**Journal of Mechanical and Construction Engineering, 2024,**
Vol. 04, Iss. 02, S. No. 063, pp. 1-8
ISSN (Online): 2583-0619

# Travel Recommendation System Using Content and Collaborative Filtering

**Shaurya Goel[1], S.W.A. Rizvi[2]**

[1,2]Department of Computer Science & Engineering, Amity School of Engineering & Technology Lucknow, Amity University Uttar Pradesh, India
[1]avigoyal544@gmail.com, [2]swarizvi@lko.amity.edu

**How to cite this paper:** S. Goel, S.W.A. Rizvi, "Travel Recommendation System Using Content and Collaborative Filtering," *Journal of Mechanical and Construction Engineering (JMCE), Vol. 04, Iss. 02, S. No. 063,* pp. 1–7, 2024.

https://doi.org/10.54060/a2zjournals.jmce.63

**Received:** 01/05/2024
**Accepted:** 06/08/2024
**Online First:** 16/08/2024
**Published:** 25/11/2024



## Abstract

*Tourism significantly impacts a nation's economy, yet there remains a void in platforms offering tailored information on local attractions. In our study, we propose a hybrid recommendation system amalgamating content and collaborative filtering methods to provide personalized tourist suggestions. This approach mitigates individual methods' drawbacks, enhancing recommendations' accuracy. To gauge item similarity, we employ cosine similarity while integrating SVD within a model-based collaborative filtering framework for improved outcomes. By utilizing a weighted hybridization technique, we effectively merge the outputs of both approaches. We collected tourist attraction and user data for implementation, yielding superior results compared to standalone content-based and collaborative filtering methods.*

## Keywords

*Collaborative Filtering, Content Filtering, Information Filters*

## 1. Introduction

The burgeoning tourism industry underscores the need for efficient and personalized travel recommendation systems. However, existing platforms often fall short of delivering tailored suggestions to users. In response, this research paper introduces a novel approach that combines content and collaborative filtering techniques to address this gap effectively [1-3]. By leveraging both methodologies, our system aims to overcome the limitations inherent in individual recommendation methods, thereby providing users with more accurate and personalized recommendations [3]. Using cosine similarity for measuring item similarity and the incorporation of Singular Value Decomposition (SVD) within a model-based collaborative filtering framework, we strive to enhance the efficacy of our recommendation system [5]. Additionally, we employ a weighted hybridization approach to seamlessly integrate the outputs of both content and collaborative filtering methods. This paper presents the methodology, implementation, and evaluation of our hybrid recommendation system, highlighting its superiority over traditional content-based and collaborative filtering approaches.

Recommender systems have found applications across various domains, including social media platforms like Facebook, entertainment platforms like Netflix and YouTube, and e-commerce giants such as Amazon and Flipkart. Within the travel domain, platforms like TripAdvisor and Trivago also utilize recommender systems. The paper is structured as follows: Section I introduces recommender systems, Section II provides an overview of travel recommender systems, and Section III discusses related work in recommender systems specifically within the travel and tourism domain [1].

## 2. Literature Review

Content-based filtering methods rely on the attributes of items and user preferences to make recommendations [1]. These methods have been widely utilized in various domains, including travel and tourism, due to their simplicity and effectiveness in recommending items like those previously liked by the user. However, they often suffer from the "cold start" problem, where recommendations for new users or items are challenging to generate due to limited historical data [3]. In contrast, collaborative filtering techniques analyze user-item interactions and similarities between users or items to make recommendations. Model-based collaborative filtering, such as Singular Value Decomposition (SVD), has shown promising results in capturing complex patterns in user preferences and item characteristics. Despite its effectiveness, collaborative filtering approaches may face challenges in scalability and sparsity of data, particularly in the travel domain where user preferences can be diverse and dynamic. [5] Recent research has focused on hybrid approaches combining content and collaborative filtering methods to leverage their strengths and mitigate their limitations. By integrating these approaches, hybrid recommender systems can provide more robust and accurate recommendations tailored to individual user preferences. [14] Weighted hybridization techniques, such as combining recommendation scores from content-based and collaborative filtering models, have been explored to optimize recommendation quality.

The product of the recommendation system would be based on a Web UI and a pickled Algorithm running in the background using the pre-processed data resources as is Ostensible in the provided schematic diagram.[16] The recommender system implemented for this product is a Utility-based recommender system i.e. it uses the user's provided input and matches it with the components of the existing dataset to provide you with the perfect utility that you desire, instead of providing recommendations based on your past visits or what type of places other customers belong in the same cluster as you chose after visiting the previous place.[9] The latter would require a customer database to perform customer segmentation and implement pattern recognition algorithms to provide recommendations. This is known as the Content-based recommender system.[2]

Sometimes you might want to integrate the two types to produce a recommender system that considers Past choices, Friends' choices, and utility based. This is known as Hybrid recommender system. [13]

We can either use the Word2Vec transformer trained by Google, or we can use Hugging. Face framework to use their provided NLP APIs.[11] Of course, if we decide to train our model and have enough available data to do so, this is usually the case. With functional websites, we can use the genism module in Python to train our word to vector transformer.[7]

The user on the client side of the server will provide four parameters, location description, user's location, the location the user desires to go to, and if distance is a relevant factor.[12] By matching the description with the reviews and the positive or negative sentimentality we will provide recommendations to the user based on their filtered preferences.[1] The recommender function at the end will provide a data frame filtered with the most optimal recommendations along with a unique ID which will then be rendered on the web UI as cards.[9]

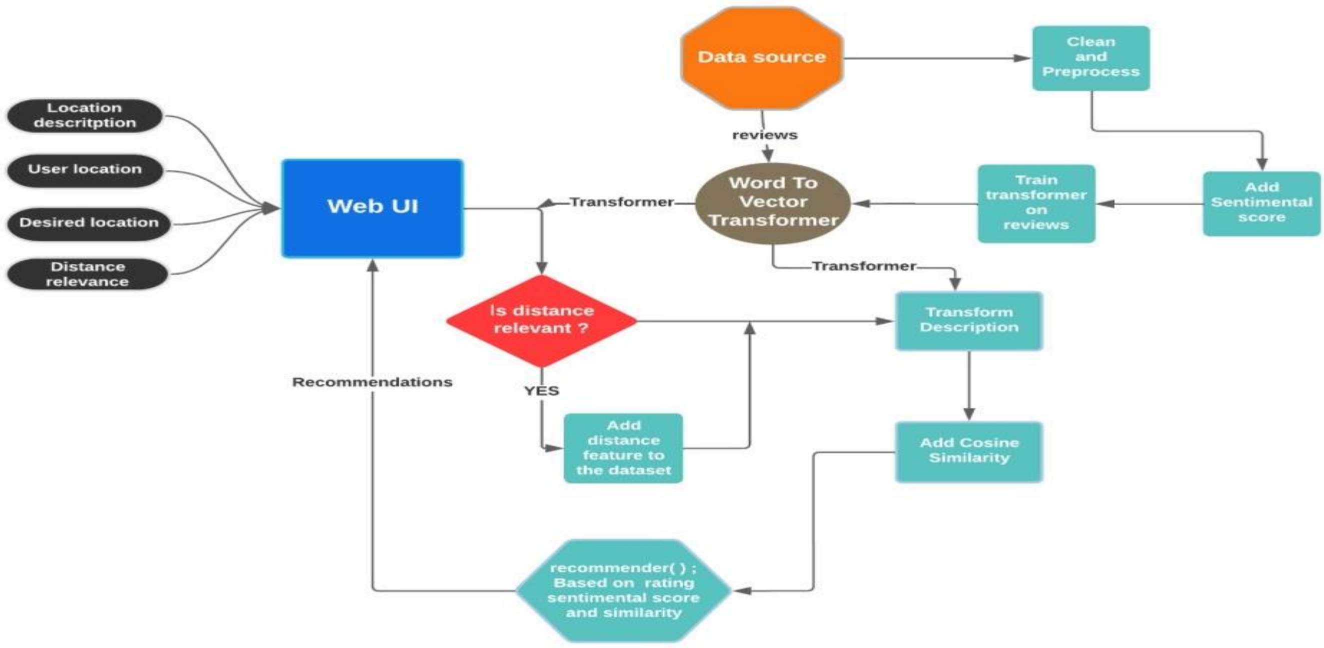


**Figure 1.** Prototype Working in graphical format

## 3. Methodologies and Algorithms

### 3.1. Word2Vec

Word2vec is a combination of models used to represent distributed representations of words in a corpus C. Word2Vec (W2V) is an algorithm that accepts text corpus as an input and outputs a vector representation for each word.[3] Here corpus is defined as the collection of all the sentences that are included in the vocabulary of the dataset. [7] There are two versions of this model used to produce neural word embeddings (i.e., the vectorized format of the provided word) based on the diverse types of neural. Architectures and how they learn the vectors and capture the semantic meanings.[9]

### 3.2. Skip-Gram transformer

The skip-gram neural network architecture loops over the corpus of the dataset and uses individual words as input to predict the provided context. Every time, the model can't predict the context words from a given input vector w it adjusts the weights of itself and tries again until it gets it right.[11] This helps in transforming every single word in the corpus by conserving the semantic meaning of the word.[1]

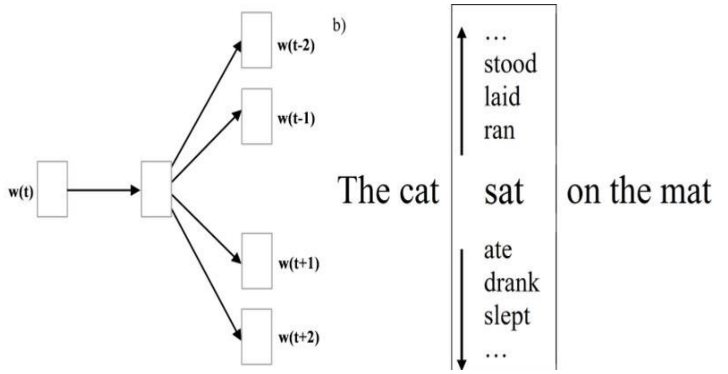


**Figure 2.** Skip-Gram Transformer

## 3.3. CBOW (continuous bag of words)

CBOW is a technique which could be thought of as an inverted version of skip-gram. This network uses the provided context of the word to predict the word in question. The architecture of this one is as follows. If, however, we do not have enough data to predict the vectors we can also use Google's pre-trained Word2Vec model or BERT which is another popular model used for NLPmodels.[8] However, it is to be noted that skip-gram works better on large datasets.[18]

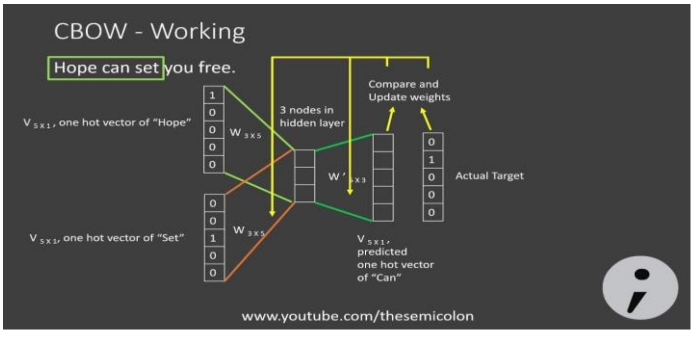


**Figure 3.** CBOW-Working

## 3.4. Cosine Similarity

After converting the words to vectors we will be required to compute the similarity between the description vector provided by the user and the best description available in the corpus of places. We can use cosine similarity to find out the most similar descriptions from the collections of reviews so we can find the document that will best match our user's request. [1]

We have all studied fundamental trigonometry to know that the angle between two vectors can be calculated using the following formula.[13]

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum\limits_{i=1}^{n} A_i B_i}{\sqrt{\sum\limits_{i=1}^{n} A_i^2} \sqrt{\sum\limits_{i=1}^{n} B_i^2}}$$

**Figure 4.** Cosine Similarity

The smaller the angle between two vectors the more parallel they are to each other In terms of word vectors, the smaller the angle is between two vectors the more similar the words are to each other.[2]

## 4. Results and Discussion

Naturally, the first step would be to train a Neural Network that will convert words to vectors to better encompass the semantic meaning of the individual words. But then the question arises Why would we do that?[5] Other models like the TfIdf vectorizer and Count vectorizer also are transformers that convert words into a numerical vector based on how frequently they are repeated in the sentence.

So why would we use a neural network when we can avoid overcomplications? [5]

There are two problems with traditional approaches that use Bag of Words.

a. **Data Sparsity:** The vectors used to represent the sentences in Bag of Words are sparse matrices with no limitation as to how many columns will there be. [3]

b. **Lack of Flexibility:** English is a language in which one word has several adaptions.

This could confuse the model while making recommendations. [7]

The four parameters passed by the user will be the user's location, the description of the desired place around which vicinity should the desired place be, and whether the distance is going to be an issue.[8] The distance relevance factor will come into play when the desired city is unknown. If the city, you desire is variable then we will compute distances from each city, and if the distance is relevant then we will consider making suggestions where the distance is a minimum. We will define a special type of city option called "None" which would allow variable optimal suggestions with minimum distance from the city you live in. [5]

Once the inputs are passed from the client side to the server, the description will be converted into a vector which is trained on our pre-existing dataset and then we will use utility-based filtering to obtain the optimal city locations for the user.[11]

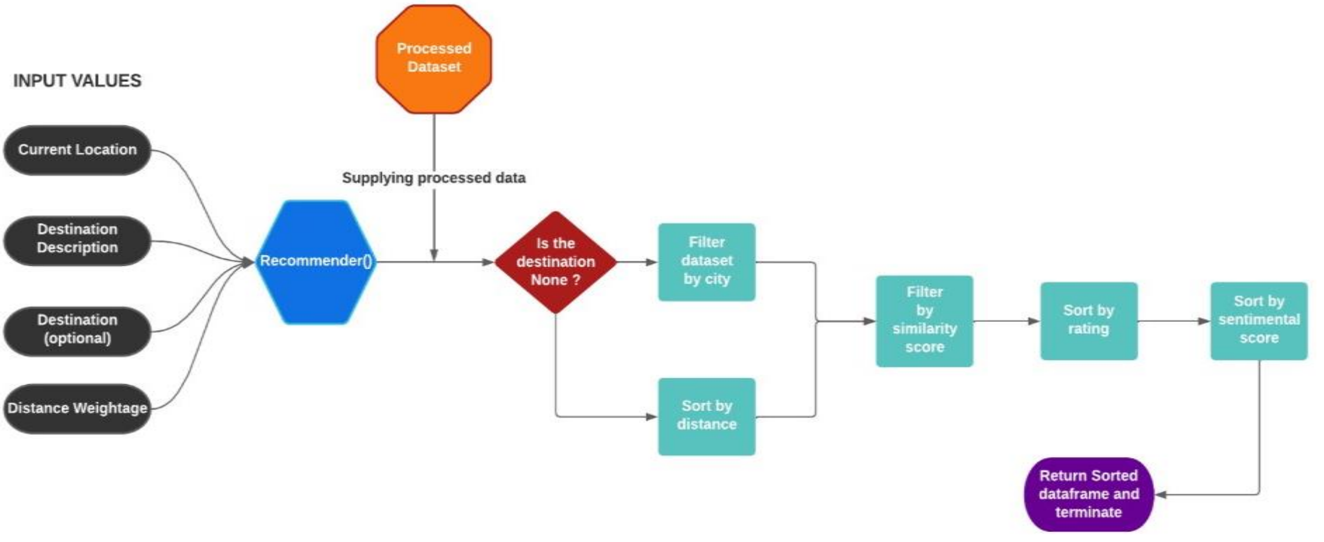


**Figure 5.** Graphical Interpretation of how Inputs And output

The recommender function would take in processed user inputs and data frames and return a set of recommendations that will be most appropriate to the user based on their preferences.[5]

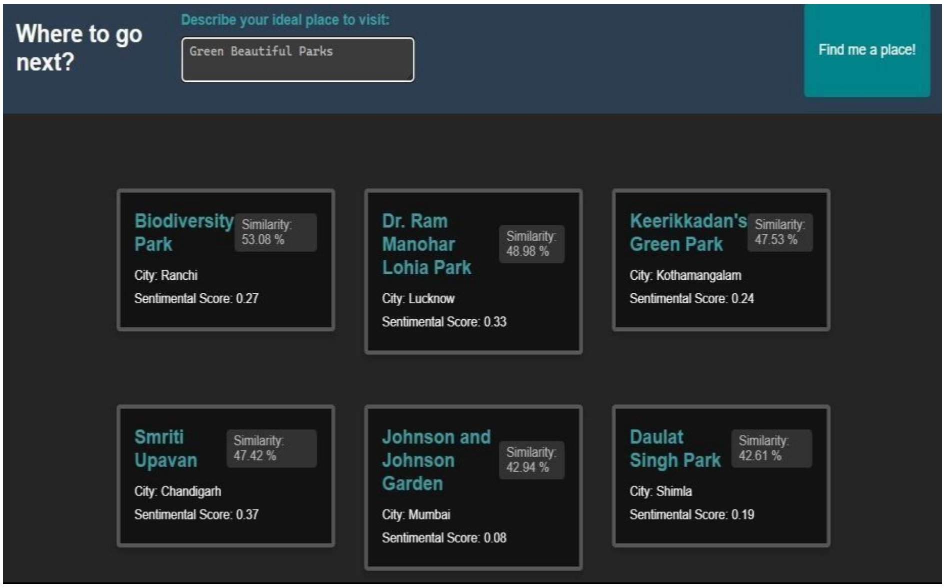


**Figure 6.** Working Prototype Screenshot 1

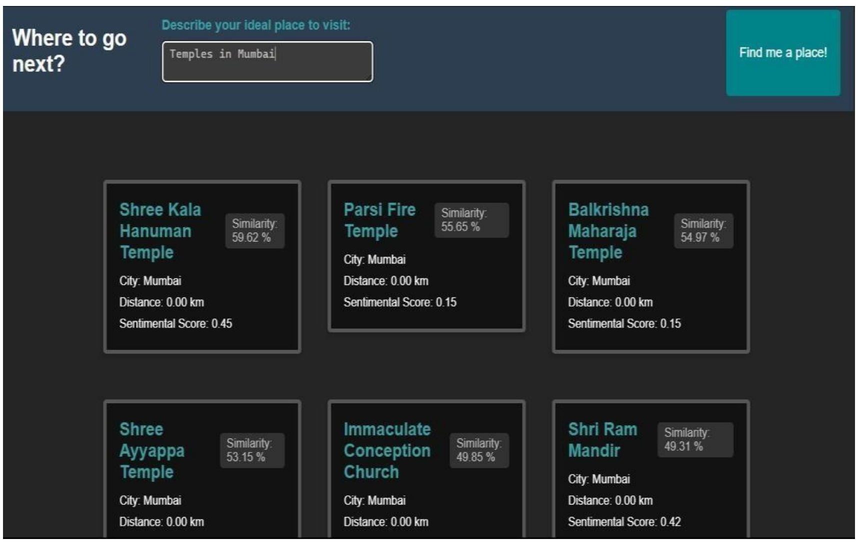


**Figure 7.** Working Prototype Screenshot 2

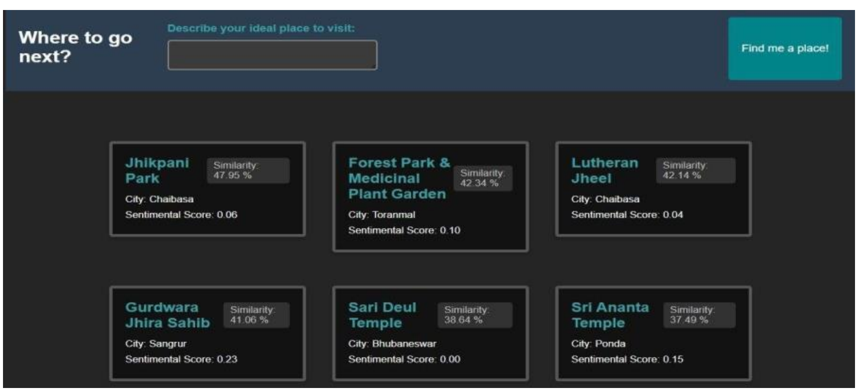


**Figure 8.** Working Prototype Screenshot 3

## 5. Conclusion

Despite the title of the paper the idea obviously will have a few costs that have to be undertaken to establish a decently functioning business. The point I was trying to make using the title is that with only minimal costs, we can produce a business that would be comparatively more effective than say opening a physical travel agency and providing suggestions simply based on your contacts and biases which could potentially lead the user to have a terrible experience.[3] However, the two worlds do not always need to collide. The algorithm is not necessarily here to take away the jobs of already existing physical travel agencies.[11] The best way of implementing this algorithm would be by collaborating with agencies using an organizational cloud.[5] We can use the money of several small agencies to invest a tiny sum of money and adopt a cloud computing service to store user data in different chunks and collaborate and produce one unified effective website that will use the contacts of travel agencies to arrange a fully packaged trip to a particular place and then integrating it with the website, hopefully even automating the entire procedure someday if the website becomes large enough.[6]

## 6. Future Scope

The future of travel recommendation systems using content and collaborative filtering holds promise for several advancements. These include refining personalization through advanced machine learning, incorporating context-aware recommendations, implementing dynamic adaptation mechanisms, exploring multi-modal recommendation techniques, and extending recommendations across domains for a broader user experience.

## References

[1.] V. Garipelly; P. T. Adusumalli; P. Singh, "12th International Conference on Computing Communication and Networking Technologies (ICCCNT)", *IEEE*,06-08 July 2021.

[2.] A. Rula, A. S. Hamid, and J. K. Albahri, "How smart is e-tourism? A systematic review of smart tourism recommendation system applying data management," vol. 39, no. 2021, pp. 100337, 2021.

[3.] P. Yochum, L. Chang, T. Gu, and M. Zhu, "Linked open data in location-based recommendation system on tourism domain: A survey," *IEEE Access*, vol. 8, pp. 16409–16439, 2020.

[4.] Z. A. Moud and H. V. Nejad, "Tourism recommendation system based on semantic clustering and sentiment analysis", vol. 167, no. 2021, pp. 114324. 2021.

[5.] C.-S. Jeong, K.-H. Ryu, and J.-Y. Lee, "Deep Learning-based Tourism Recommendation System using Social Network Analysis," *International Journal of Internet Broadcasting and Communication*, vol. 12, no. 2, pp. 113–119, 2020.

[6.] M. Wang, "Applying Internet information technology combined with deep learning to tourism collaborative recommendation system," *PLoS One*, vol. 15, no. 12, p. e0240656, 2020.

[7.] Shini and A. Sreekumar, "An extensive study on the evolution of context-aware personalized travel recommender systems", *Information Processing Management*, vol. 57, no. 1, pp. 102078, 2020.

[8.] C. T. Gu, Y. Sun and L. Chang, "A Travel Route Recommendation System Based on Smart Phones and IoT Environment," *Wireless Communications and Mobile Computing*, vol. 2019, 2019.

[9.] Z. Pu, H. Du, S. Yu, and D. Feng, "Improved tourism recommendation system," in *Proceedings of the 2020 12th International Conference on Machine Learning and Computing*, pp. 121-126, 2020.

[10.] M. Dhaware and P. Vanwari, "A tourism and travel recommendation system based on user-location vector," in *Lecture Notes in Electrical Engineering*, Singapore: Springer Singapore, 2020, pp. 1429–1437.

[11.] J. Akilandeswari, G. Jothi, K. Dhanasekaran, K. Kousalya, and V. Sathiyamoorthi, "Hybrid fire-fly-ontology-based clustering algorithm for analyzing tweets to extract causal factors," *Int. J. Semant. Web Inf. Syst. (IJSWIS)*, vol. 18, no. 1, pp. 1–27, 2022.

[12.] P. Ni Wayan, "Literature Review Recommendation System Using Hybrid Method by Utilizing Social Media as Marketing," *Computer Engineering and Applications Journal*, vol. 10, no. 2, pp. 105–113, 2021.

[13.] E. Omrani, B. Khoshnevisan, S. Shamshirband, H. Saboohi, N. B. Anuar, and M. H. N. Nasir, “Potential of ra-dial basis function based support vector regression for apple unwellness detection,” *J. Measuring*, pp. 233–252, 2014.

[14.] W. Yin, Y. Sun, and J. Zhao, “Personalized tourism route recommendation system based on dynamic clustering of user groups,” in *2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*, 2021.

[15.] K. A. Fararni, F. Nafis, B. Aghoutane, A. Yahyaouy, J. Riffi, and A. Sabri, “Hybrid recommender system for tourism based on big data and AI: A conceptual framework,” *Big Data Min. Anal.*, vol. 4, no. 1, pp. 47–55, 2021.