

**Minor Project -**

**News Daily**

Submitted By: Submitted To:

Rajat Juneja – 16103009 Dr. Archana Purwar

Mansi Mehta - 16103091

Ajay Kumar – 16103177

Chaitanya Kumar – 16103244

**Introduction to the Project:**

Life is more than a series of isolated incidents. Things relate to other things, problem continue, issue arise, persist, are sometime resolved. If we don’t know this intuitively most of us figure it out by the time we are adults that keeping the knowledge of all these happenings around not only keep us technologically advanced but also helps us taking measures to prevent us from getting into such a worse situation.

With our project News Daily we aim at keeping our users with the current on goings in the world. We provide them with a personalised interface to go through various categories of news including business, sports, technology, health, science, etc. With the personalised interface, comes the facilities of getting news recommended as per the interest of the logged in users.

Going through the news one is able to express his views by commenting on the respective article and can also enter into discussions with the people who are willing to join or are already the part of discussion through a live chat.

We also provide the feature of the comments being classified through sentiment analysis as positive or negative. We provide the web interface to the users where the website is built in a modular fashion using React.js and Node.js.

The following languages/technologies have been used to create this Project:

1. Node.js
2. React.js
3. Socket.io
4. Python
5. Flask
6. MongoDB

The news is categorised in six categories:

1. Technology
2. Sports
3. Business
4. Science
5. Health
6. Entertainment

And all these six categories have been further subdivided into different news sources for each category.

**Literature Survey:**

1. Source: Recommender Systems – The Text Book, Chapter 2

Link: <http://www.springer.com/978-3-319-29657-9>

Algorithm Used: Collaborative Filtering using K-Nearest Neighbours

Dataset: IMDB Movie Ratings for Various Users

Author: Charu C. Aggarwal

Advantages: Collaborative Filtering using K-Nearest Neighbours is advantageous as it considers the similarity between a set of users and recommends data based on the similarity between the users.

Disadvantages: The Efficiency in K-Nearest Neighbours is less as compared to other users and it does not take into consideration the variation in the data of the users.

1. Source: Sentiment Analysis with Long Short-Term Memory networks, Research Paper Business Analytics Vrije Universiteit Amsterdam

Dataset: 50,000 movie reviews from the IMDB Dataset.

Algorithm Used: Bidirectional Long-short Term Memory Networks(BiLSTMs)

Author: Fenna Miedema

Date: August 1, 2018

Advantages: Traditional neural networks fail to persist the previous events which might end up giving result based on only the present event

Whereas the bilstm is able to rememeber the previous events and forget the not so important events which help in analysing the entire sentence keeping the knowledge from scratch.

Disadvantages:The models may be overfitted.

1. Source: Medium – Building a Node.js Websocket Chat App with Socket.io and React

Author: Vincent Mühler

Algorithm Used: Server and client using web-sockets

Date: February 23, 2018

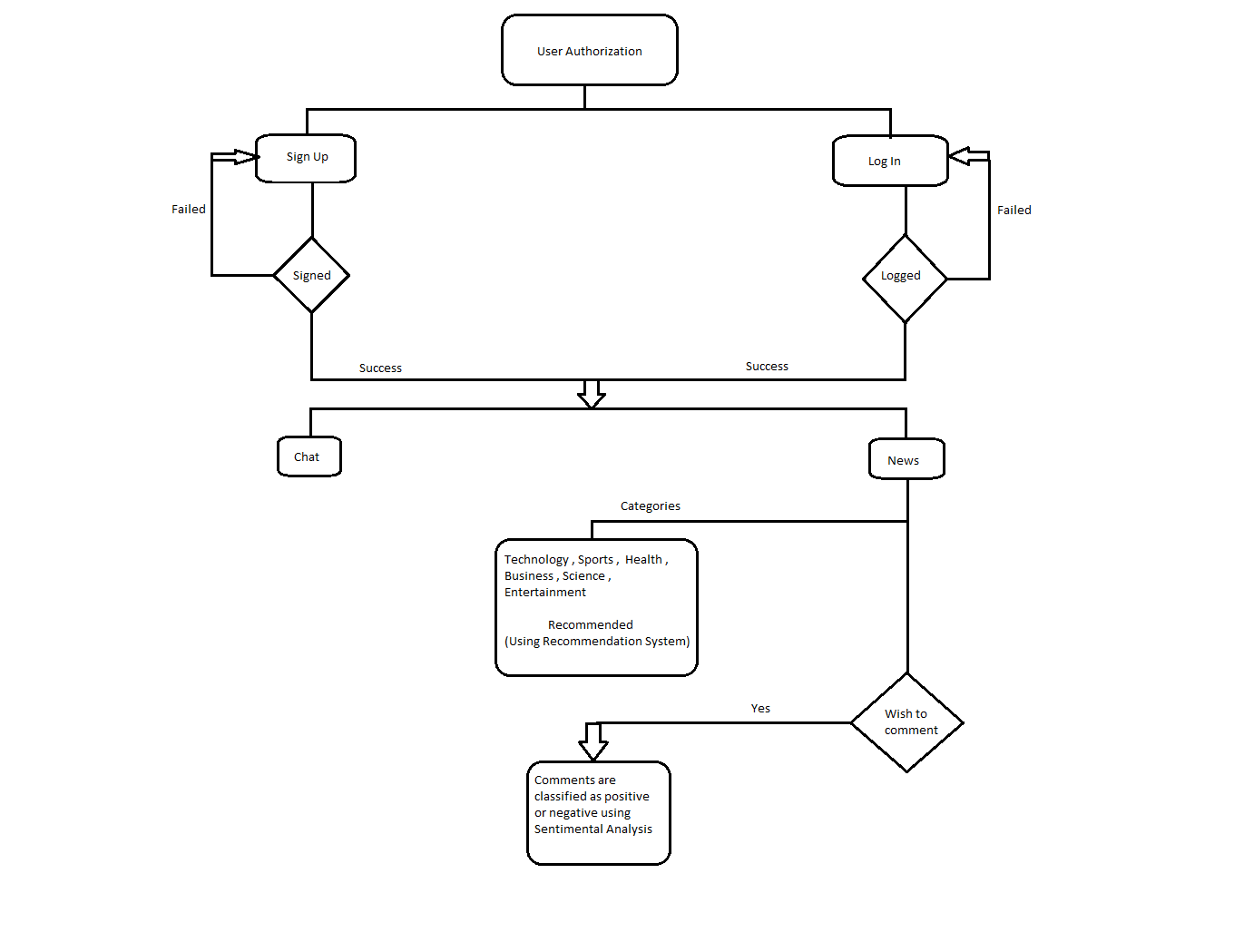
1. Source: Medium – Machine Learning for Recommender System Part-1 and Part-2

Author: Pavel Kordik

Date: June 4, 2018

Algorithm Used: Collaborative Filtering

**Flow Chart:**

****

**Module Description:**

* Sign Up : The user signs up and gives his information which is stored in the database and password is encrypted using hashing.
* Log In : The user logs in and backend system checks if the user exists or not.
* Chat : The user is able to discuss, and present his/her views regarding any topic or the topic going on currently in the discussion as multiple users can join the conversation.
* News : The user gets general news upfront. Later, he can choose from the categories and get the respective news. There’s also one section of recommended news, where user will get news based upon recommendation system. Comments can be added to the news where each comment will go through sentimental analysis and gives us the state of comment, which can either be positive or negative.

**Results:**

Datasets Used for Sentiment Analysis –

1. Twitter’s tweet dataset
2. Amazon Product Reviews
3. IMDB movie reviews

Total Count ~ 51000

Data Used for Recommendation System –

1. The count of articles read by users from each category of the provided news categories and sources.

Hardware Requirements – Working System with internet connection with basic requirements as follows:

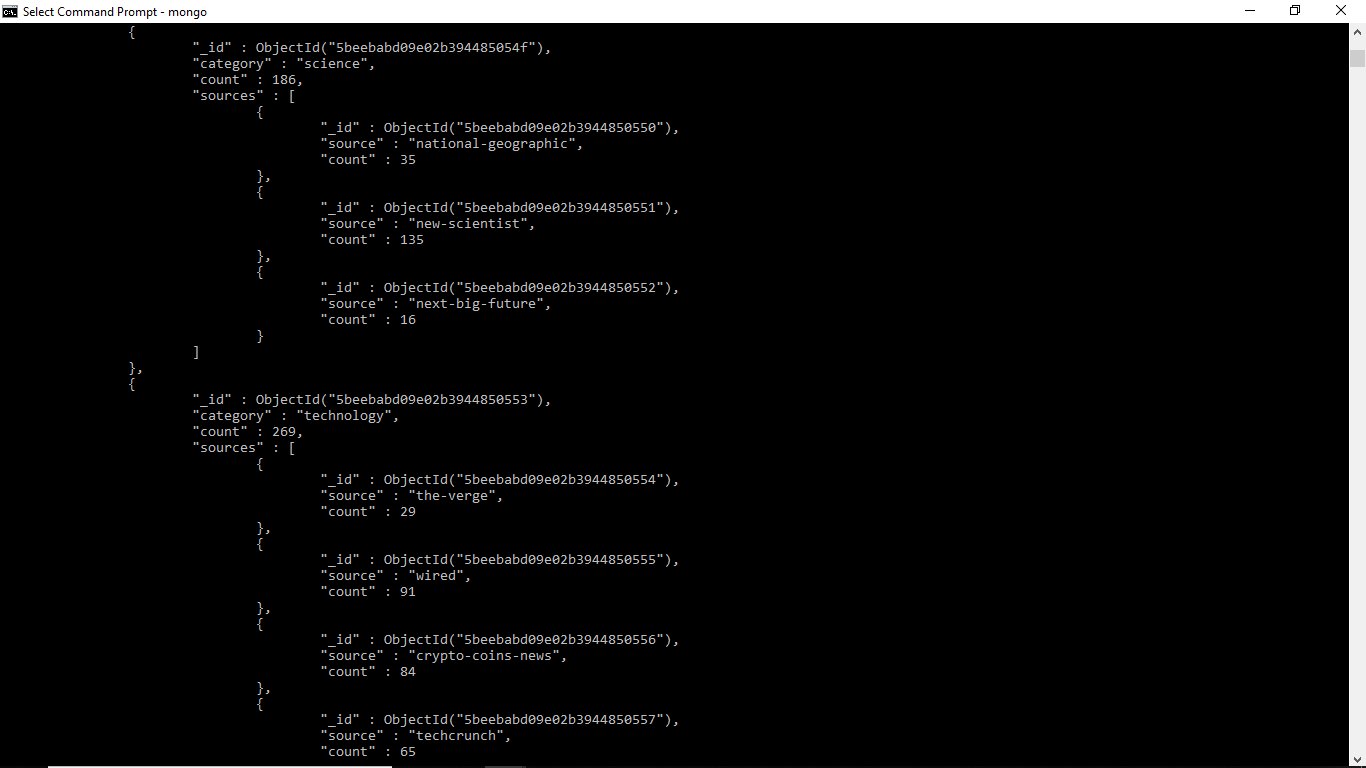
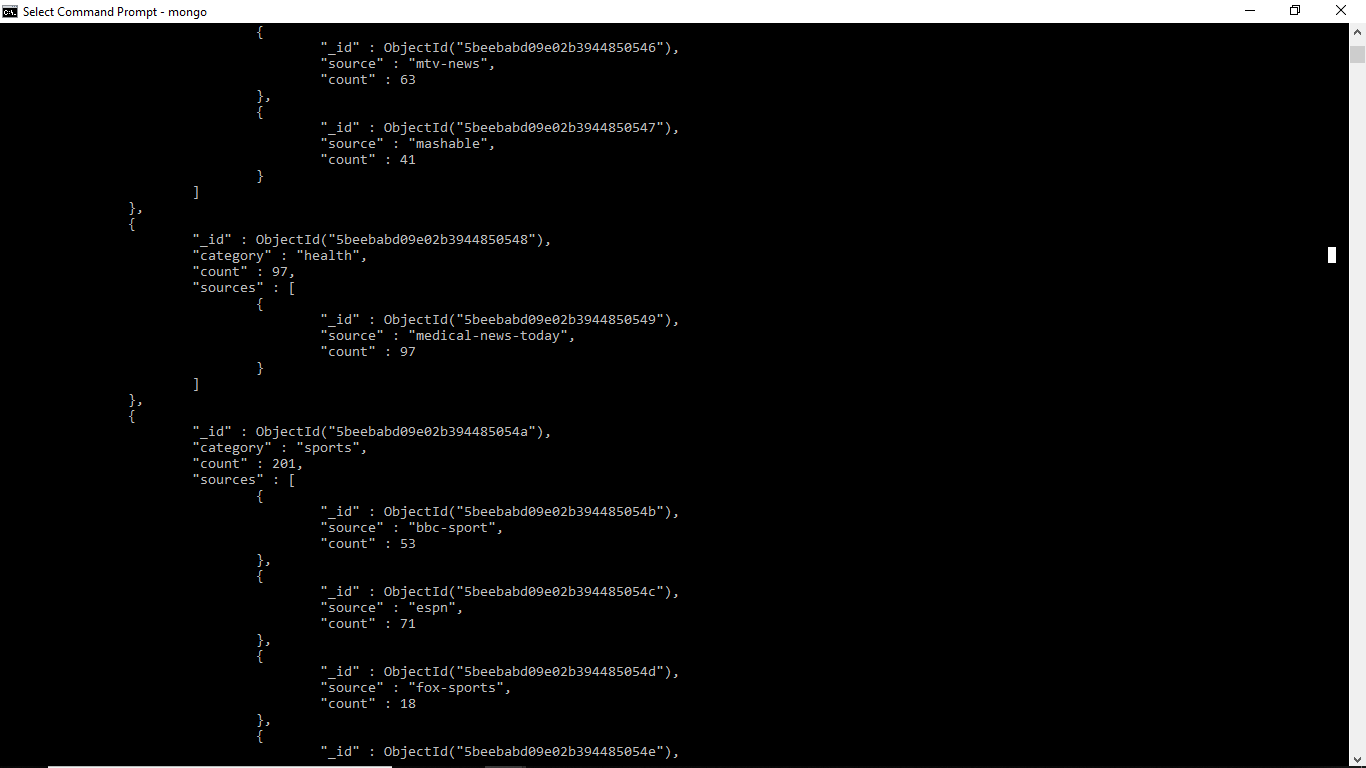
1. Processor – Intel Pentium CPU @ 1.60 GHz
2. System Type – 64 bit operating system
3. RAM – 4 gigabytes

Software Requirements -

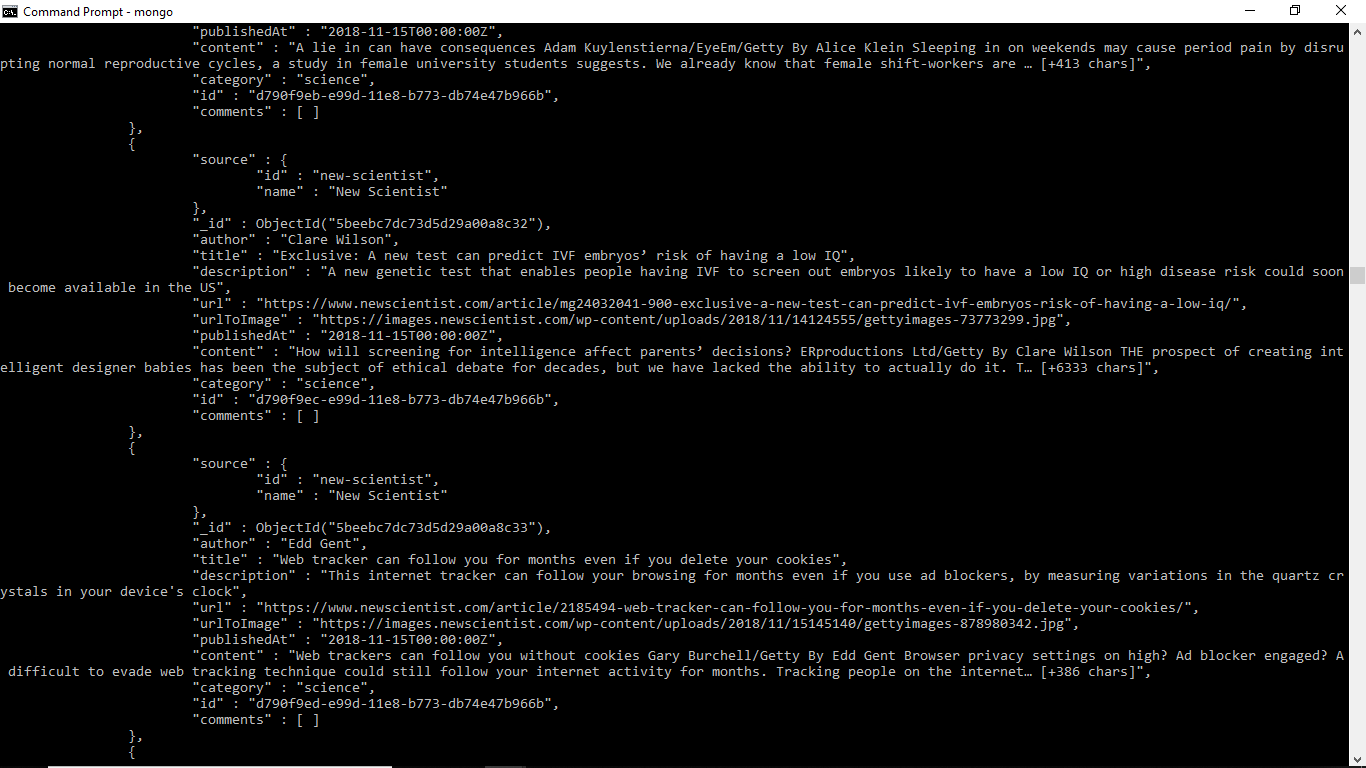
1. Visual Studio Code
2. Google Colab
3. Anaconda – Jupyter Notebook

**Screenshots:**

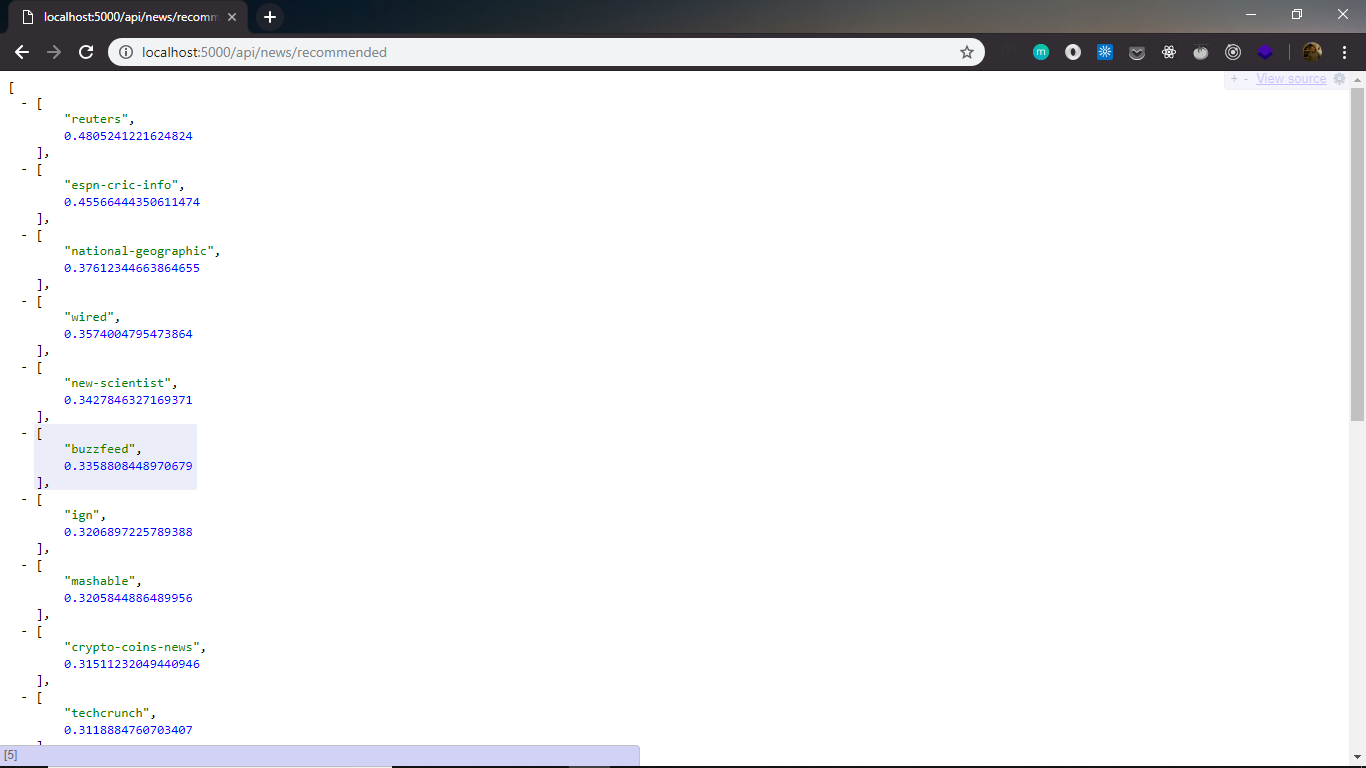
User Database –



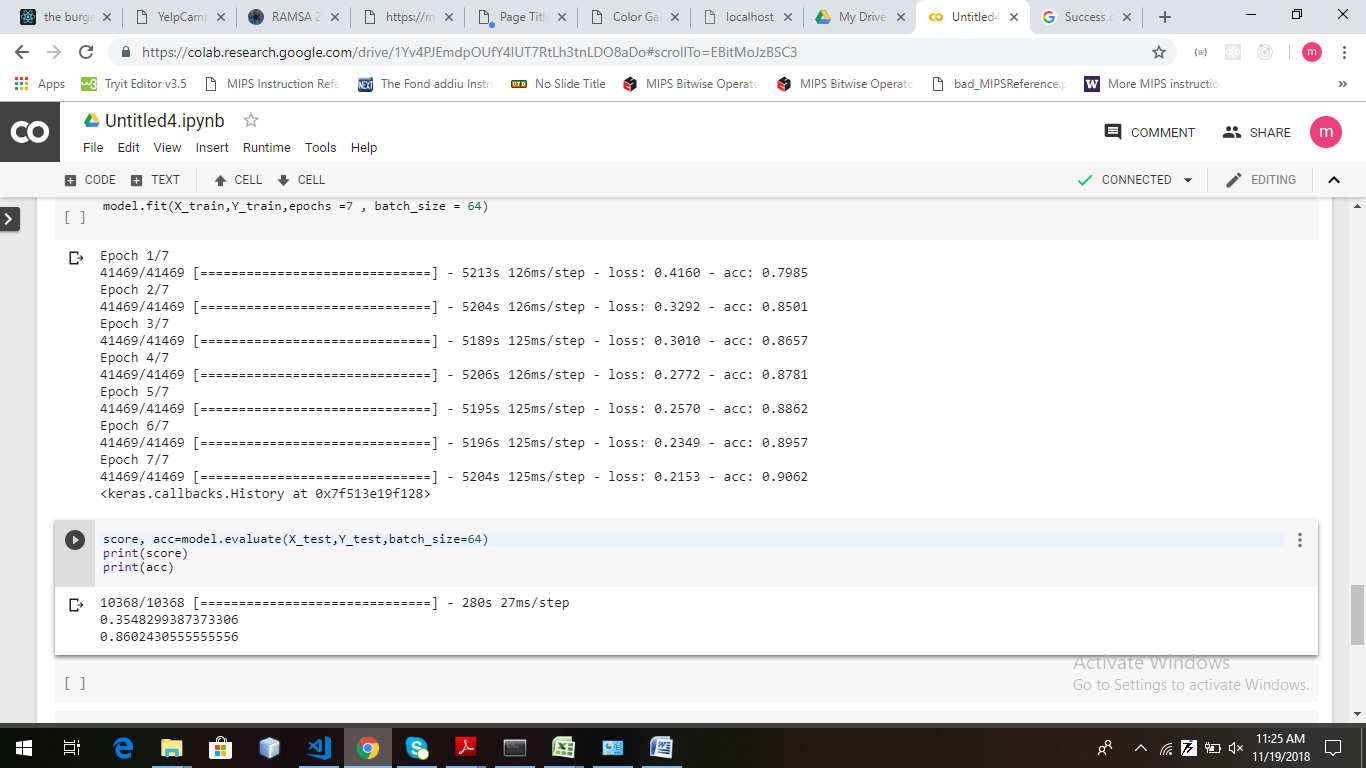
News Database –



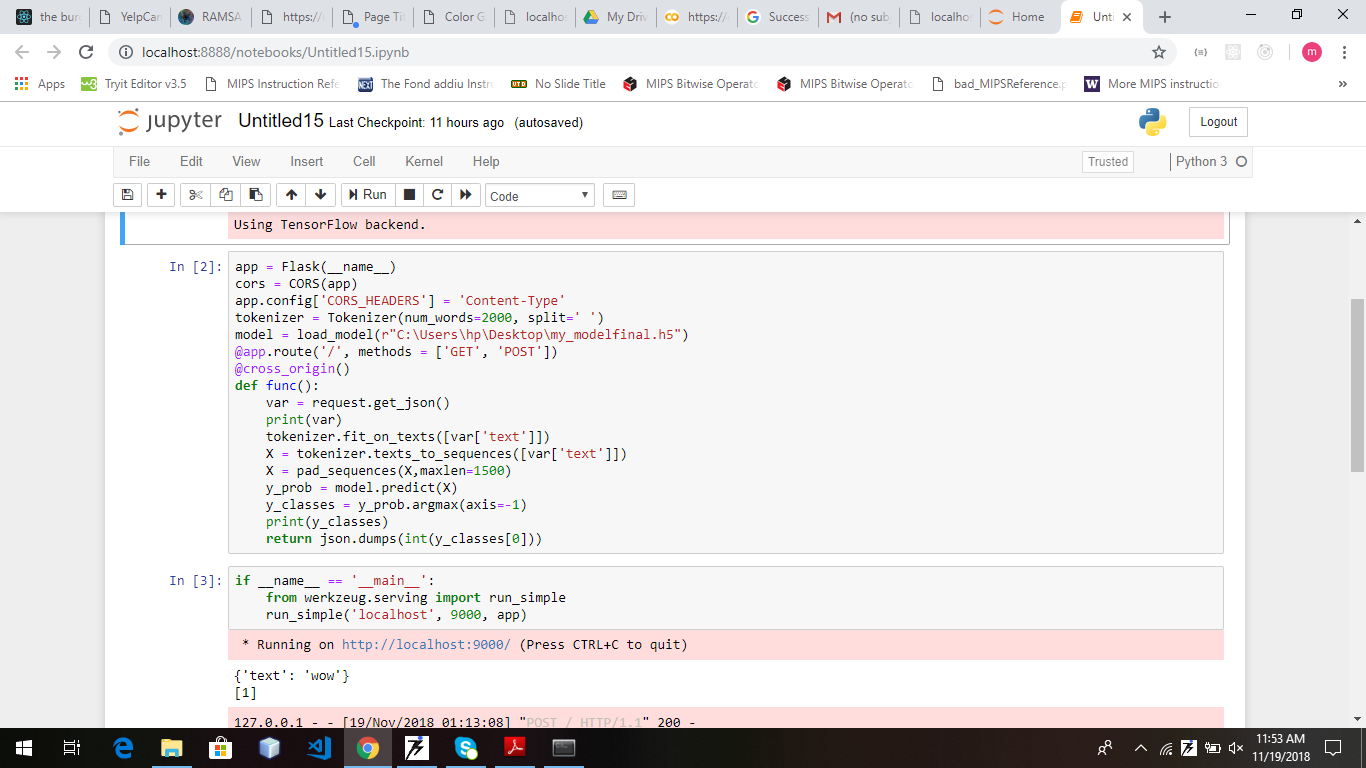
Recommendation System – The user id of the logged in user is provided as the input to the recommendation system and the probability of each news source which the user reads is provided as the output by the system.



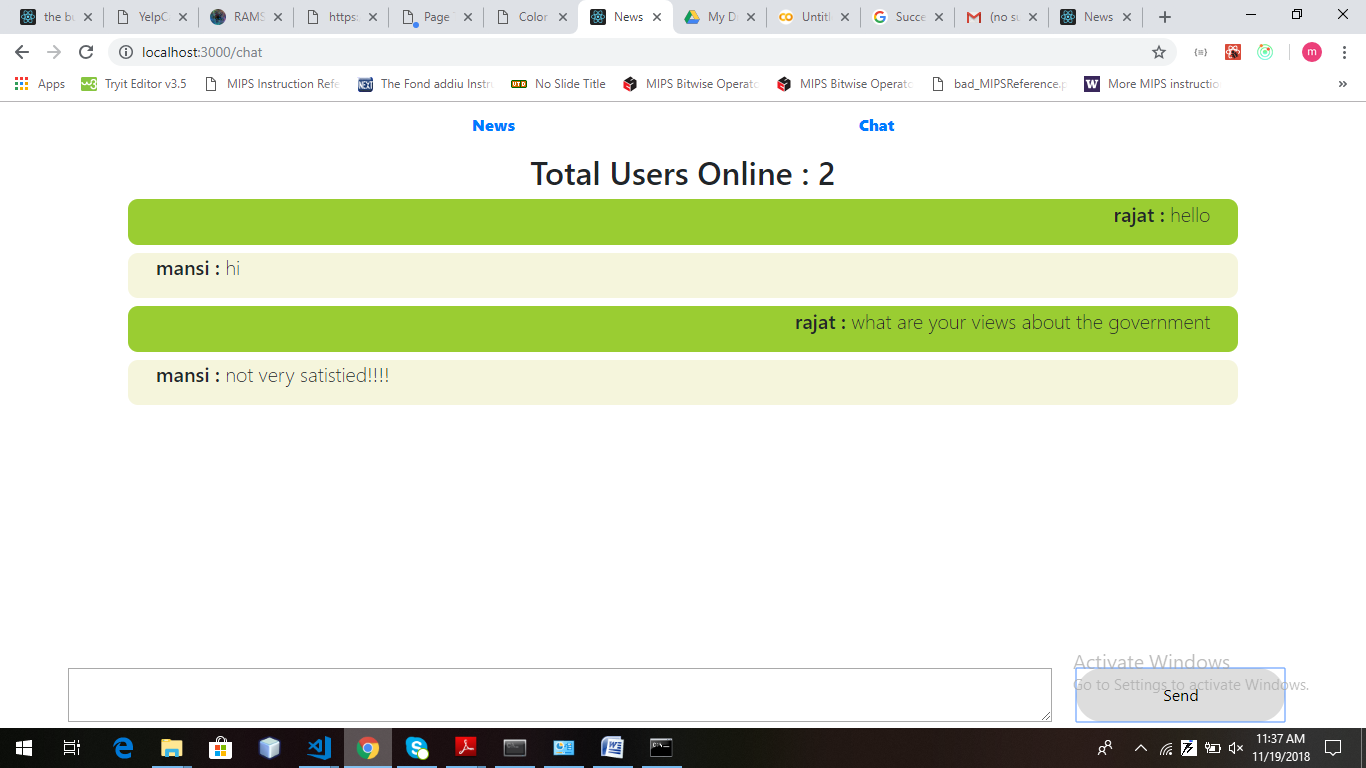
Sentiment Analysis – The model takes the dataset for training and is trained up to 7 epochs and the out is the trained model which is then tested and further and used to predict the polarity of the comment on the article.



The comment sent by the user is sent to the node.js backend and then further sent to a flask server on which we will predict the polarity of the comment and then send it back to the node server where it will be saved in the database with the predicted polarity.



Chat Application –



Test Case:

Out of 51000 sentences in our training dataset the program randomly chooses 10,000 sentences to test once the model is trained. The accuracy on this test set comes out to be 86%.

**Conclusion and Future Work:**

The project aimed at providing a personalised interface to the user helping him getting aware of the happenings around and help all to share views and discuss together and upto a certain level we have achieved the prime goal. With further research and studies we aim to improve the project with the following features-

1. We will be able to filter out the negative comments from the articles.
2. Provide a better recommendation system.
3. The data set will be much better.
4. Accuracy too will be tried to increased.

**References:**

* Stackoverflow
* Github
* Medium
* Geeksforgeeks
* W3Schools