment of Bangla text.”**

**Author,**  
**Shoaib Alam**  
**Reg. no : 2013331064**  
**Dept. Of CSE, SUST.**  
**Email : shoaibalam.cse@gmail.com**

# Abstract

Sentiment Analysis is an area of important research over the last decade. The basic task in sentiment analysis is classifying a given text. Sentiment analysis is an important type of text analysis process that aims to measure how much positive or negative a text is. Though Bangla is one of the top ten most spoken languages in the world and is spoken by more than 200 million people, it still lacks significant research efforts in the area of Bangla Sentiment Analysis. In the first phase of our thesis work we collected raw data of sentences, which were used as comments by users at Facebook.

In the later phases, we used different machine learning techniques to determine whether the given query is positive or negative.

# Overview

## Sentiment Analysis



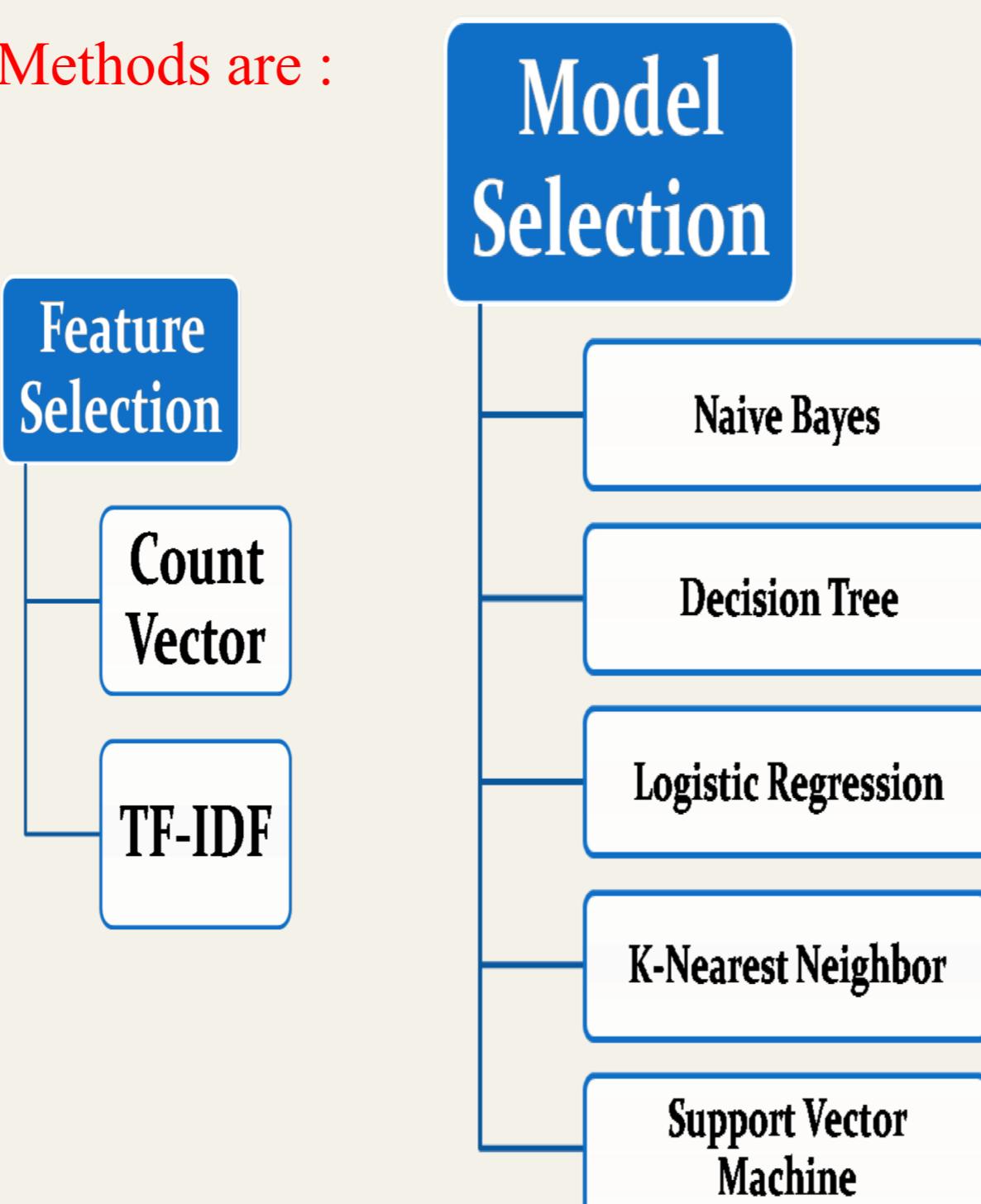
Estimate the  
acceptance of  
comment

Polarize human statement and build statistical report.

Mainly because of the web  
huge volumes of opinionated  
text

# Methodology

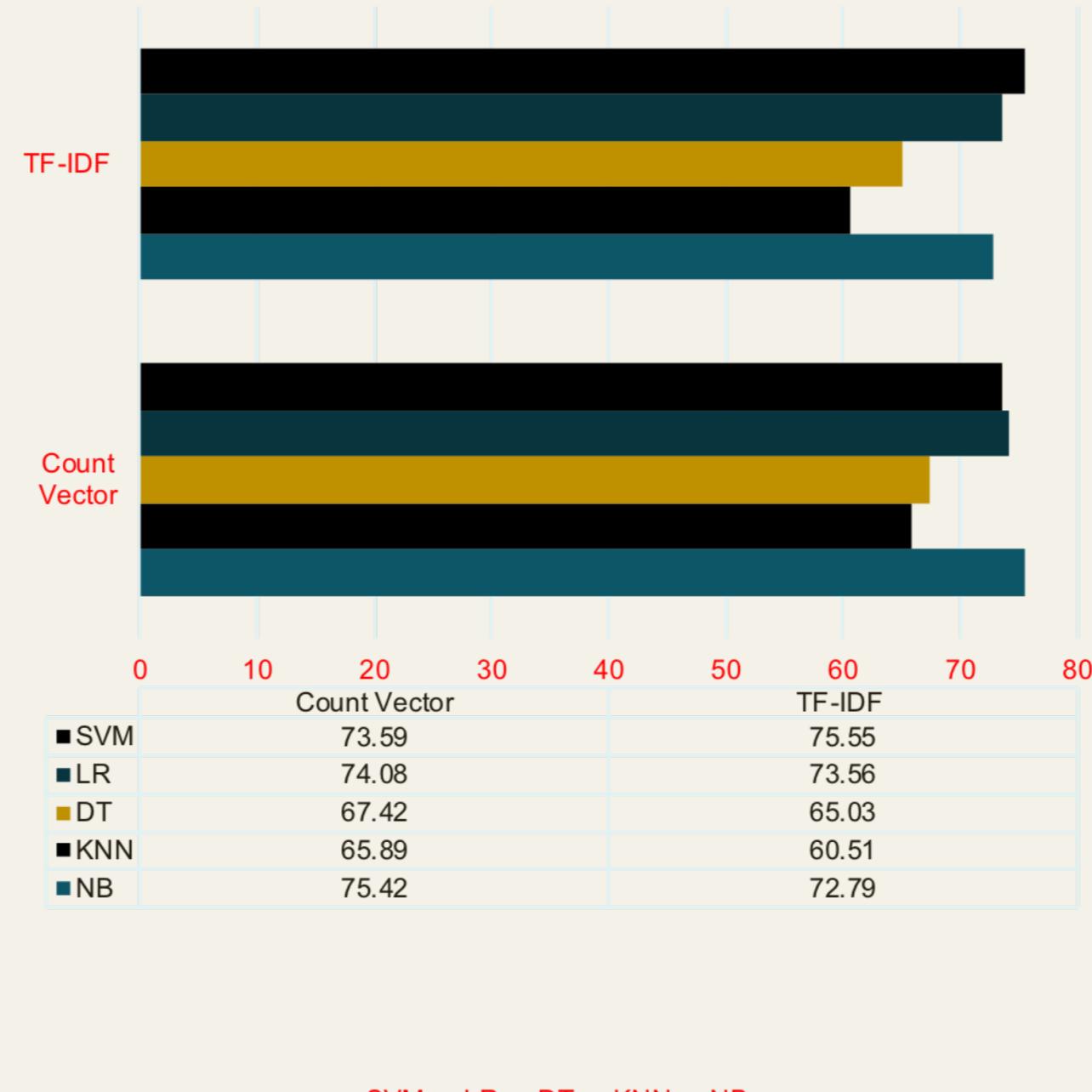
**Methods are :**



# Sample of Data Set (Present)

pos	পুরান কথে মনে পরে গোল..... ) ..... দেখতে হবে
pos	সাকিব ৮ বলে ১৮ করার পরও আমি সাকিবের পক্ষে।কাল দিনের প্রথম ওভারে ছক্কা মারতে গিয়ে আউট হলেও আমি সাকিবের পক্ষেই থাকবো।আমি অঙ্গ ক
pos	পৃথিবীর সকল নারীদের বিশ্ব নারী দিবসের শুভেচ্ছা ।
pos	প্রপ্রথিতীর সকল নারীর জন্য অত্যন্ত আনন্দের দিন, অত্যন্ত সম্মানের দিন। নারীর সাফল্যের স্বীকৃতির দিন।
pos	যুবহই ভালো উদ্যোগ। জুনিয়রদের প্রত্যেকটা কাজে আমাদের উৎসাহ দেওয়া উচিত। ভিডিওর লিংকটায় "প্রাউট টু বি কুয়েটিয়ান, হ্যাট্স অফ জুনিয়র্স" লিখে।
pos	ভোরের আলোয় ঘুজে নিয়ে চেনা ঠিকান!!!!!!
pos	শুভ জন্মদিন পিতা। বাঙালি জাতি তোমার কাছে চির কৃতজ্ঞ।
pos	অপৰ্যব অমিয়াধুম
pos	এমন সুন্দর জ্যোতিশাখা বসে সারাদিন কাটিয়ে দেয়াই যায় ^_^
pos	এইসব জ্যোতি পরিষ্কার-পরিষ্কৃত রাখা আমাদের দায়িত্ব।
pos	সকলের দোষা প্রাণী।
pos	মিথ তবু হার মাননেন না
neg	কালুরা এবার নুঙ্গি তুলে নাচতে থাকবে মনে হবে ওরা ওয়ার্ক কাপে প্রথমবার সেমিফাইনালে উঠেছে
neg	#মুণ্ডাফিজুর_রহমান আমাদের দেশ ও জাতীয় অহংকার। কিন্তু কিন্তু অসাধু ব্যক্তি, যাদের দেশ প্রেম নাই, তারা তাকে ভারতীয় দের দেওয়া নাম #The_Fizz ব
neg	একজন বাংলাদেশী হিসেবে আমি গবিংত আর বালকাতার পরাধীন ভিয়ারি বাঙালীরা কি বুঝবে স্বাধীনতার কি মূল্য। তোরা দিল্লির চামচামি কর নয়ত ভারতবর্ষ
neg	আমাদের মোন্টাফিজি কে শুধু রেকর্ডের খাতায় দেখা যায়, কিন্তু জাতীয় দলের হয়ে খেলা দেখা যায় না। অনেক টাঁ মোসির মত।
neg	মামিনুলকে ২০১৪, ১৫ তে সেরা টেস্ট ক্রিকেটার দেয়া হয়নি অর্থ সে ধাওয়ানের চেয়ে এগিয়ে ছিল...।
neg	আলহামদুলিল্লাহ, সুখবর, আরে ভাই সমস্যা ছিল ত এদের বোলিং আকশনের। শোধেরে নিয়েছে আলহামদুলিল্লাহ, কবে চিল্লাচিল্লি করা ধামাবেন বলেন ত?? দে
neg	ফহিমি ICC এর কথা কি আর বলি? তারা সুবিধাবাদী ইতিয়াকে Win করানোই তাদের মূল লক্ষ্য। কথায় আছে,, সংবিধান মানি কিন্তু তাল গাছ টাঁ আমার
neg	এখন সবই বৈধ। যখন আবার কোন বড় ট্রুম্পেটে ভারতের সাথে বাংলাদেশের খেলা পরবে তখন আবার রাতারাতি সব আবেধ হয়ে যাবে।
neg	আলহামদুলিল্লাহ, সুখবর, সেই সাথে তীব্র নিম্না ও ধিঙ্কার জনাই জারা বাংলাদেশ টিমকে হেঁয়ে করার জন্যে, এমন ঘূন্ঘ জবন্য ষড়যন্ত্র করেছিলো। ধানবাদ ত
neg	তাহলে অভিযোগকারী আবেধ..... এই দুটি প্রেসার তথ্য বাংলাদেশকে কে যারা এত প্রেসারে রাখলো আই সি সি তাদের কী করবে

# Performance Analysis



# Discussion

# **MD. Masum**

## **Associate Professor**

### **Dept. Of CSE, SUST.**

Focused mainly on the variety of words used in comments for each of the both kinds of labels.

Selecting count vector and TF-IDF as feature for both of them. The result showed that SVM with tf-idf performed really well on classifying the comments depending on sentiment.

# Future Scope

- (1) Sentiment analysis is the most interesting and newly emerged research topic.
  - (2) It will open a new door for the writers, bloggers, and businessman.
  - (3) One can easily know the percentage of product acceptance and make their strategy to improve the product quality.
  - (4) Our model can also be used to accurately predict stock market and other business strategical decisions.
  - (5) We think different type of Machine learning model will enforce the result to a Significant level

# Conclusion

In this Thesis Report, we have designed Sentiment Analyzer which detects the sentiment of a text or statement. We used several supervised machine learning methods. It gives us approximately satisfactory accuracy. The approach presented here is flexible and suggests promising avenues for further research. So, We conclude our hypothesis as one of the most flexible and accurate implemented model in the field of sentiment analysis.

