

## NewsRec System using NLP for revolution 4.0 secured by blockchain

Anjali Pise  
Department of CSE  
Bdce,  
Wardha, India.

Dr.Amit Thakre  
Head of Department of CE  
Bdce,  
Wardha, India

*Abstract*—News is the strongest conversation device for this present day time. It can unfold the information all throughout the world in a couple of minutes. It is a satisfactory medium to get related with the whole world. We can research the whole thing from it. It is one of the exceptional sources to get understanding and analyze higher skills. We can get entry to the information on the net for free. All throughout the world and each vicinity have to get admission to to the news. The web is a very robust medium to broadcast news. Nowadays everybody examines information because of technology,

humans are studying them on their cell or computer, however, it is nevertheless equal news. News let us study so many things, specifically tradition,culture, modern-day world news.It supports awareness.

Awareness is an important thing for farmers' related agricultural updates. The farmers get the exact information that they require for their field.

NewsRec system help farmers to find the news of their choices based on their own news experience in an efficient and effective manner without wasting time in purposeless browsing.

**Keywords-AI, NLP, Recommendation system, Metamask, Ether**

### I. INTRODUCTION

1.1 .Machine learning is the subpart of artificial intelligence. IT enables a machine to automatically learn from data, it has improved performance from past predict the things without explicitly programmed.

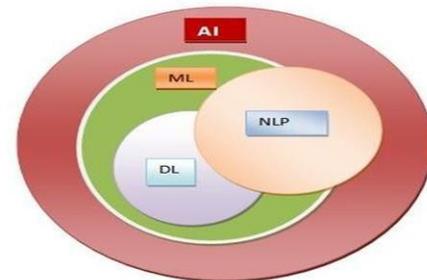


Fig - NLP ,ML are the part of AI

### 1.2 Recommendation System

Recommendations usually pace up searches and make it less complicated for customers to get admission to the content material they're fascinated in, and shock them with presents they would have in no way searched for. The consumer begins to experience acknowledged and understood and is greater probably to purchase extra merchandise or consume greater content. By understanding what Characteristic information. This is statistics about gadgets (keywords, categories, etc.) and customers (preferences, profiles, etc.). User-item interactions. This is data such as ratings, a number of purchases, likes, etc.

Types of Recommendation System. There are mainly 4 types of recommendation systems.

1. Content- based system
2. Collaborative system.
3. Knowledge-based system
4. Hybrid system.

We simply want to make sure that we assign them a crew in accordance with their features.

### 1.4 Natural language processing

The cause of natural language processing strives to

design machines that recognize and react with textual content or voice records and reply with textual content or speech on their very own in tons the equal manner human beings do. It is a semi-discipline of synthetic intelligence that makes NLP languages like dialects comprehensible for machines. It is capable of creating wise machine the usage of system mastering set of rules that could intelligently carry out repetitive analytical and predictive tasks .

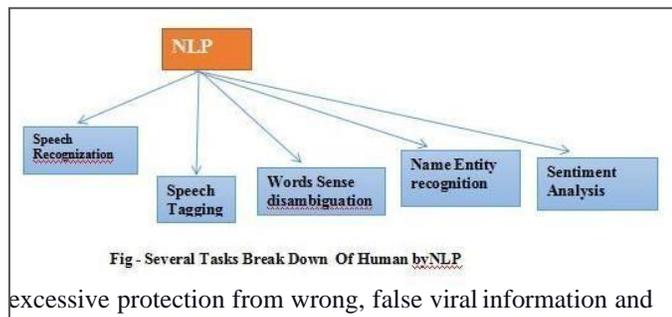


Fig - Several Tasks Break Down Of Human by NLP

excessive protection from wrong, false viral information and creates a social disturbance insidesociety.

## 2. Objectives

To design a news recommendation system for spreading the awareness related towards the smart farming and connect them to the digital globe.

It provide knowledge,multiple technique which are used in recently in the market through mobileapp.

## 3. Related Work

Every researcher done some research in recommendation system field. There are multiple recommendation systems are use. Some researcher was work on stereotyping recommendation system .

[1] Beel et.al. In this recommendation system is used for psychology which allowed the psychologists to quickly judge by people based on few characteristics called as “Facets”.

[2] Goldberg et.al proposed that “information filtering process”. He was used the collaborative recommendation system.

[3] Rensick.et.al was proposed that “Interaction” between users and their items was established through authorship. He collected personal collection, adding social tags.

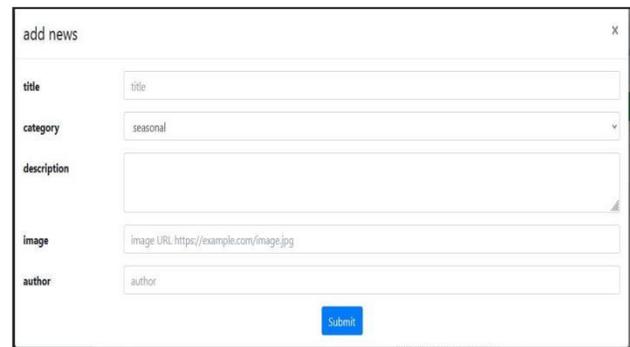
[4] Zarrinkalam and Kahani considered authors as featured and determine similarities by the number of authors two items share.

[5] “Published in”, “authorship”, “relatedness between genes it depending on the entities in the graph.

[6] TechLens is one of the most famous research paper recommender systems although it was not just first like some have claimed. currently TechLens recommender system active in other recommendation system field .

## 4. System Experiment & Implementation Work

[1] For this news recommendation system, Dataset of news samples of 120. The rating file has 5 categories.



[2] This recommendation system workflow likewise  
 1. Moduled dependent 2. Library Importer 3. Data frames creator 4. Similarity matrix creator 5. New recommendation.

```
[1]: def library_importer():
import pandas as pd
import nltk
import numpy as np
from sklearn.metrics.pairwise import cosine_similarity
from sklearn.feature_extraction.text import CountVectorizer
import warnings
from rake_nltk import Rake
r=Rake()
return pd,nltk,np,cosine_similarity,CountVectorizer,Rake,r

def dataframe_creator(path):
pd,nltk,np,cosine_similarity,CountVectorizer,Rake,r=library_importer()
df=pd.read_json(path)
RecommendedNews = df['available_news'].values.tolist()
df = pd.DataFrame(RecommendedNews,columns=['row_id','title','category','de

df['Key_words'] = ''
r = Rake()
for index, row in df.iterrows():
r.extract_keywords_from_text(row['description'])
key_words_dict_scores = r.get_word_degrees()
row['Key_words'] = list(key_words_dict_scores.keys())

df['category_list']=df['category'].apply(lambda x:[x])
```

**Fig .Code Segments**

Newsrec System recommend on the basis of bag of words using RAKE.

1	Govt pegs wheat procurement at record 444 lakh...	seasonal government estimated procurement reco...	The government has estimated procurement of a ...	seasonal
2	Dairy industry's revenue to rebound 12 this f...	seasonal revenue india organised dairy industri...	Revenue of India's organised dairy industry wi...	seasonal
3	PM Modi suggests seven pointers from budget pr...	government policy emphasising agriculture budg...	Emphasising that the agriculture budget alloca...	government policy
4	Budget 2022 focusses on making agriculture sma...	government policy agricultural loans farmers l...	Agricultural loans for farmers have been incre...	government policy
5	Wholesale price of Mustard oil drops 10	seasonal prices mustard oil come since harvest...	Prices of mustard oil have come down since har...	seasonal
6	Tax on Agriculture GST applicable on seeds us...	government policy law applicable grain seed di...	The law applicable to grain and seed will be d...	government policy
7	Government releases buffer stock of onion to s...	government policy buffer stock released even l...	The buffer stock is being released even in Las...	government policy
8	Pulses importers land in trouble after govt s...	seasonal making representation union ministry...	We are making representation to the Union mini...	seasonal
9	Who will pay for the technology drones flown	technology drones flown	Drones have flown into technology	technology

**5. Storing & securing news data using public blockchain.**

An ethereum blockchain platform is used in this process. The metamask wallet is used for transaction. User is used the metamask wallet. The ether faucet is used for purchasing the ether for the testing purpose. It prevents from malicious actors. Each time the value of ether will change based on network traffic. The Unit of ether is Wei. Today price of ether is \$3,122.06 is frequently changed.

**6. System Architecture of NewsRec system**

In this system architecture

1. Store the raw news in blockchain and then transfer to the central server.
2. Central server store the raw data and send back to the admin.
3. If admin get success response from the

server then it will call the ml model read the update data from central server.

AdminLTE 3	1	Tripura launches Rs 153 crore pineapple & jackfruit mission for next 5 years	seasonal	AGARTALA: With the encouraging response from the farmers and demand in the international market over the past three years, the Tripura government on Wednesday announced to set up Pineapple and Jackfruit Mission for the next five years to develop a hort-based business model for a sustainable rural economy.		TOI
	2	Avocado Market Size in 2022: 3.3% CAGR	seasonal	Mar 14, 2022 (The ExpressWire) — Increase In Demand (2022) At 3.3% CAGR, The Avocado Market Size is expected to grow from 11980 million USD in 2020, to reach 15000 Million USD by 2026. The base year considered for the study is 2021, and the forecast period is between 2022 and 2026.		Marketwatch news, March 14, 2022
	3	Avocado shortage spotlights need for strong bilateral supply chain	seasonal	The recent ban on Mexican avocado importation into the U.S. has highlighted the importance of a strong bilateral supply chain that promotes economic growth and prosperity for both countries, said Texas A&M AgriLife experts. The recent ban on Mexican avocado imports showed the importance of a strong bilateral Mexico-U.S. supply chain. (Texas A&M AgriLife photo) Inspections of		AgriLife

[3] **Cosine similarity** - It is the content based recommendation logic. In this logic will compute the following 3 matrix

Similarity(S), Correlation(C), Freshness(F). If users like to read articles of similar category or topics then set the higher value for "Similarity metric". It computes the values range={0 to 1}. Computed page by this mathematical formula -  $S(\text{Article 1}, \text{Article 2}) = \frac{|\text{Article 1} \cap \text{Article 2}|}{|\text{Article 1} \cup \text{Article 2}|}$  and Bag of words algorithm is used for finding the keywords..

```
words += ' '.join(row[col]) + ' '
row['Bag_of_words'] = words

# strip white spaces in front and behind, replace multiple whitespaces (if any)
df['Bag_of_words'] = df['Bag_of_words'].str.strip().str.replace(' ', '')
df_new = df[['title', 'Bag_of_words', 'description', 'category']]
return df_new, df

def similarity_matrix_creator():
count = CountVectorizer()
count_matrix = count.fit_transform(df['Bag_of_words'])
# to generate the cosine similarity matrix (size 250 x 250)
# rows represent all movies; columns represent all movies
# cosine similarity: similarity = cos(angle) = range from 0 (different) to 1
# all the numbers on the diagonal are 1 because every movie is identical to itself
# matrix is also symmetrical because the similarity between A and B is the same as B and A
# for other values eg 0.1578947, movie x and movie y has similarity value of 0.1578947

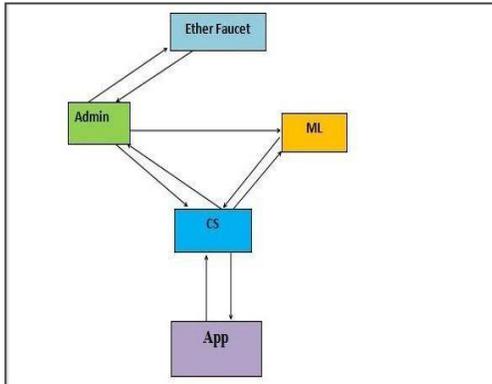
cosine_sim = cosine_similarity(count_matrix, count_matrix)
#print(cosine_sim)
# to create a Series for news titles which can be used as indices (each index is a title)
indices = pd.Series(df['title'])

return cosine_sim, indices

def News_recommender(title, cosine_sim):
recommended_description = []
```

4. ML send the the update data to the central server for storing the updated news.
5. App request the for updated news which is add by user.
6. Central server respond to the application

with the updated news.



## 7. Conclusion & Future Scope

Recently recommendation system was used since is a very common system. Multiple companies are used in multiple domains. In the agriculture domain, a recommendation system is used in a very less manner, especially in the news domain. Fake news are in trend, it create disturbance in the society. When NLP is used in this newsrec system it gives the very proper knowledge about the globe and provides security from the blockchain. From the multiple malicious proliferation. For being as smart, developing, and digital India it is an essential part to always aware about upcoming techniques and connects with the globe. The proper news has the power to shape the world in new direction.

## 8. REFERENCES

- [1] Yun Xu (April 2015) IEEE , “Recent Machine learning Applications to Internet of Things”. [2]Nallapaneni Manoj Kumar ,Pradeep Kumar Mallick, ( 2018) Elsevier“Blockchain technology for security issues and Challenges”, International Conference On Computational Intelligence andData Science.
- [3] Renick,Paul and Hal R.Varian,”Recommender Systems”Communications of the ACM40,no.3(1997)
- [4] G. Sambasivam, Geoffrey Duncan Opiyo ( 2020) IEEE “A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks”,Egyptian Informatics Journal.
- [5] Fran Casino, Thomas K. Dasaklis,Constantinos Patsakis , Elsevier (2019), “A Systematic literature review of blockchain-based applications: Currentstatus”, Telematics and Informatics.
- [6] Jitin Jagannath,Nicholas Polosky,Anu Jagannath,Francesco Restuccia and Tommaso Melodia , (July2021) IEEE “Machine Learning for wireless Communications in the Internet of Things: A Comprehensive Survey”.Ad Hoc Networks.
- [7] Shaik Vaseem Akram,Praveen Kuma Malik, Rajesh Singh Anita Sudeep Tanwar, IEEE, (June 2019 ) “Adoption of blockchain Technology in various Realms: Opportunities and Challenges”.
- [8] Omar Said,Masud( 2013) , “Towards Internet of things: Survey and Future Vision”,International Journal of computer networks(IJCN),Volume 5.issue(1):2013.
- [9] Karlgren,Jussi,1990.”An Algebra fo Recommendations Syslab Working paper.Retrieved 27 october 2017.
- [10] L.Terveen,W.Hill,D.Mcdonald et.al:A system forsharingrecommendations,Communication ,ACM 40930(1997)
- [11]D.H.Wng, Y.C.Liang,DXu,X.Y.Feng.(2018)),A content-based recommender system for computer Sciencepublications,”Knowledge-based systems, 157:1-9 [12].<https://www.google.com/search?q=electronics+hardware+parts&client=firefoxbd&sxsrf=AOaem>

vLTLskwbWcE43PzQQknT0FXCyXmA:164083  
8765971&source=Inms&tbn=shop&sa=X&ved=2a  
hUKEwjJyZ3e2Ir1AhWvF6YKHbfjCLYQ\_AUoA  
noECAIQBA&biw=1366&bih=643&dpr=1. [13]Harikumar  
Pallathadka, Malik Mustafa,  
Domenic T. Sanchez, Guna Sekhar Sajja, Sanjeev Gour,  
Mohd Naved(2020) Elsevier, "Impact Of machine learning  
on Management, healthcare and agriculture.  
[14] Sudeep Tanwar,Quasim Wai- Chnag Hong(Senior  
member of IEEE) e.tl "machine learning adoption in  
blockchain based smart application:: challenges and the way  
forward " IEEE (2020)  
[15] Prem Melvill and Vikas Sindhvani, Recommender  
systems,Encyclopedia of machine Learning,2010.

[16] Donghui Wang, Yanchun Liang, Dong  
Xu et al., "A content-based recommender system for computer  
science publication" (2018) Elsevier.  
[17] Beel, J., Gipp, et al., "Research paper Recommender  
Systems: A Literature  
Survey", International Journal on Digital  
libraries. (2015).  
[18] R.J. Mooney & L. Roy (1999). Content-based book  
recommendation using learning for text Categorization. in  
workshop recom.Sys.Algo and evaluation.  
[19] "Playboy Lead Rise of Recommendation engines -" TIME  
TIME.com "27 May 2010  
[20] J. Lu, D. Wu, M. Mao, et al., "Recommender system  
application developments: a survey, Decis, Support  
Syst" (2015).