MINOR PROJECT

NEWS RECOMMENDATION SYSTEM

MENTOR: DR. ANURAG SINGH BAGHEL

SUBMITTED BY:

DUSHYANT PRATAP SINGH (14/ICS/020)

NAMITA SINGH CHOUHAN (14/ICS/030)

PRATEEK KALRA (14/ICS/037)

SARTHAK TEOTIA (14/ICS/045)

TABLE OF CONTENT: PAGE NO.

INTRODUCTION

PURPOSE:

With the abundance of electronic information on the World Wide Web and ease to access such information online, users increasingly feel the need of automated systems that can help them navigate this huge information and media landscape. Be it music, articles,

products or books, there is vast amount of data online, and filtering it according to preference

at user level is difficult. Recommendation Systems have emerged as a solution to this information overload by providing users with personalized recommendation of content suited to their needs and preferences. A **recommender system** or a **recommendation system** (sometimes replacing "system" with a synonym such as platform or engine) is a subclass of information filtering system that seeks to predict the "rating" or "preference" that a user would give to an item. A Recommendation System takes into account the user preference and user's history of choice, and recommends similar content which may interest the user. Recommender systems have become increasingly popular in recent years, and are utilized in a variety of areas including movies, music, news, books, research articles, search queries, social tags, and products in general. A typical example of this information overload is the news industry, which seems to be turning fully online and trying to follow the developments in Web publishing. Most of the news publishers have introduced electronic versions of their content, making it easier for users to read and access latest news updates online. In general, people look for instantaneous access to fresh news updates. When surfing on the internet, user looks for interesting news among a large amount of news articles

In this project we have designed, analysed and implemented such a recommendation system for news articles. Almost all the online news service providers use one or other kind of recommendation system nowadays. The basic concept is the same - if we have history of user's previously read articles, we can use that information and decide if any new article will be of his interest or not. Based on these decisions, we can recommend new articles to the user. Our application is rather simple. We provide n article titles to the user, and based on his choice, we recommend m new articles, where n and m are user selected. The original n articles are picked

from test dataset, where topics are unknown at the moment, and m recommendations are given from entire dataset. User can select the articles based on title, and can read the text by further options.

SCOPE:

The targeted software product is a News Recommendation System. The system makes an extensive use of user data to come up with reasonable news predictions about user preferences and there are two main paradigms in this sense: Content-based and collaborative filtering. As a beginning, the scope of the system is scaled down to content based. Content-based filtering, also referred to as cognitive filtering, recommends items based on a comparison between the content of the items and a user profile. The content of each item is represented as a set of descriptors or terms, typically the words that occur in a document. The user profile is represented with the same terms and built up by analyzing the content of items which have been seen by the user. The system is aimed to have additional solutions for big data and sparse information. The test data will be provided from BBC NEWS dataset. The test data for the system are planned to contain millions of users and irregular data. The project limits the improvement of the recommendation system around these challenges and favours accuracy over performance while giving the final suggestions. As a milestone, the system should follow the methods to enhance accuracy, the performance enhancements are not included in the scope. The Recommender System will be a desktop application and the end users will be directly exposed to the suggestions made for them.

DEFINITIONS, ACRONYMS AND ABBREVIATIONS:

1.) **USER PROFILE:** The implementation of an information filtering system requires the creation of a user model, and subsequently a user profile. *User* modeling is defined as a discipline that deals with both how information about the user can be acquired and used by an automated system.

The description of what information is of interest to a user is commonly referred to as a **user profile**.

2.) **RECOMMENDATION SYSTEM:** Recommender system is defined as a decision making strategy for users under complex information environments.

3.) **RECOMMENDER PERSISTENCE** – In some situations, it is more effective to re-show recommendations, or let users re-rate items, than showing new items. There are several reasons for this. Users may ignore items when they are shown for the first time, for instance, because they had no time to inspect the recommendations carefully.

4.) **LABELLING** – User satisfaction with recommendations may be influenced by the labelling of the recommendations. For instance, in a study click-through rate (CTR) for recommendations labeled as "Sponsored" were lower (CTR=5.93%) than CTR for identical recommendations labeled as "Organic" (CTR=8.86%). Interestingly, recommendations with no label performed best (CTR=9.87%) in that study.

5.) **RISK DEFINITION-**The risk in recommender systems is the possibility to disturb or to upset the user which leads to a bad answer of the user.

6.) **CONTENT-BASED FILTERING -** Content-based filtering methods are based on a description of the item and a profile of the user's preferences. In a content-based recommender system, keywords are used to describe the items and a user profile is built to indicate the type of item this user likes.

7.) TRUST – A recommender system is of little value for a user if the user does not trust the system. Trust can be built by a recommender system by explaining how it generates recommendations, and why it recommends an item.

8.) PREFERENCE- Provide personalised and relevant recommendations to users according to their preference are the ultimate aim of recommendation systems.

Data Type	Example
Demographic	Age, marital status, gender etc.kk
Geographic	Location, city, country etc
Psychographic	Stable: interests, lifestyle, personality etc. Fluid: mood, attitude, opinions etc.

Generally, a news recommender system consists of three key components: users, items and user-item matching algorithms. So, the background of our product is becoming more important than foreground component. At the background, the system will gather data from both users and items. To be more specific and precise, those data can be classified as follows:

● User Modeling: Suggest that user data can be categorized into 2 domains: User Profile Data and User Reading Experience Data

1.) USER PROFILE DATA

Sports	Sports enthusiasts, interested in knowing updates in the field of sports
Business	Actively indulged in business
Entertainment	Easy going readers ,reading content just for fun or out of curiosity
National	News for the common man and concerns almost everyone
International	Competitive people
Tech-news	Concerns the techies and the nerds

2.) USER READING EXPERIENCE DATA

● Item Modeling: The second component of recommender systems is news item. The news metadata classified into three categories as in table below:

Topic metadata (TM)	Sports, national, international, entertainment, etc
Area metadata (AM)	Global, local, etc

DESIGN / MODEL:

LOGIN PAGE

SIGNUP PAGE

(7)

MAIN PAGE

IMPLEMENTATION:

TF-IDF based Recommender System

Recommender System based on tf-idf as vector representation of documents

TF-IDF Based Recommender

1. Represent articles in terms of bag of words
2. Represent user in terms of read articles associated words
3. Generate TF-IDF matrix for user read articles and unread articles
4. Calculate cosine similarity between user read articles and unread articles
5. Get the recommended articles

Describing parameters:

1. PATH_NEWS_ARTICLES: specify the path where news_article.csv is present
2. ARTICLES_READ: List of Article_Ids read by the user
3. NO_RECOMMENDED_ARTICLES: Refers to the number of recommended articles as a result

```
In [1]:  PATH_NEWS_ARTICLES="/Users/Dell/Music/articles1.csv"
         ARTICLES_READ=[2,7,8,17,18,34]
         NUM_RECOMMENDED_ARTICLES=50
```

```
In [4]:  try:
             import numpy
             import pandas as pd
             import pickle as pk
             from sklearn.feature_extraction.text import TfidfVectorizer
             from sklearn.metrics.pairwise import cosine_similarity
             import re
             from nltk.stem.snowball import SnowballStemmer
             import nltk
             stemmer = SnowballStemmer("english")
         except ImportError:
             print('You are missing some packages! ' \
                   'We will try installing them before continuing!')
             !pip install "numpy" "pandas" "sklearn" "nltk"
             import numpy
             import pandas as pd
             import pickle as pk
             from sklearn.feature_extraction.text import TfidfVectorizer
             from sklearn.metrics.pairwise import cosine_similarity
             import re
             from nltk.stem.snowball import SnowballStemmer
             import nltk
             stemmer = SnowballStemmer("english")
             print('Done!')
```

(10)

1. Represent articles in terms of bag of words

1. Reading the csv file to get the Article id, Title and News Content
2. Remove punctuation marks and other symbols from each article
3. Tokenize each article
4. Stem token of every article

```
In [5]: news_articles = pd.read_csv(PATH_NEWS_ARTICLES)
        news_articles.head()
```

Out[5]:

	id	article_no	user rating	login	title	publication	author	date	year	month	url	content	Unnamed: 12
0	0	17283	1	886	House Republicans Fret About Winning Their Hea...	New York Times	Carl Hulse	2016-12-31	2016	12	NaN	WASHINGTON — Congressional Republicans have...	NaN
1	1	17284	2	777	Rift Between Officers and Residents as Killing...	New York Times	Benjamin Mueller and Al Baker	2017-06-19	2017	6	NaN	After the bullet shells get counted, the blood...	NaN
2	2	17285	0	915	Tyrus Wong, 'Bambi' Artist Thwarted by Racial ...	New York Times	Margalit Fox	2017-01-06	2017	1	NaN	When Walt Disney's "Bambi" opened in 1942, cri...	NaN
3	3	17286	3	793	Among Deaths in 2016, a Heavy Toll in Pop Musi...	New York Times	William McDonald	2017-04-10	2017	4	NaN	Death may be the great equalizer, but it isn't...	NaN
4	4	17287	0	335	Kim Jong-un Says North Korea Is Preparing to T...	New York Times	Choe Sang-Hun	2017-01-02	2017	1	NaN	SEOUL, South Korea — North Korea's leader, ...	NaN

(11)

```
In [6]:  #Select relevant columns and remove rows with missing values
         news_articles = news_articles[['id','title','content']].dropna()
         #articles is a list of all articles
         articles = news_articles['content'].tolist()
         articles[0] #an uncleaned article
```

Out[6]: 'WASHINGTON — Congressional Republicans have a new fear when it comes to their   health care lawsuit against the
        Obama administration: They might win. The incoming Trump administration could choose to no longer defend the executiv
        e branch against the suit, which challenges the administration's authority to spend billions of dollars on health ins
        urance subsidies for   and   Americans, handing House Republicans a big victory on   issues. But a sudden loss of th
        e disputed subsidies could conceivably cause the health care program to implode, leaving millions of people without a
        ccess to health insurance before Republicans have prepared a replacement. That could lead to chaos in the insurance m
        arket and spur a political backlash just as Republicans gain full control of the government. To stave off that outcom
        e, Republicans could find themselves in the awkward position of appropriating huge sums to temporarily prop up the Ob
        ama health care law, angering conservative voters who have been demanding an end to the law for years. In another twi
        st, Donald J. Trump's administration, worried about preserving executive branch prerogatives, could choose to fight i
        ts Republican allies in the House on some central questions in the dispute. Eager to avoid an ugly political pileup,
        Republicans on Capitol Hill and the Trump transition team are gaming out how to handle the lawsuit, which, after the
        election, has been put in limbo until at least late February by the United States Court of Appeals for the District o
        f Columbia Circuit. They are not yet ready to divulge their strategy. "Given that this pending litigation involves th
        e Obama administration and Congress, it would be inappropriate to comment," said Phillip J. Blando, a spokesman for t
        he Trump transition effort. "Upon taking office, the Trump administration will evaluate this case and all related asp
        ects of the Affordable Care Act. " In a potentially   decision in 2015, Judge Rosemary M. Collyer ruled that House Re
        publicans had the standing to sue the executive branch over a spending dispute and that the Obama administration had
        been distributing the health insurance subsidies, in violation of the Constitution, without approval from Congress. T
        he Justice Department, confident that Judge Collyer's decision would be reversed, quickly appealed, and the subsidies
        have remained in place during the appeal. In successfully seeking a temporary halt in the proceedings after Mr. Trump
        won, House Republicans last month told the court that they "and the 's transition team currently are discussing pote
        ntial options for resolution of this matter, to take effect after the 's inauguration on Jan. 20, 2017. " The suspen
        sion of the case, House lawyers said, will "provide the   and his future administration time to consider whether to c
        ontinue prosecuting or to otherwise resolve this appeal. " Republican leadership officials in the House acknowledge t
        he possibility of "cascading effects" if the   payments, which have totaled an estimated $13 billion, are suddenly st
        opped. Insurers that receive the subsidies in exchange for paying   costs such as deductibles and   for eligible con
        sumers could race to drop coverage since they would be losing money. Over all, the loss of the subsidies could destab
        ilize the entire program and cause a lack of confidence that leads other insurers to seek a quick exit as well. Antic
        ipating that the Trump administration might not be inclined to mount a vigorous fight against the House Republicans g

(12)
```

```python
In [7]: def clean_tokenize(document):
            document = re.sub('[^\w_\s-]', ' ',document)   #remove punctuation marks and other symbols
            tokens = nltk.word_tokenize(document)          #Tokenize sentences
            cleaned_article = ' '.join([stemmer.stem(item) for item in tokens])   #Stemming each token
            return cleaned_article
```

```python
In [8]: cleaned_articles = list (map(clean_tokenize, articles[0:50]))
        cleaned_articles   #a cleaned, tokenized and stemmed article
```

Out[8]: ['washington congression republican have a new fear when it come to their health care lawsuit against the obama adm inistr they might win the incom trump administr could choos to no longer defend the execut branch against the suit which challeng the administr s author to spend billion of dollar on health insur subsidi for and american hand hous republican a big victori on issu but a sudden loss of the disput subsidi could conceiv caus the health care program to implod leav million of peopl without access to health insur befor republican have prepar a replac that could lea d to chao in the insur market and spur a polit backlash just as republican gain full control of the govern to stave off that outcom republican could find themselv in the awkward posit of appropri huge sum to temporarili prop up the obama health care law anger conserv voter who have been demand an end to the law for year in anoth twist donald j t rump s administr worri about preserv execut branch prerog could choos to fight it republican alli in the hous on so me central question in the disput eager to avoid an ugli polit pileup republican on capitol hill and the trump tran sit team are game out how to handl the lawsuit which after the elect has been put in limbo until at least late febr uari by the unit state court of appeal for the district of columbia circuit they are not yet readi to divulg their strategi given that this pend litig involv the obama administr and congress it would be inappropri to comment said phillip j blando a spokesman for the trump transit effort upon take offic the trump administr will evalu this case and all relat aspect of the afford care act in a potenti decis in 2015 judg rosemari m collyer rule that hous repub lican had the stand to sue the execut branch over a spend disput and that the obama administr had been distribut th e health insur subsidi in violat of the constitut without approv from congress the justic depart confid that judg c ollyer s decis would be revers quick appeal and the subsidi have remain in place dure the appeal in success seek a temporari halt in the proceed after mr trump won hous republican last month told the court that they and the s tran sit team current are discuss potenti option for resolut of this matter to take effect after the s inaugur on jan 20

(13)

2. Represent user in terms of read articles associated words

```
In [9]:  #Get user representation in terms of words associated with read articles
         user_articles = ' '.join(cleaned_articles[i] for i in ARTICLES_READ)
```

```
In [10]:  user_articles
```

Out[10]: 'when walt disney s bambi open in 1942 critic prais it spare haunt visual style vast differ from anyth disney had d
one befor but what they did not know was that the film s strike appear had been creat by a chines immigr artist who
took as his inspir the landscap paint of the song dynasti the extent of his contribut to bambi which remain a mark
for film anim would not be wide known for decad like the film s titl charact the artist tyrus wong weather irrevoc
separ from his mother and in the hope of make a life in america incarcer isol and rigor interrog all when he was st
ill a child in the year that follow he endur poverti discrimin and chronic lack of recognit not onli for his work a
t disney but also for his fine art befor find acclaim in his 90s mr wong die on friday at 106 a hollywood studio ar
tist painter printmak calligraph illustr and in later year maker of fantast kite he was one of the most celebr arti
st of the 20th centuri but becaus of the margin to which were long subject he pass much of his career unknown to th
e general public artist recognit when mr wong did find it was all the more noteworthi for the fact that among chine
s immigr men of his generat profession prospect were larg limit to menial job like houseboy and laundryman train as
a painter mr wong was a lead figur in the modernist movement that flourish in california between the first and seco
nd world war in 1932 and again in 1934 his work was includ in group show at the art institut of chicago that also f
eatur picasso matiss and paul klee as a staff artist for hollywood studio from the 1930s to the 1960s he drew story
board and made vibrant paint as detail as ani architectur illustr that help the director envis each scene befor it
was shot over the year his work inform the look of anim pictur for disney and film for warner brother and other stu
dio among them the sand of iwo jima 1949 rebel without a caus 1955 and the wild bunch 1969 but of the dozen of film
on which he work it was for bambi that mr wong was belat most renown he was truli involv with everi phase of produc
t john canemak an anim and a historian of anim at new york univers said in an interview for this obituari in march
he great an art direct that had realli never been seen befor in anim in 2013 and 2014 mr wong was the subject of wa

(14)

3. Generate TF-IDF matrix for user read articles and unread articles

In [11]:
```python
#Generate tfidf matrix model for entire corpus
tfidf_matrix = TfidfVectorizer(stop_words='english', min_df=2)
article_tfidf_matrix = tfidf_matrix.fit_transform(cleaned_articles)
article_tfidf_matrix #tfidf vector of an article
```

Out[11]: <50x2926 sparse matrix of type '<class 'numpy.float64'>'
           with 15553 stored elements in Compressed Sparse Row format>

In [12]:
```python
#Generate tfidf matrix model for read articles
user_article_tfidf_vector = tfidf_matrix.transform([user_articles])
user_article_tfidf_vector
```

Out[12]: <1x2926 sparse matrix of type '<class 'numpy.float64'>'
           with 1423 stored elements in Compressed Sparse Row format>

In [13]:
```python
user_article_tfidf_vector.toarray()
```

Out[13]: array([[0.04243785, 0.02178228, 0.02468441, ..., 0.00773386,
           0.   , 0.01546771]])

(15)

4. Calculate cosine similarity between user read articles and unread articles

```
In [14]: articles_similarity_score=cosine_similarity(article_tfidf_matrix, user_article_tfidf_vector)
```

```
In [15]: recommended_articles_id = articles_similarity_score.flatten().argsort()[::-1]
```

```
In [16]: recommended_articles_id
```

```
Out[16]: array([ 7,  2, 17,  8, 18, 26,  9, 38, 34, 37, 27,  3, 43, 12, 40, 29, 22,
                 1, 23, 21, 30, 10, 31, 46, 44, 24,  6, 45, 42, 15, 16, 32, 48, 13,
                41, 36, 28, 39, 49, 33,  4, 11, 19, 20, 25,  0, 14, 35, 47,  5], dtype=int64)
```

```
In [17]: #Remove read articles from recommendations

        final_recommended_articles_id = [article_id for article_id in recommended_articles_id
                                         if article_id not in ARTICLES_READ ][:NUM_RECOMMENDED_ARTICLES]
```

(16)

5. Get the recommended articles

In [18]: `final_recommended_articles_id`

Out[18]: [26,
 9,
 38,
 37,
 27,
 3,
 43,
 12,
 40,
 29,
 22,
 1,
 23,
 21,
 30,
 10,
 31,
 46,
 44,
 24,
 6,
 45,
 42,
 15,
 16,
 32,
 48,
 13,
 41,
 36,
 28,

(17)

```
39,
49,
33,
4,
11,
19,
20,
25,
0,
14,
35,
47,
5]
```

In [19]:
```python
#Recommended Articles and their title

print ('Articles Read')

print (news_articles.loc[news_articles['id'].isin(ARTICLES_READ)]['title'])

print ('\n')

print ('Recommender ')

print (news_articles.loc[news_articles['id'].isin(final_recommended_articles_id)]['title'])
```

```
Articles Read

2    Tyrus Wong, 'Bambi' Artist Thwarted by Racial ...

7    After 'The Biggest Loser,' Their Bodies Fought...

8    First, a Mixtape. Then a Romance. - The New Yo...

17   Modi's Cash Ban Brings Pain, but Corruption-We...

18   Suicide Bombing in Baghdad Kills at Least 36 -...

34   Riot by Drug Gangs in Brazil Prison Leaves at ...

Name: title, dtype: object
```

(18)

```
Recommender
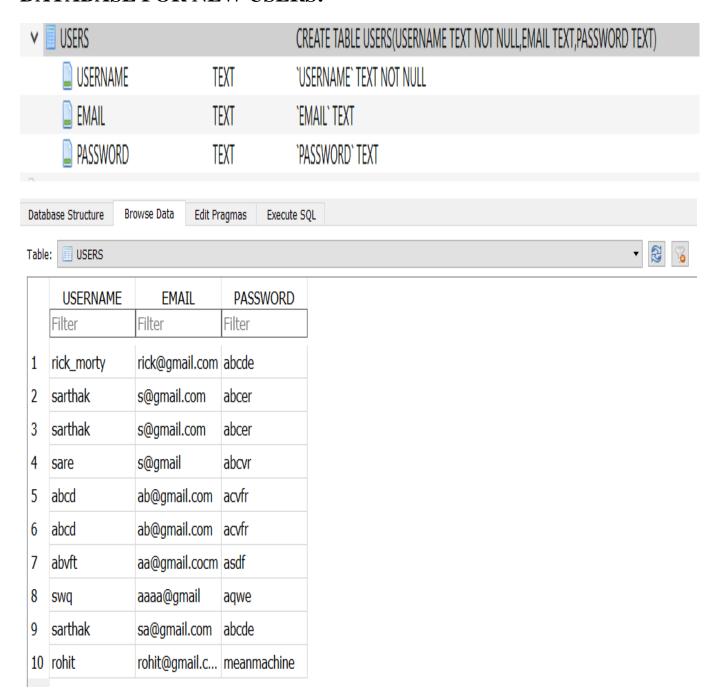0     House Republicans Fret About Winning Their Hea...
1     Rift Between Officers and Residents as Killing...
3     Among Deaths in 2016, a Heavy Toll in Pop Musi...
4     Kim Jong-un Says North Korea Is Preparing to T...
5     Sick With a Cold, Queen Elizabeth Misses New Y...
6     Taiwan's President Accuses China of Renewed In...
9     Calling on Angels While Enduring the Trials of...
10    Weak Federal Powers Could Limit Trump's Climat...
11    Can Carbon Capture Technology Prosper Under Tr...
12    Mar-a-Lago, the Future Winter White House and ...
13    How to form healthy habits in your 20s - The N...
14    Turning Your Vacation Photos Into Works of Art...
15    As Second Avenue Subway Opens, a Train Delay E...
16    Dylann Roof Himself Rejects Best Defense Again...
19    Fecal Pollution Taints Water at Melbourne's Be...
20    N.F.L. Playoffs: Schedule, Matchups and Odds -...
21    Mariah Carey's Manager Blames Producers for Ne...
22    Damaged by War, Syria's Cultural Sites Rise An...
23    George Michael's Freedom Video: An Oral Histor...
24    With New Congress Poised to Convene, Obama's P...
25    Republicans Stonewalled Obama. Now the Ball Is...
26    Istanbul, Donald Trump, Benjamin Netanyahu: Yo...
27    Inside Trump Defense Secretary Pick's Efforts ...
28    ISIS Claims Responsibility for Istanbul Nightc...
29    The Afghan War and the Evolution of Obama - Th...
30    Tips for Your Post-Holiday Clutter Purge - The...
31    Benjamin Netanyahu Questioned in Israel Graft ...
32    France Lets Workers Turn Off, Tune Out and Liv...
33    Airline Pilot, Believed to Be Drunk, Is Pulled...
35    SpaceX Says It's Ready to Launch Rockets Again...
36    A 'World Unto Itself' in New York Area Yeshiva...
37    Inside Jonathan Lethem's Oddball Trove - The N...
38    An Immigrant's Instinct to Aid Others Collides...
39    House Republicans, Under Fire, Back Down on Gu...
40    Republican Party, Megyn Kelly, Titanic: Your W...
41    Turmoil Overshadows First Day of Republican-Co...
```

(19)

DATASET:

This a New York times 50,000 articles dataset collected from Kaggle.com

id	article_no	user rating	login	title	publication	author	date	year	month	url	content
0	17283	1	886	House Rep	New York	Carl Hulse	########	2016	12		WASHINGTON â€" Congressional Republicans have a new fear when it comes to their health care lawsuit against the Obama admin…
1	17284	2	777	Rift Betwe	New York	Benjamin	########	2017	6		After the bullet shells get counted, the blood dries and the votive candles burn out, people peer down from windows and see crime sc…
2	17285	0	915	Tyrus Won	New York	Margalit Fe	1/6/2017	2017	1		When Walt Disneyâ€™s â€œBambiâ€ opened in 1942, critics praised its spare, haunting visual style, vastly different from anything Disn…
3	17286	3	793	Among De	New York	William M	########	2017	4		Death may be the great equalizer, but it isnâ€™t necessarily evenhanded. Of all the fields of endeavor that suffered mortal losses in 20…
4	17287	0	335	Kim Jong-u	New York	Choe Sang	1/2/2017	2017	1		SEOUL, South Korea â€" North Koreaâ€™s leader, Kim said on Sunday that his country was making final preparations to conduct its…
5	17288	1	386	Sick With a	New York	Sewell Cha	1/2/2017	2017	1		LONDON â€" Queen Elizabeth II, who has been battling a cold for more than a week, missed a New Yearâ€™s Day church service at …
6	17289	2	492	Taiwanâ€™	New York	Javier C. H	1/2/2017	2017	1		BEIJING â€" President Tsai of Taiwan sharply criticized Chinaâ€™s leaders on Saturday, saying they had resorted to military and eco…
7	17290	4	649	After â€"T	New York	Gina Kolat	2/8/2017	2017	2		Danny Cahill stood, slightly dazed, in a blizzard of confetti as the audience screamed and his family ran on stage. He had won Season 8…
8	17291	1	421	First, a Mi	New York	Katherine	########	2016	12		Just how is Hillary Kerr, the founder of a digital media company in Los Angeles? She can tell you what song was playing five years a…
9	17292	2	362	Calling on	New York	Andy New	########	2016	12		Angels are everywhere in the MuÃ±iz familyâ€™s apartment in the Bronx: paintings of angels on the wall, ceramic angels flanking the a…
10	17293	2	27	Weak Fed	New York	Justin Gillis	1/3/2017	2017	1		With Donald J. Trump about to take control of the White House, it would seem a dark time for the renewable energy industry. After all…
11	17294	0	690	Can Carbo	New York	John Schw	1/5/2017	2017	1		THOMPSONS, Tex. â€" Can one of the most promising â€" and troubled â€" technologies for fighting global warming survive durin…
12	17295	4	59	Mar-a-Lag	New York	Maggie Ha	1/2/2017	2017	1		WEST PALM BEACH, Fla. â€" When Donald J. Trump rang in the new year this weekend, he did it in opulence, joined by the actor Sy…
13	17296	3	763	How to fo	New York	Charles Du	1/2/2017	2017	1		This article is part of a series aimed at helping you navigate lifeâ€™s opportunities and challenges. What else should we write about? C…
14	17297	1	926	Turning Yo	New York	Stephanie	########	2017	4		Itâ€™s the season for family travel and photos â€" and perhaps enlarging some of those images of snowy landscapes or tropical geta…
15	17298	0	540	As Second	New York	Emma G. F	1/2/2017	2017	1		Finally. The Second Avenue subway opened in New York City on Sunday, with thousands of riders flooding into its polished stations to …
16	17300	1	426	Dylann Ro	New York	Kevin Sack	1/2/2017	2017	1		pages into the journal found in Dylann S. Roofâ€™s car â€" after the assertions of black inferiority, the lamentations over white po…
17	17301	2	172	Modiâ€™s	New York	Geeta Ana	1/2/2017	2017	1		MUMBAI, India â€" It was a bold and risky gamble by Prime Minister Narendra Modi of India that quickly seemed to backfire. The ann…
18	17302	1	736	Suicide Bo	New York	The Associ	1/3/2017	2017	1		BAGHDAD â€" A suicide bomber detonated a pickup truck loaded with explosives on Monday in a busy Baghdad market, killing at leas…
19	17303	1	211	Fecal Pollu	New York	Brett Cole	1/3/2017	2017	1		SYDNEY, Australia â€" The annual beach pilgrimage during the height of summer in Melbourne, Australiaâ€™s city, is threatened by a…
20	17305	3	368	N.F.L. Play	New York	Benjamin	1/3/2017	2017	1		When the Green Bay Packers lost to the Washington Redskins in Week 11, dropping to Aaron Rodgers vowed to â€œrun the tableâ€ i…
21	17306	2	567	Mariah Ca	New York	Patrick He	1/2/2017	2017	1		Mariah Carey suffered through a performance train wreck in Times Square on New Yearâ€™s Eve as malfunctions left her at a loss voc…
22	17307	4	429	Damaged I	New York	Marlise Sir	1/1/2017	2017	1		PARIS â€" When the Islamic State was about to be driven out of the ancient city of Palmyra in March, Yves Ubelmann got a call from…
23	17308	2	782	George Mi	New York	Guy Treba	1/1/2017	2017	1		Pop music and fashion never met cuter than in George Michaelâ€™s â€œFreedom! â€™90â€ video. Sure, other catwalk favorites starr…
24	17309	0	530	With New	New York	Jennifer St	1/2/2017	2017	1		WASHINGTON â€" The most powerful and ambitious Congress in 20 years will convene Tuesday, with plans to leave its mark on virt…
25	17311	2	862	Republican	New York	Carl Hulse	1/3/2017	2017	1		WASHINGTON â€" Itâ€™s or time for Republicans. After a tumultuous decade that has seen profound changes in the makeup and …
26	17312	3	123	Istanbul, D	New York	Charles M	1/3/2017	2017	1		Good morning. Hereâ€™s what you need to know: â€¢ The Turkish authorities are hunting for the gunman who opened fire at an Istar…
27	17313	2	67	Inside Trur	New York	Sheri Fink	1/5/2017	2017	1		The body of the Iraqi prisoner was found naked and badly bruised in 2003, outside a detention center in southern Iraq run by United St…

(20)

DATABASE FOR NEW USERS:

USERS		CREATE TABLE USERS(USERNAME TEXT NOT NULL,EMAIL TEXT,PASSWORD TEXT)
USERNAME	TEXT	`USERNAME` TEXT NOT NULL
EMAIL	TEXT	`EMAIL` TEXT
PASSWORD	TEXT	`PASSWORD` TEXT

Database Structure | Browse Data | Edit Pragmas | Execute SQL

Table: USERS

	USERNAME	EMAIL	PASSWORD
	Filter	Filter	Filter
1	rick_morty	rick@gmail.com	abcde
2	sarthak	s@gmail.com	abcer
3	sarthak	s@gmail.com	abcer
4	sare	s@gmail	abcvr
5	abcd	ab@gmail.com	acvfr
6	abcd	ab@gmail.com	acvfr
7	abvft	aa@gmail.cocm	asdf
8	swq	aaaa@gmail	aqwe
9	sarthak	sa@gmail.com	abcde
10	rohit	rohit@gmail.c...	meanmachine

(21)

Code to implement the database:

```python
import sqlite3

def createTable():
    connection = sqlite3.connect("login.db")

    connection.execute("CREATE TABLE USERS(USERNAME TEXT NOT NULL,EMAIL TEXT,PASSWORD TEXT)")

    connection.execute("INSERT INTO USERS VALUES(?,?,?)",('rick_morty','rick@gmail.com','abcde'))

    connection.commit()
    result = connection.execute("SELECT * FROM USERS")

    for data in result:
        print("Username:",data[0])
        print("Email:",data[1])
        print("Password:",data[2])

    connection.close()

createTable()
```

RESULT:

These are the most Read Articles

0 House Republicans Fret About Winning Their Hea...
1 Rift Between Officers and Residents as Killing...
2 Among Deaths in 2016, a Heavy Toll in Pop Musi...
3 Kim Jong-un Says North Korea Is Preparing to T...
4 Sick With a Cold, Queen Elizabeth Misses New Y...
5 Taiwan's President Accuses China of Renewed In...
6 Calling on Angels While Enduring the Trials of...
7 Weak Federal Powers Could Limit Trump's Climat...
8 Can Carbon Capture Technology Prosper Under Tr...
9 Mar-a-Lago, the Future Winter White House and ...
10 How to form healthy habits in your 20s - The N...
11 Turning Your Vacation Photos Into Works of Art...
12 As Second Avenue Subway Opens, a Train Delay E...
13 Dylann Roof Himself Rejects Best Defense Again...
14 Fecal Pollution Taints Water at Melbourne's Be...
15 N.F.L. Playoffs: Schedule, Matchups and Odds -...
16 Mariah Carey's Manager Blames Producers for Ne...
17 Damaged by War, Syria's Cultural Sites Rise An...
18 George Michael's Freedom Video: An Oral Histor...
19 With New Congress Poised to Convene, Obama's P...
20 Republicans Stonewalled Obama. Now the Ball Is...
21 Istanbul, Donald Trump, Benjamin Netanyahu: Yo...
22 Inside Trump Defense Secretary Pick's Efforts ...
23 ISIS Claims Responsibility for Istanbul Nightc...
24 The Afghan War and the Evolution of Obama - Th...
25 Tips for Your Post-Holiday Clutter Purge - The...
26 Benjamin Netanyahu Questioned in Israel Graft ...
27 France Lets Workers Turn Off, Tune Out and Liv...
28 Airline Pilot, Believed to Be Drunk, Is Pulled...
29 SpaceX Says It's Ready to Launch Rockets Again...
30 A 'World Unto Itself' in New York Area Yeshiva...
31 Inside Jonathan Lethem's Oddball Trove - The N...
32 An Immigrant's Instinct to Aid Others Collides...
33 House Republicans, Under Fire, Back Down on Gu...
34 Republican Party, Megyn Kelly, Titanic: Your W...
35 Turmoil Overshadows First Day of Republican-Co...
36 With Choice of Trade Negotiator, Trump Prepare...
37 Trump Appears to Side With Assange Over Intell...

REFERENCES:

- kaggle.com (to get dataset)
- courser.org
- learning data mining with python by Robert Layton
- https://www.safaribooksonline.com/library/view/programming-collective-intelligence/9780596529321/ch02s02.html
- Programming Collective Intelligence by Toby Segaran
- https://www.youtube.com/watch?v=KeqVL-0vSQg&list=PLseNcwx1RJ4WdgtrMTXndw4B4nlf4-pgS
- https://www.youtube.com/watch?v=39vJRxIPSxw&t=3769s
- https://www.youtube.com/watch?v=4Y5zjTJ9LV4&t=2587s

DOCUMENT APPROVAL

By appending my signatures to this report I agree that the minor project(news recommendation system) is implemented properly

Name of mentor: DR. ANURAG SINGH BAGHEL

Signature:

Date:

Name of group member: DUSHYANT PRATAP SINGH

Signature:

Date:

Name of group member: NAMITA SINGH CHOUHAN

Signature:

Date:

Name of group member: PRATEEK KALRA

Signature:

Date:

Name of group member: SARTHAK TEOTIA

Signature:

Date: