JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Developing a Collaborative Filtering Recommendation System for E-commerce Using k-means Clustering

P. Silpa Chaitanya[1*], A. Eesha Pravalli[2],K. Leela Rani[3],G.Nandini[4] , B.Venkateswari[5]
1* Associate professor: Vignan's Nirula Institute of Technology and Science for Women
2.3.4.5 B.Tech Scholar: Vignan's Nirula Institute of Technology and Science for Women.

Abstract

Nowadays, e-commerce websites are one of the main rising trends that facilitate online selection, and selling of products suggesting an algorithm that aims to forecast the preference (or rating) that a user would something. Collaboration filtering is one such effective way to offer suggestion for the users. We present two types of collaborative filtering techniques in this paper: memory-based and model-based. Algorithms based on memory: item-based collaborative filtering (IBCF) and user-based collaborative filtering (UBCF) filtration (IBCF). RandomForest Classification and K-Means are examples of model-based algorithms. Despite several drawbacks, the K-means clustering algorithm is a popular partitional clustering method for recommender systems due to its low complexity and ease of use.

Keywords: Recommendation System, Collaborative Filtering, K-means clustering, Item- based recommendation, User-based recommendation.

1. Introduction

Collaborative filtering takes into account the user's past purchases, selections, and/or numerical ratings of those things, in addition to other users' comparable choices. The user's potential interest in certain items or item ratings is predicted by this model. The only data used by the CF system to provide recommendations is the rating profiles of various persons or things. Conversely, however, Items must be described by features according to CBF, which uses an item's pre-tagged attributes to make recommendations more objects with comparable characteristics. Collaborative filtering, as opposed to content-based filtering, can be used with any kind of substance, but it can also convey ideas like quality that are difficult to express.

Furthermore, it has been established that the most effective and extensively used recommendation technique is collaborative filtering. For these reasons, in this study, we shall concentrate on the collaborative filtering technique. Model-based and memory-based collaborative filtering techniques are the two main categories into which they fall. The Memory-based Approach makes the assumption that the user and the item they have previously rated in the database must be in the system memory at the time the recommendation algorithm is executing. By first exploring the user and item database and then combining neighboring interests as recommendations, the similarities between various users (or objects) are computed. In the model based method, user ratings of unrated goods are predicted using CF models that are created using machine learning algorithms.

User-based collaborative filtering(UBCF) and item-based collaborative filtering(IBCF) are two components of the memory-based collaborative filtering technique. Simple similarity measurements like Pearson Correlation and Cosine Similarity are largely relied upon by memory-based techniques to group similar users or items together. The memory-based approach has a few shortcomings, including scalability issues and cold start. Model-Based Collaborative Filtering: The Model-Based approach involves estimating a user's level of interest in a product that they have never seen before. More machine learning algorithms can be trained to anticipate a user's preference (rating) for a newly added item to the system training to the vector of items for that particular user. It fixes the issue that the memory-based approach. When it comes to query time, model-based methods are usually faster than memory-based ones. Therefore, they are more suited for recommendations for coldstarts. Model-based collaborative filtering includes Random Forest and K-Means clustering techniques. This paper introduces an effective machine learning recommendation system that uses k-means clustering.

2. Literature Survey

In an effort to address the issues of scalability and sparsity in collaborative filtering, Song Jie Gong, the author, proposed "A Collaborative Filtering Recommendation Algorithm Based on User Clustering and Item." To do this, the system employs a personalized recommendation approach that combines user and item clustering technologies before using item cluster collaborative filtering to generate the recommendation.

In this study by Liao, Q. et al., the MapReduce concept is implemented through parallelism technique to improve K-means clustering. Reducing the number of repetitions and increasing the processing speed per iteration improves the K-means performance. Two approaches that are deemed reliable in light of the data scattering are suggested for the distance measure and the first selection of centroids. The experimental findings demonstrated that the suggested method outperforms the conventional K means in terms of stability.

Clustering algorithms are a useful tool for overcoming duplication and ambiguity, negotiating data sparsity and scalability, and fostering customization and personalization (Demissie and Shashi 2018; Chen et al. 2018b; Koohi and Kiani 2017). The goal of this thesis is to use a hybrid technique to enhance the clustering-based CF recommender system.

On big datasets,it has been demonstrated that the K-means split-merge clustering technique by Wang, Huang, and Bosneag (2018) is more effective than K-means clustering with the same flexibility and convenience (Boeva, Angelova, and Tsiporkova 2019; Wang, Huang, and Bosneag 2018). The final outcome of the procedure is formed by merging the closest clusters and dividing poor clusters after the data is first randomly partitioned into K clusters. According to Boeva, Angelova, and Tsiporkova (2019), the algorithm can be applied when fresh data is introduced to the database, aiding in the development of a strong recommender system for incoming real-time data applications.

Nevertheless, Belacel et al. (2018) failed to take use of the chance to apply item clustering, creating avoid in the clustering implementation and failing to capitalize on the benefits of item- based collaborative filtering over user-based collaborative filtration, as demonstrated by Azzopardi's research (2017). One of the best methods for recommender systems is collaborative filtering (CF)(Chen etal. 2018b; JiandShen2015; MoradiandAhmadian2015). A common technique in collaborative filtering recommender systems is to utilize a similarity metric, such as cosine similarity or Pearson correlation coefficient (PCC), to locate neighbors.

Heryanto, (2018) The recommender system emerges as a crucial technological solution to address this issue, assisting consumers in locating goods that align with their interests. In actuality, the recommender system makes predictions about user preferences by utilizing past.

Choosing the appropriate technique for developing a recommender system might be challenging. Various recommender system strategies have been developed, including collaborative, content-based, and hybrid filtering (Prayoga & Kusnawi, 2022).

According to Khusna, Dela Sano, and Saputra (2021), Collaborative filtering is a popular and effective technique for constructing recommendation systems. This technique recommends products based on user interests and commonalities.

To increase the efficacy and calibre of suggestions, YH Cho, JKKim, and SH Kim suggest a personalized recommendation technique. Numerous data-mining techniques, such as decision trees, k-means clustering, association rules, online usage, and product taxonomy, are used in this paper's methodology. It makes product recommendations based on browsing patterns and purchasing behavior. By applying decision tree approaches, it prevents providing inappropriate recommendations. Marketers change product suggestions based on explicit knowledge. This method is explained in detail for accurate RS.

A system for intelligent, knowledge-based product recommendations was proposed by Steven Ashaya. The products are being recommended using artificial intelligence found in processing components and subjective product information that is either supplied by customers or entered into the systems during initial setup. Using this method, customer requirements are discovered using keyword searches rather than just by looking through purchase records.

(Azzopardi 2017; Li, Lu, and Xue Feng 2005; Koohi and Kiani 2017). Therefore, it would be interesting to investigate whether applying the K-means split-merge clustering technique to the item-based collaborative filtering recommender system can reduce the performance gap compared to K-means clustering-based recommender systems and address scalability issues.

3. Proposed Methodology

Recommendation approach based on collaborative filtering Recommender systems based on k-means clustering.

K-Means Clustering is an unsupervised learning algorithm that is used to solve the clustering problems in machine learning. This divides the dataset without labels into many clusters. In this case, K indicates how many pre-defined clusters must be formed during the process; for example, if K=2, there will be two clusters; if K=3, there will be three clusters, and so on. The unlabeled dataset is divided into k distinct clusters using an iterative process, with each dataset belonging to a single group with comparable attributes. It enables us to arrange the data into distinct groups and provides a practical means of independently identifying the group categories in the unlabeled dataset without the requirement for training. Each cluster in this technique is linked to a centroid, which is based on centroid theory. This algorithm's primary goal is to reduce the total distance between each data point and its associated cluster. The unlabeled dataset is fed into the algorithm, which then separates it into k number of clusters and keeps going until it finds no optimal clusters. In this procedure, the value of k ought to be predetermined.

Algorithm

Step-1: Select the number K to decide the number of clusters.

Step-2: Select random K points or centroids. (It can be other from the input dataset).

Step-3: Assign each data point to their closest centroid, which will form the predefined K clusters.

Step-4: Calculate the variance and place a new centroid of each cluster.

Step-5: Repeat the third steps, which means reassign each datapoint to the new closest centroid of each cluster.

Step-6: If any re assignment occurs, then go to step-4 else go to FINISH.

Step-7: The model is ready.

The K-means clustering algorithm's performance is dependent on the formation of highly efficient clusters. However, determining the ideal number of clusters is a difficult process. There are other approaches to determining the ideal number of clusters, but in this case, we'll focus on the most suitable technique for determining the number of clusters, or K value. The following is the method:

Elbow Method

One of the most widely used techniques for determining the ideal number of clusters is the elbow method. The WCSS value is a notion used in this procedure. Within Cluster Sum of Squares, or WCSS, is a statistical measure that characterizes the total variation inside a cluster. The following formula can be used to get the WCSS value (for three clusters):

$WCSS=\sum_{PiinCluster1} distance(P_iC_1)^2+\sum_{PiinCluster2} distance(P_iC_2)^2+\sum_{PiinCLuster3} distance(P_iC_3)^2$In the above formula of WCSS, ∑Pi in Cluster1 distance (Pi C1)2: For each data point within a cluster1, it is the sum of the squares of the distances to its centroid. The same holds true for the other two terms. We can use any method, such as the Manhattan distance or the Euclidean distance, to calculate the distance between the data points and the centroid. The elbow approach uses the following procedures to get the ideal value of clusters:

It conducts the K-means clustering on a supplied dataset for different K values (ranges from 1- 10). The WCSS value is computed for each value of K. Creates a curve connecting the number of clusters (K)with the computed WCSS values. The point on the plot that resembles an arm or a sharp bend is regarded as the best value for K.

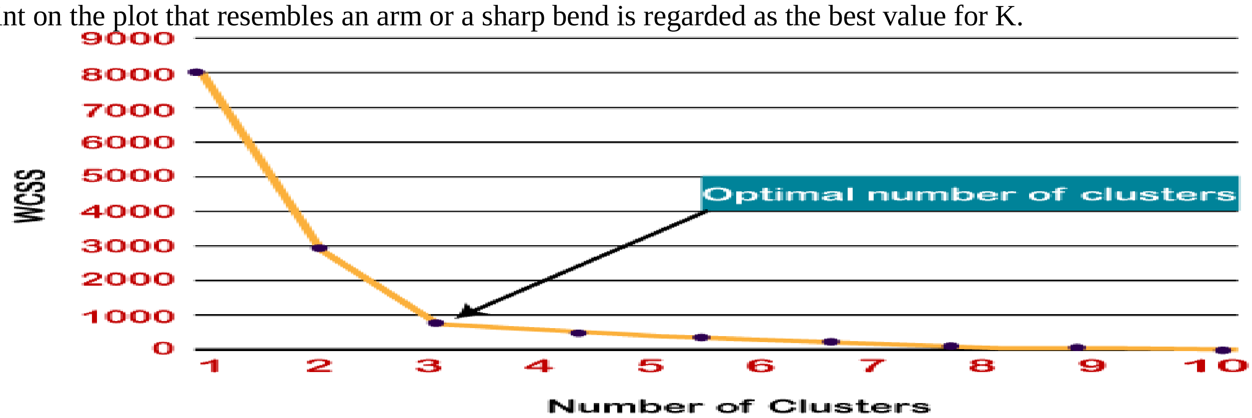

Fig:1elbow method for predicting k value

Elbow technique is used to find K (numbers of clusters). The Elbow Method is a method we employ to figure out how many centroids (k) to include in a k-means clustering algorithm. Using this strategy, we repeatedly iterate from k=1 to k=n (where n is the hyper parameter we select based on our needs)to get the k-value. We determine the within-cluster sum of squares (WCSS)value for each value of k.

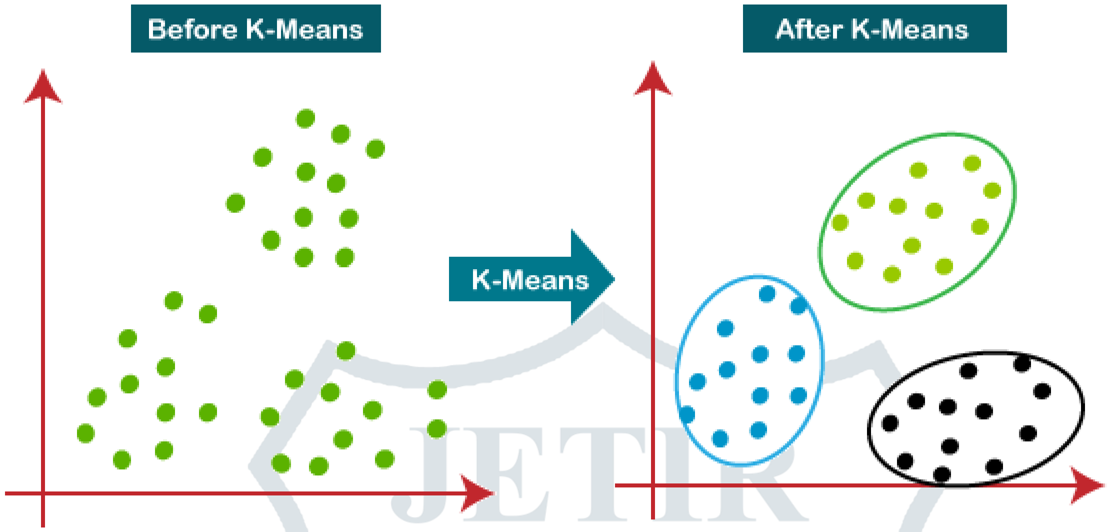

Fig2: k-means clustering in collaborative filtering for e-commerce

This process involves selecting k points from training dataset, using elbow method after predicting the k value we implement k-means clustering. The unlabeled dataset is divided in to k distinct clusters using an iterative process, with each dataset belonging to a single group with comparable attributes.

4. Result

The Elbow Method is used in the K-Means clustering to determine the optimal number of clustering. The elbow approach is a well-liked technique for figuring out the ideal value of K to run the K-Means Clustering Algorithm. This method's main concept is to visualize the different cost values as k changes. There will be fewer items in the cluster as K rises. Consequently, average distortion will drop. Less items indicate a closer proximity to the centroid. Therefore, the elbow position is where this distortion decreases the most.

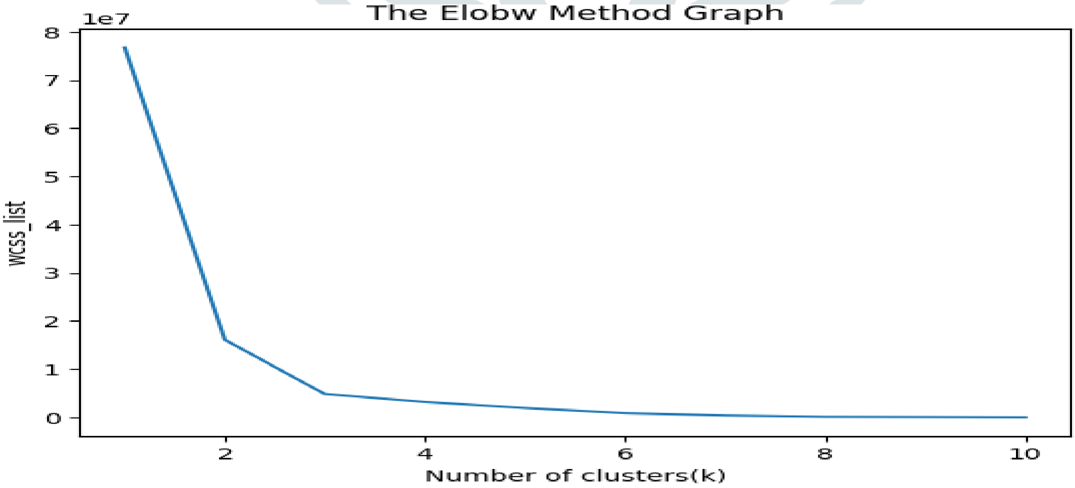

Fig3: Elbow method for predicting k value

In order to establish an empty list that will hold the value of wcss computed for various values of k ranging from 1to10, we constructed the wcss_list variable. The for loop was the n initiated for the iteration on a new value of k between 1 and 10; as Python's for loop excludes the out bound limit, 11 is assumed to contain the10th value. From the above plot. We can see the elbow point at 3. Therefore, the number of clusters here will be 3.

This is intended to show where clusters that make sense intuitively can be produced by k- means.

Fig4: Clusters Of Customers

The output image makes the three distinct clusters with varying colors quite visible. The dataset's two parameters- spending price and rating form the clusters. Depending on the situation or preference, we can alter the labels and colors. Additionally, we can see the following points from the a fore mentioned patterns:

Cluster1 shows the customers with average spending price and average rating so we categorize these products have low quality

Cluster2 shows the customers with average spending price and above average rating so we categorize these products are useful and having high quality.

Cluster3 shows the customers with high spending price and high rating so we categorize these products are useful, long-lasting, and having best quality.

5. Conclusion

In this research, we have covered a simple k-means clustering technique with an acceptable example. This study applies the k-means clustering method to the E-commerce dataset analysis. The clustering algorithm was analyzed qualitatively in this paper, and the distance between items was calculated using Euclidean distance. Using 4590 transaction records from an e-commerce dataset, we conducted the k-means clustering technique. Based on the theoretical analysis and experimental results in this study, the k-means method performs more efficiently when grouping large volumes of data. The system in the proposal uses various factors together with area information to provide users with more tailored and up-to-date recommendations.

6. References

[1] Vejendla, L.N., Bysani, B., Mundru, A., Setty, M., & Kunta, V. J. (2023). Score-based support vector machine for spam mail detection. Proceedings of the IEEE International Conference on Electronics, Computing and Communication Technologies (ICOEI), 915–920.

[2] Raval, U. R.; Jani, C. (2015): Implementing and Improvisation of K-Means Clustering. International Journal of Computer Science and Mobile Computing, vol. 4, pp. 72-76.

[3] Utkarsh, Gupta, and Nagamma Patil. Chameleon is a hierarchical clustering algorithm-based recommender system. Pages 1006–1010, IEEE International Advance Computing Conference (IACC), 2015.

[4] Madhulatha and TagaramSoni.2011 sawa comparison of the clustering methods k-means and k-medoids. Pages 472–481, Springer International Conference on Advances in Computing and Information Technology.

[5] John Riedl, Joseph Konstan, Ben Schafer, and "Recommender systems in e-commerce." The First ACM Conference on Electronic Commerce Proceedings. 1999ACM.

[6] Chidambaram, M. and Balasubramaniam, KA. 2018. A new method for recommender system clustering supervised learning (csl). Science & Technology, International Journal of Simulation-Systems, 19(4).

[7] Yao G. and Cai L., "Designing Collaborative Filtering Recommendation Algorithms: A User-Based and Item-Based Approach," University of California, San Diego, 2017.

[8] Joel Azzopardi. 2017.Collaborative recommendation predictions: item-based versus user- based. Pages 165–170 of Semantic Keyword-based Search on Structured Data Sources, Springer.

[9] Khurana P., Parveen S., “A Successful Hybrid Recommender Method Using Enhanced K- means and Similarity,” International Journal of Computer Trends and Technology (IJCTT), 36(3), 2016.

[10] Belacel, Nabil, Cajetan Bouchard, Guillaume Durand, and Serge Leger. 2018. scalable splitting-merging clustering-based collaborative filtering system. Pages 290-311, Springer, International Conference on Agents and Artificial Intelligence.

[11] Anusha, P., Ravikiran, A., Lakshman Narayana, V., & Maddumala, V. R. (2020). Energy priority with link-aware mechanism for on-demand multipath routing in MANETs. Journal of Electrical Engineering and Automation, 29(3), 8979–8991.

[12] Narayana, V. L., & Bharathi, C. R. (2017). Identity-based cryptography for mobile ad hoc networks. International Journal of Engineering and Technology Innovation, 95(5), 1173–1181

[13] Narayana, V. L., & Midhunchakkaravarthy, D. (2020). A time interval-based block chain model for detection of malicious nodes in MANET using network block monitoring node. Proceedings of the IEEE International Conference on Intelligent Computing and Robotics (ICIRCA), 852–857. https://doi.org/10.1109/ICIRCA48905.2020.9183256

[14] Bharathi Vejendla, C. R., Narayana, L., & Ramesh, L. V. (2020). Secure data communication using Internet of Things. Journal of Advanced Research in Dynamic and Control Systems, 9(4), 3516–3520.

[15] Narayana, V. L., Sujatha, V., Sri, K. S., Pavani, V., Prasanna, T. V. N., & Ranganarayana, K. (2023). Computer tomography image-based interconnected antecedence clustering model using deep convolutional neural network for prediction of COVID-19. Traitement du Signal, 40(4), 1689–1696. https://doi.org/10.18280/ts.400437

[16] Patibandla, R. S. M. L., Rao, B. T., & Narayana, V. L. (2022). Prediction of COVID-19 using machine learning techniques. In Handbook of Machine Learning for Computational Biology and Bio-informatics (pp. 219–231). https://doi.org/10.1016/B978-0-12-824145-5.00007-1

[17] Pavani, V., Sri, K. S., Krishna, P. S., & Narayana, V. L. (2021). Multi-level authentication scheme for improving privacy and security of data in decentralized cloud server. Proceedings of the IEEE International Conference on Systems and Electronics Engineering (ICOSEC), 391–394. https://doi.org/10.1109/ICOSEC51865.2021.9591698

[18] Arepalli, P. G., Jairam Naik, K., & Rout, J. K. (2024). Aquaculture water quality classification with sparse attention transformers: Leveraging water and environmental parameters. In ACM International Conference Proceeding Series (pp. 318-325). https://doi.org/10.1145/3651781.3651829

[19] Kumar, S. A., Babu, E. S., Nagaraju, C., &Gopi, A. P. (2015). An empirical critique of on-demand routing protocols against rushing attack in MANET. International Journal of Electrical and Computer Engineering, 5(5), 1102-1110. https://doi.org/10.11591/ijece.v5i5.pp1102-1110

[20] Arepalli, G., Erukula, S. B., Gopi, A. P., & Nagaraju, C. (2016). Secure multicast routing protocol in MANETs using efficient ECGDH algorithm. International Journal of Electrical and Computer Engineering, 6(4), 1857-1865. https://doi.org/10.11591/ijece.v6i4.9941

[21] Arepalli, P. G., Akula, M., Kalli, R. S., Kolli, A., Popuri, V. P., & Chalichama, S. (2022). Water quality prediction for salmon fish using Gated Recurrent Unit (GRU) model. In 2022 2nd International Conference on Computer Science, Engineering and Applications, ICCSEA 2022. https://doi.org/10.1109/ICCSEA54677.2022.9936539

[22] Narayana, V. L., & Gopi, A. P. (2017). Visual cryptography for gray scale images with enhanced security mechanisms. Traitement du Signal, 34, 197-208. DOI: 10.3166/ts.34.197-208

[23] Arepalli, P. G., Narayana, V. L., Venkatesh, R., & Kumar, N. A.(2019). Certified node frequency in social network using parallel diffusion methods. Ingenierie des Systemesd'Information, 24(1), 113-117. https://doi.org/10.18280/isi.240117

[24] Peda Gopi, A., & Lakshman Narayana, V. (2017). Protected strength approach for image steganography. Traitement du Signal, 34(3-4), 175-181. https://doi.org/10.3166/TS.34.175-181

[25] Narayana, V. L., Gopi, A. P., Anveshini, D., & Lakshmi, G. V. V. (2020). Enhanced path finding process and reduction of packet droppings in mobile ad-hoc networks. International Journal of Wireless and Mobile Computing,18(4),391-397. https://doi.org/10.1504/IJWMC.2020.108539

[26] Vejendla, L. N., Naresh, A., &Arepalli, P. G. (2021). Traffic analysis using IoT for improving secured communication. In Innovations in the Industrial Internet of Things (IIoT) and Smart Factory (pp. 106-116). https://doi.org/10.4018/978-1-7998-3375-8.ch008

[27] Kanumalli, S. S., Lavanya, K., Rajeswari, A., Samyuktha, P., & Tejaswi, M. (2023, February). A scalable network intrusion detection system using bi-lstm and cnn. In 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS) (pp. 1-6). IEEE.

[28] Kanumalli, S. S., Mantena, S. J., Kandula, S., Doppalapudi, K., & Atluri, T. (2022, May). Automated Irrigation Management System using IoT. In 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 476-482). IEEE.

[29] Kanumalli, S. S., Swathi, S., Sukanya, K., Yamini, V., & Nagalakshmi, N. (2022). Classification of dna sequence using machine learning. In Soft Computing for Security Applications

[30] Prathipati, Silpa Chaitanya, and Susanta Kumar Satpathy. "A Multilevel De-Noising Approach for Precision Edge-Based Fragmentation in MRI Brain Tumor Segmentation." *Traitement du Signal* 40.4 (2023): 1715.

[31] Kosaraju, Chaitanya, et al. "Mirchi crop yield prediction based on soil and environmental characteristics using modified RNN." 2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS). IEEE, 2023

[32] Prathipati, Silpa Chaitanya, and Susanta Kumar Satpathy. "Transforming 3D Brain Tumour Image Segmentation: An Enhanced V-Net Approach for Precise Diagnosis and Treatment Planning." 2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI). IEEE, 2024.

[33] Sujatha, V., Prasanthi, Y., Pravallika, C.H., ... Ayesha Banu, S.K., Sahithi, M.,K(23)

[34] "A Computer Vision Method for Detecting the Lanes and Finding the Direction of Traveling the Vehicle Lecture Notes in Networks and Systems", 2023, 612, pp. 373–382

[35] Sri, L. Akhila, K. Manvitha, G. Amulya, I. Sai Sanjuna, and V. Pavani. "FBI crime analysis and prediction using machine learning." Journal of Engineering Sciences 11, no. 4 (2020): 441-448.

[36] P. V, L. S. K, P. Vyshnavi A, M. Ch and S. B. G, "Students Community Portal using Machine Learning," 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2023, pp. 1109-1113, doi: 10.1109/ICEARS56392.2023.10085516.

[37] Sirisha, Aswadati, et al. "Intrusion detection models using supervised and unsupervised algorithms-a comparative estimation." International Journal of Safety and Security Engineering 11.1 (2021): 51-58.

[38] Krishna, KVSS Rama, et al. "Vehicle Number Plate Detection using Deep Learning." 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS). IEEE, 2024.

[39] Krishna, Komanduri Venkata Sesha Sai Rama, et al. "Classification of Glaucoma Optical Coherence Tomography (OCT) Images Based on Blood Vessel Identification Using CNN and Firefly Optimization." Treatment du Signal 38.1 (2021).

[40] Rayachoti, Eswaraiah, Sudhir Tirumalasetty, and Silpa Chaitanya Prathipati. "Watermarking system for telemedicine based on FABEMD." Multimedia Tools and Applications 81.30 (2022): 44383-44404.

[41] Chaitanya, P. Silpa, KV Narasimha Reddy, and G. Madhavi. "Effective Search of Color-Spatial Image Using Semantic Indexing." International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol 2 (2012): 9-19.

[42] Alapati, N., Prasad, B. V. V. S., Sharma, A., Kumari, G. R. P., Veeneetha, S. V., Srivalli, N., ... & Sahitya, D. (2022, November). Prediction of Flight-fare using machine learning. In 2022 International Conference on Fourth Industrial Revolution Based Technology and Practices (ICFIRTP) (pp. 134-138). IEEE.

[43] Alapati, N., Prasad, B. V. V. S., Sharma, A., Kumari, G. R. P., Bhargavi, P. J., Alekhya, A., ... & Nandini, K. (2022, November). Cardiovascular Disease Prediction using machine learning. In 2022 International Conference on Fourth Industrial Revolution Based Technology and Practices (ICFIRTP) (pp. 60-66). IEEE.

[44] Srikanth Kilaru " A Novel Approach to Human Iris Recognition And Verification Framework Using Machine Learning Algorithm" 2023 6th International Conference on Contemporary Computing and Informatics (IC3I),

[45] DOI: 10.1109/IC3I59117.2023.10397886, ISBN Information: Electronic ISBN:979-8-3503-0448-0 Print on Demand(PoD) ISBN:979-8-3503-0449-7

[46] Srikanth Kilaru "Analytical models for collaborative autonomous mobile robot solutions in fulfillmentcenters" in Applied Mathematical Modelling,Volume 91, March 2021, Pages 438-457,https://doi.org/10.1016/j.apm.2020.09.059

[47] Gopi, A. P., & Naik, K. J. (2022). An IoT model for fish breeding analysis with water quality data of pond using modified multilayer perceptron model. 2022 International Conference on Data Analytics for Business and Industry (ICDABI), 448-453. https://doi.org/10.1109/ICDABI56818.2022.10041617

[48] Arepalli, P. G., & Naik, K. J. (2024). A deep learning-enabled IoT framework for early hypoxia detection in aqua water using lightweight spatially shared attention-LSTM network. Journal of Supercomputing, 80*(2)*, 2718-2747. https://doi.org/10.1007/s11227-023-05580-x

[49] Arepalli, P. G., & Naik, K. J. (2023). An IoT-based water contamination analysis for aquaculture using lightweight multi-headed GRU model. Environmental Monitoring and Assessment, 195*(12)*, Article 1516. https://doi.org/10.1007/s10661-023-12126-4

[50] Gopi, A. P., Gowthami, M., Srujana, T., Gnana Padmini, S., & Durga Malleswari, M. (2023). Classification of denial-of-service attacks in IoT networks using AlexNet. In Smart Innovation, Systems and Technologies (Vol. 316, pp. 349-357). https://doi.org/10.1007/978-981-19-5403-0_30

[51] Bikku, T., Gopi, A. P., & Prasanna, R. L. (2019). Swarming the high-dimensional datasets using ensemble classification algorithm. In Advances in Intelligent Systems and Computing (Vol. 815, pp. 583-591). https://doi.org/10.1007/978-981-13-1580-0_30

[52] Sharma Shivani. Shivani Sharma developed a recommendation system based on an improvised K-means clustering method, which was published in July 2018 in the International Journal of Research in Advanced Technology (Vol. 6, No. 7, E-ISSN: 2321-9637).

[53] Boeva, Milena Angelova, Veselka, and Elena Tsiporkova. 2019. A split-merge evolutionary clustering technique. The ICAART 2019 proceedings, pages 337–346.