2022 International Conference on Futuristic Technologies (INCOFT)
Karnataka, India. Nov 25-27, 2022

Drug Recommendation System Using LDA

Satwik Reddy Sripathi
Department of Computer Science and Engineering,
Amrita School of Computing, Bengaluru,
Amrita Vishwa Vidyapeetham, India
satwikreddysripathi@gmail.com

Nadella Venkata Sai Pradyumna
Department of Computer Science and Engineering,
Amrita School of Computing, Bengaluru,
Amrita Vishwa Vidyapeetham, India
pradyumnanadella10@gmail.com

Akula Dhanush
Department of Computer Science and Engineering,
Amrita School of Computing, Bengaluru,
Amrita Vishwa Vidyapeetham, India
dhanushakula18@gmail.com

Subramani R
Department of Mathematics,
Amrita School of Engineering,
Bengaluru, India
r_subramani@blr.amrita.edu

***Abstract* – These days, people are very much concerned about their health. Many want to lead a healthy and long life. However, studies reveal that errors due to doctor's and pharmacist's experience resulted in many people's death. These medical errors are costing us our lives. Technical algorithms like LDA, PCA and other machine learning algorithms can help in topic modelling that would play a major role in the industry. These technologies can aid in the investigation of medical history and, by being doctor-friendly, can help to prevent medical errors. In this paper, a medicine is proposed based on the review taken from the patient as it would be identified with our model created with LDA. Recommendation system that analyses sentiment analysis of patient review data to discover the best medication for a disease is done in our work.**

Key words: LDA, PCA, Recommendation system, Topic Modelling.

I. INTRODUCTION

For a human, one of the most important things would be their health and to maintain it without any issues there are many searches done on the internet. These searches are completely about health information, and this information tells them what to do for their condition. There are many situations where many people died because of the misdiagnosis done by the self-proclaimed doctors, who put their medical clinics illegally. Because of this issue many people do not get proper diagnosis. Over the previous few decades, a considerable clinical data reflecting health of patient (e.g., physician reports, disease treatment Plans, test results etc.) has been gathered.

The digital data available for making patient centered decision has increased drastically as a result of this. Even though there is so much data available, such digital content is usually distributed across several websites which makes it difficult for patient or the user to find information that will help them improve their health. Furthermore, medical personnel have to access to more medications, tests, and treatment recommendations frequently. It becomes difficult for them to select treatments for patients appropriately [17].

The motivation behind this work is that implementing medical recommender systems should help bridge these gaps and help patients and doctors make better healthcare decisions. Our work provides physicians with a drug recommendation system that they can utilize when prescribing pharmaceuticals. A recommender framework is a standard framework that provides users with such a list of items that they can use to meet their own needs.

In contrast to a variety of other frameworks, health advice is mostly based on the patients emotional, bodily, and mental concerns. A pharmaceutical recommendation system is a comparable system that uses patient evaluations to propose drugs for a specific ailment. In today's quickly expanding technological environment, this technology is vital since it can save lives by supporting doctors. The suggested pharmaceutical recommendation system and its operation are described in this study, and it is deployed using Artificial Intelligence. Here we are using the LDA (Latent Dirichlet allocation) for creating out Model.

The model generated is done using topic modelling after borrowing the concepts from natural language processing. With our model, a patient can directly give their feedback or review regarding their condition to successfully get a recommended drug for that condition. This recommended drug is given after we successfully build our model and after we take the review from a patient. This review taken from the patient is further sent back into the model to get the most probable drug that can be considered for the patient. In our work we will be describing about the already existing works in section 2 I.e. the literature survey. After that we would be discussing about our model architecture and the results along with our conclusion and future scope.

II. LITERATURE SURVEY

With the rise in artificial intelligence, applications using deep learning and machine learning are playing an important role in all of the sectors present. And one of the most interesting and dominating application is the recommendation system. For this application there are various types of implementations including the concepts of reinforcement learning and deep reinforcement learning, ideas of machine learning, different models of deep learning and also using the natural language processing using topic modelling[24]. One of the most booming sectors for these applications is the medical field. But unfortunately, in the present world there aren't many recommendation systems available for the medical field like for the drug recommendation part as referred the number of pages from [5].

In order to filter required information according to preferences and discriminate it, Recommended Systems were

978-1-6654-5046-1/22/$31.00 ©2022 IEEE

created to help people. A method to automatically label documents based on LDA is proposed by J.C. Bailón-Elvira [4].

AI was applied to a national health database by Hung [3], in order to develop dental care recommendations. Data from the 2013–2014 National Health and Nutrition Examination Survey was obtained, and prediction models were created using LASSO and R to establish the best regression models.

Recently, the concept of self-aid has received a lot of popularity. In this sense, there has been an upsurge in research towards individualized wellness support systems [12]. Calorie-in-take plays a major role in one's well-being. It is very important to balance the calorie content. In this regard, the author proposed a hybrid recommendation framework that delivers physical activity and diet recommendation to meet a user's wellness needs.

Another way of recommendation is by matching data in database. An Android application that recommends possible medication, diet, and exercise to assist patients control their diabetes has been designed and built suing collaborative filtering [13]. Input data from user is matched with data in database and different things are recommended based on it.

Xiaoyao Zheng [10], proposes a tourism destination recommendation system that uses user emotion and a temporal dynamic to reflect user choice and destination popularity.

In this work, model, multiple regression analysis and social network analysis have been implemented for a book recommendation network [1]. The attributes linked to increase the sale have been taken into consideration and follows Pareto's law distribution.

Lee [11], proposed a text-mining approach for recommendation services which is robot-based. Flexible responses are generated when a robot comes across linguistic expressions. Two methods TF-IDF and LDA were combined in this work so that the robot responds with a recommended content and not what is pre-programmed.

A user can find information most relevant to his/her requirement with help of a recommendation system. A multi-period product recommendation system is proposed in this work which implements RNNs for analyzing time series data over collaborative filtering by which the model can learn customers' purchasing pattern [2].

DonghuiWang [9], developed a model to assist authors in determining where to submit their work. They use softmax regression along with chi-square feature selection to recommend a conference along with a context-based journal in computer science in their system.

This paper focuses on arranging the medication recommender framework as well as collecting data from medicinal case expertise in the medical dataset, which covers a number of novel recommender ideas. Through on-line social networking, communication is obscenely improved, and completely distinct information is presented on the internet essentially at the open pace[21]. The fully various information must communicate in order to record peaks of possible edges and availability of benefit bits of knowledge and people practices.

Leilei Sun [14], In order to find the most probable and optimal therapy prescription for patients, researchers looked at a vast number of treatment data. The plan was to evaluate how comparable treatment data were using a fast semantic clustering algorithm. Similarly, the author created a framework for evaluating the treatment's efficacy. Based on their demographics and medical conditions, this framework can offer the best treatment strategies for new patients. At present there is a publication present to recommend the drugs based on the topic modelling [20]. In this paper, the authors worked on generating the most probable drugs for a respective disease using topic modelling. They made their model by dividing it into 5 different components.

1. The dataset module which consists of reviewing the dataset.
2. Data creating and preparing module, where the data preprocessing is done.
3. Creating the recommendation module, with their logic for recommendation
4. Evaluating the model
5. Visualization of that model

The model which they used for their recommendation model is the pure topic modelling using n gram model. The n-gram is a group of words that occur together. This approach can be used to build features for supervised machine learning models like decision trees and naive bayes.[20]

These are the steps which they observed for the n-gram model

Step 1: Starting with unigram part and finding the collection of co-occurring terms in the reviews.

Step 2: Computing for the 1-gram model, here they used a single corpus to analyses the text. However, it does not adequately classify the required topics

Step 3: Computing for the 2-gram model, utilizing bi-gram to categories positive and negative reviews is very difficult. So we proceed for our next step

Step 4: Computing for the 3-gram model, The Tri-grams have yet to be able to distinguish between positive and negative feedback so we proceed to get the 4-gram models

Step 5: Computing for the 4-gram model, when compared to other gram models, the 4-gram classifies emotions significantly better. As a result, the deep learning model is built using 4-gram.

Step 6: Now they calculated the sentiments for each review, with sentiment=1 when the rating is greater than 5, and sentiment=0 when the rating is less than 5.

Step 7: Using the given sentiments and the updated review, they created a deep learning model.

They even used light gbm for their recommendation system, and on how exactly hey worked with this gbm is explained in the steps below.

Step 1: They started with the uni-gram and found the collection of co-occurring terms in the reviews.

Step 2: Then they proceeded for the 1-gram model. Here they used a single corpus to analyses the text. However, it does not adequately classify the required emotion.

Step 3: Then they proceeded for the 2-gram model: utilizing bi-gram to categorize positive and negative reviews is difficult.

Step 4: Then they proceeded for the 3-gram model: while using the tri-grams have yet to be able to distinguish between positive and negative feedback.

Step 5: Then they proceeded for the 4-gram model: When compared to other gram models, the 4-gram classifies emotions significantly better. As a result, the deep learning model is built using 4-gram.

Step 6: After that, they calculated the sentiment for each review, with sentiment=1 when the rating is greater than 5, and sentiment = 0 when the rating is less than 5.

Step7: And finally using the given sentiments and cleansed review, they created a deep learning model.

And finally, the steps with which they proceeded are

Step 1: They determined the accuracy of the N-gram model, which was found out to be 80 percent accurate.

Step 2: Then they proceeded to determine the correctness of the light gbm model, which was found out to be 90% accurate.

Step 3: After that, they compared the accuracies of both models and choose the one with the highest accuracy. The Light gbm model was used for recommendation since its accuracy is higher than the N-gram model.

Step 4: Then they created a module for visualization. It essentially provides a graphical representation to display a few key details behind the taken decision.

The development in IT has helped improve quality of students and their ability through intelligent classroom models [6]. It is proposed to develop a recommendation model with adaptive learning for intelligent classroom. A system based on TR-LDA is developed.

A method for question classification based on unsupervised Latent Dirichlet Allocation is suggested for community-based question answering. [8]. This method begins by extracting themes from a vast amount of unlabeled data using unsupervised topic modelling. Following that, the topics learned are used in the training step to determine their relationship with training data's accessible category labels.

Text representation based on the latent topic model is a non-Gaussian job since the observed words and latent topics are multinomial variables, and the topic proportional are Dirichlet variables. To characterize topics proportional, the traditional way of topic model is built by incorporating only one Dirichlet prior. The words in a written document are represented by a random mixture of semantic themes. By incorporating the prior, a new latent variable model which is based on the Dirichlet mixture model learns latent themes and their proportionate. The latent Dirichlet Mixture Model (LDMM) that results is used for both topic as well as to cluster a document. In learning representation, multiple Dirichlet offers a means to build structural latent variables over a large range of topics. [7].

In another work, by calculating individual term ratings and using them to rank and select phrases for the final summary. The challenges of automatic summary of changes made in dynamic text collection are addressed in this work [15]. Using an approach based on the Latent Dirichlet Allocation (LDA) model, the latent topic structures of changes are found. The LDA model is used in order to discover distinct topics where each topic's modified phrases are expected to convey at least one substantial modification.

III. SYSTEM ARCHITECTURE

LDA is a three-layer model that focusses on probability generation. In 2003, Andrew Ng, David Blei and Micheal jordaon proposed the same [19]. This algorithm is developed very much in the recent past with many models similar to this. The parameters in the LDA used are alpha, beta, theta, phi, z and w. Alpha signifies the parameter or Dirichlet prior for the pre-document topic distributions. Beta represents the Dirichlet prior for the per-topic word distribution. Theta reflects a document's subject dispersion. The word distribution for a subject k is represented by phi, and z represents topic for every n^{th} word in the document [16].

The reason for use-to-use LDA is that we have done research on the comparative study with LDA, PCA and GDA [18]. When the number of samples per class is limited or the training data samples the underlying distribution nonuniformly, LDA can be outperformed by PCA. Many real-world applications, particularly in the field of facial recognition, the underlying distributions for the various classes are never known in advance.

As a result, one may argue that determining whether the provided training data is appropriate for the task would be challenging in practice claim is supported by the tests we present. Several of our trials suggest that PCA outperforms LDA, while others show LDA is superior to PCA. When PCA outperforms LDA, the number of training samples per class is modest, but this is to be expected given the data sets utilized in previous experiments.

Where did a comparison for Logistic Regression and the naïve Bayes approach along with the LDA approach, and found out the accuracy of the models. The table I represents the accuracy score comparison between these models.

TABLE I. MODELS & THEIR ACCURACY

Models	Accuracy
LDA	80.80
Logistic Regression	76.93
Naïve Bayes Approach	79.87

We developed our model with many different modules, this was done after through research and taking this into consideration. Our system architecture can be divided into 5 different parts. This can be observed in the Fig. 1.

The parts are:

1. Data Pre-processing
2. Corpus Generation
3. LDA Algorithm
4. Saving the topics
5. Recommendation Algorithm

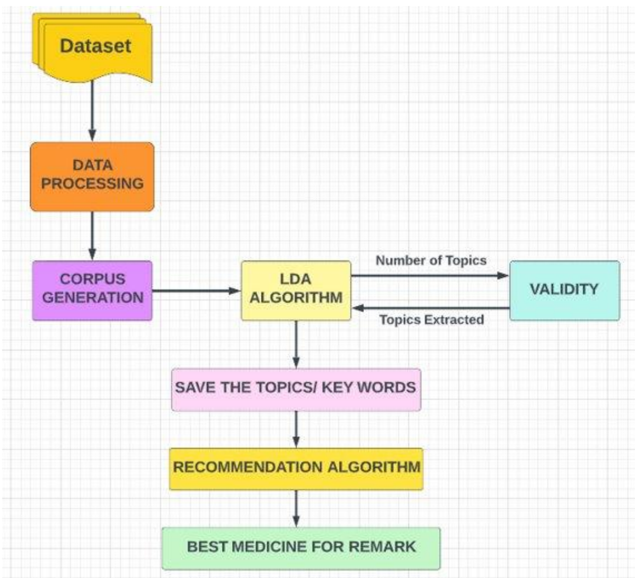

Fig. 1. System Architecture

Now let's discuss about each and every module.

A. *Data Pre processing*

In the data pre-processing, first we consider our data from the dataset. For this the dataset is taken from uci ml repository and this dataset consists of 5 columns and 226526 rows. The five columns present can be classified as an respective feature. These five columns are the drug name, review, condition, rating and useful count. After the extraction of this data from the dataset, we proceed for the pre-processing. In the pre-processing, we neglect and change all the positions of cells that do not have any values and fill these values with NAN. After this process being done be remove all the unwanted elements present and also characterize our dataset to have data that has very high rating. This is done to overcome the faulty prescriptions present.

B. *Corpus generation*

After the data pre-processing, the data is sent to the model for corpus generation. In this step all the unnecessary words such as the stop words present in the review column of the dataset is removed. This is done s recommendation system algorithm uses this review corpus for the classification and the analysis.

C. *LDA Algorithm*

In this paper, we used the Latent Dirichlet allocation we mine the topics for the patient review system. Basically, LDA is a three-layer model that focusses on probability generation. In 2003 by Andrew Ng, David Blei and Micheal jordaon proposed the same [19]. This algorithm is developed very much in the recent past with many models similar to it. The parameters in the LDA used are alpha, beta, theta, phi, z and w[25]. Alpha signifies the parameter or Dirichlet prior for the pre-document topic distributions. Beta represents the Dirichlet prior for the per-topic word distribution. Theta reflects a document's subject dispersion. The word distribution for a subject k is represented by phi, and z represents topic for every n^{th} word in the document [16].

For our model let Q be the set of reviews that are for a specific disease and by A = {a1, a2…aN} the set (consisting of N reviews). That contains only the reviews that are considered as the probable disease for the drugs to be recommended. Let Wi denote the vocabulary used in the extended collection of articles, which includes the researcher's corpus as well as the recommendation corpus (Qi A). The set of words w that appear in reviews Di's corpus Qi and/or the suggestion corpus A is called Wi. The formal representation of a drug model is a topic model generated by applying the LDA algorithm to the texts made up of the reviews for each and every drug and then proceeding for that drug. We should point out that each issue is formally represented by a probability.

Pi(w|k) is a distribution over the studied vocabulary, where w is a range of words. Among the Ki subjects is the word Wi and k. We, particular, get the probability pi(w|1 : Ki) the topic model associated with disease Di. as in equation (1).

$$p(w|d_j) = \frac{nocc(w,d_j) + \frac{\mu nocc(w,Q)}{\sum_w nocc(w,Q)}}{\sum_w nocc(w,d_j) + \mu} \quad (1)$$

To find out the probability of each and every review-based drug, we proceed to find it where nocc indicates the number of times it was used in relation to the words w and Q.

The topic modelling's goal is to uncover topics from a set of automatically papers that have been observed while the topic structure has been hidden. We now go on to the computational task of computing the posterior. Given the observed documents, this is the conditional distribution of the topic structure. On the left side of the equation, the probability of the document topic distribution, the word distribution of each topic, and the topic labels given all words (in all documents) and the hyperparameters and and. We are particularly interested in predicting the probability of topic (z) for a given word for all words and subjects (w). The prior is given by the equation (2).

$$p(\theta, \emptyset, z \mid w, \alpha, \beta) = \frac{p(\theta,\emptyset,z,w,\mid \alpha,\beta)}{p(w,\mid \alpha,\beta)} \quad (2)$$

The LDA technique will be used to extract latent topics from data[26]. After that, the most frequently recurring subjects are gathered and the Jaccard Similarity between the themes of different drugs is used to determine the Jaccard Similarity.

The Jaccard similarity coefficient compares two finite sample sets A and B by dividing the total number of unique items they have in common by the number of items they share. This way, based on remedy or based on patient's symptoms, common words can be extracted and probability can be found. This is given by the equation (3).

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (3)$$

When the Jaccard similarity coefficient exceeds a certain threshold, there is a strong link between the latent features of two drugs.

D. Saving the topics

Here in this step all the topics that are generated after the application of the topic modelling using the LDA algorithm. These topics generated are further saved and used when we try to recommend an ideal drug based on the given review of the patient.

E. The recommendation algorithm

For the recommendation algorithm[23], as we have successfully created the corpus and the topics. Now we take an input review from the user[22]. This taken input review is considered and it gets sent into our model to the most satisfying output as the modelled topic and from that we generated and give out the most probable drug that can be used by the patient.

IV. RESULTS

As mentioned in our system architecture we worked on multiple topic division, where we divided the whole review dataset by 2 topics, then after the visualization we got the result as shown in Fig. 2.

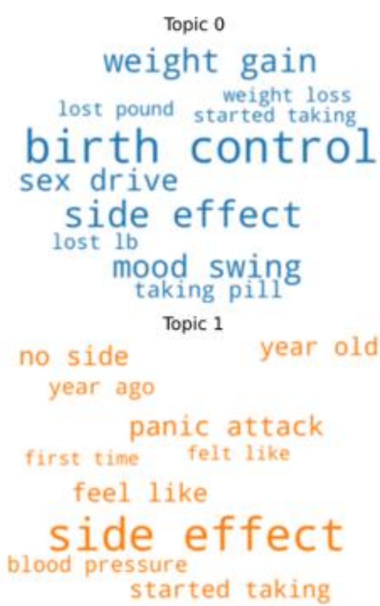

Fig. 2. Review dataset divided into 2 topics

We divided the whole review dataset by 6 topics, then after the visualization we got the topics and topic categories under them as show in Fig. 3.

The size of topics tells us the probability of that topic belongs to certain topic. More the size, more is the probability.

We even used a coherence model with the type c_v. The C v score is calculated using a moving window, one-set partitioning of the top words, and an implicit validation metric based on normalized pointwise mutual information (NPMI) and cosine similarity. With the help of this model, we were able to achieve a score of 80.82.

This coherence index of 80.82 indicates that for the considered dataset. The parameters taken are the reviews and for these parameters the topic modelling is done with the help of latent dirichilet allocation. Now after the allocation when a new review is given then the rate at which we will be able to allocate a new drug for the given review would have a good accuracy of 80.82. This score is the accuracy evaluation of our model.

When we do topic modelling, we assign each and every word into separate topics, these topics are done based on the reviews that are generated. Now as we can see in the Fig. 2, we are able to completely divide our dataset into 2 different topics. Similarly, we were able to divide the entire dataset into 6 topics as shown in Fig. 3. Also, from these topics we are able to find out the most dominant topics. As we can see form Table II the dominant topics with respect to their key words and reviews are given.

TABLE II. DOMINANT TOPICS WITH THEIR CONTRIBUTIONS

S.no	Dominant Topic	Perc Contribution	Topic Keywords	review
0	1.0	0.5922	Side effect, feel like, no side, year ago	tried antidepressant year citalopram fluoxetin…
1	2.0	0.8617	Side effect, blood pressure, dry mouth, lost lb	sor crobn disease done well asacol no complain…
2	0.0	0.2500	Birth control, mood swing, weight gain, sex drive	quick reduction symptom
3	2.0	0.4346	Side effect, blood pressure, dry mouth, lost lb	contrave combine drug used alcohol smoking opi…
4	0.0	0.7492	Birth control, mood swing, weight gain, sex drive	birth control one cycle reading review type si…

Fig. 3. Review dataset divided into 6 topics

```
>>> Progress    :100.0000%
Running time is 70.75422215461731s
>>> Best score is: 0.7499232973115646
>>> The statement given with the symptoms as remark is
She had to use birth control
>>> The most suitable durg that can be prescribed is
Ortho Tri-Cyclen Lo
```

Fig. 4. Model prediction

The Fig. 4 shows how our terminal would look like to take the user input and then get the output for the given drug. In this case the considered statement is ‘She had to use drug control’ and the most probable drug is ‘Ortho Tri-Cyclen Lo’ with a score of 74.99.

V. CONCLUSION & FUTURE SCOPE

With this we conclude that with the help of our model, when a person is not feeling well. They can directly give their review statement and with this review statement, we find out the most probable disease that affected them and also the most probable drug that cures them. Our model works with LDA to get out the most optimal disease and we even worked on why do we need to take LDA. For the future we would even explore more into having multiple features as the columns like the age and the weather conditions to get even more optimal results based on the review.

REFERENCES

[1] Kim, J. K., Jeong, C. G., Li, Q., & Choi, I. Y. (2022). The demand effect analysis of head books and tail books in book recommendation networks. Expert Systems, 39(2), [e12847]. https://doi.org/10.1111/exsy.12847

[2] Lee, H. I., Choi, I. Y., Moon, H. S., & Kim, J. K. (2020). A multi-period product recommender system in online food market based on recurrent neural networks. Sustainability (Switzerland), 12(3), [969]. https://doi.org/10.3390/su12030969

[3] Hung, Man; Xu, Julie; Lauren, Evelyn; Voss, Maren W.; Rosales, Megan N.; Su, Weicong; Ruiz-Negrón, Bianca; He, Yao; Li, Wei; Licari, Frank W. (2019). Development of a recommender system for dental care using machine learning. SN Applied Sciences, 1(7), 785–. doi:10.1007/s42452-019-0795-7.

[4] J.C. Bailón-Elvira, M.J. Cobo, E. Herrera-Viedma, A.G. López-Herrera, Latent Dirichlet Allocation (LDA) for improving the topic modeling of the official bulletin of the spanish state (BOE), Procedia Computer Science, Volume 162, 2019, Pages 207-214, ISSN 1877-0509 https://doi.org/10.1016/j.procs.2019.11.277

[5] Stark, Benjamin, Constanze Knahl, Mert Aydin and Karim O. Elish. “A Literature Review on Medicine Recommender Systems.” International Journal of Advanced Computer Science and Applications (2019):

[6] H. Lin, S. Xie, Z. Xiao, X. Deng, H. Yue, K. Cai, Adaptive Recommender System for an Intelligent Classroom Teaching Model INTERNATIONAL JOURNAL OF EMERGING TECHNOLOGIES IN LEARNING, 14 (5) (2019), pp. 51-63 doi:{10.3991/ijet. v14i05.10251}.

[7] Chien, J.T., Lee, C.H., Tan, Z.H. (2018). Latent Dirichlet mixture model. Neurocomputing, 278:12-22. https://doi.org/10.1016/j.neucom.2017.08.029

[8] [8] Momtazi, S. (2018). Unsupervised Latent Dirichlet allocation for supervised question classification. Information Processing & Management, 54(3):380-393. https://doi.org/10.1016/j.ipm.2018.01.001

[9] Donghui Wang, Yanchun Liang, Dong Xu, Xiaoyue Feng, Renchu Guan, A content-based recommender system for computer science publications, Knowledge-Based Systems, Volume 157, 2018, Pages 1-9, ISSN 0950-7051, https://doi.org/10.1016/j.knosys.2018.05.001.

[10] X. Zheng, Y. Luo, L. Sun, J. Zhang, F. Chen A tourism destination recommender system using users' sentiment and temporal dynamics JOURNAL OF INTELLIGENT INFORMATION SYSTEMS, 51 (3) (2018), pp. 557-578 doi:{10.1007/s10844-018-0496-5}.

[11] Lee, N., Kim, E., & Kwon, O. (2018). Combining TF-IDF and LDA to generate flexible communication for recommendation services by a humanoid robot. Multimedia Tools and Applications, 77(4), 5043-5058. https://doi.org/10.1007/s11042-017-5113-z

[12] Ali, S.I., Amin, M.B., Kim, S., Lee, S. (2018). A hybrid framework for a comprehensive physical activityand diet recommendation system. In Mokhtari, M., Abdulrazak, B., Aloulou, H. (Eds.) Smart homes andhealth telematics, designing a better future: urban assisted living (ICOST 2018), (pp 101–109). SpringerInternational Publishing, Cham.

[13] Bankhele, S., Mhaske, A., Bhat, S. (2017). V., s.: a diabetic healthcare recommendation system. InternationalJournal of Computer Applications,167, 14–18

[14] Leilei Sun, Chuanren Liu, Chonghui Guo, Hui Xiong, and Yanming Xie. 2016. Data-driven Automatic Treatment Regimen Development and Recommendation. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). Association for Computing Machinery, New York, NY, USA, 1865–1874. DOI:https://doi.org/10.1145/2939672.2939866

[15] Manika, K., Nunes, S., Ribeiro, C. (2015). Summarization of changes in dynamic text collections using Latent Dirichlet allocation model. Information Processing & Management, 51(6):809-833. https://doi.org/10.1016/j.ipm.2015.06.002

[16] S. Xie and Y. Feng, "A Recommendation System Combining LDA and Collaborative Filtering Method for Scenic Spot," 2015 2nd International Conference on Information Science and Control Engineering, 2015, pp. 67-71, doi: 10.1109/ICISCE.2015.24.

[17] Wiesner, M., & Pfeifer, D. (2014). Health recommender systems: Concepts, requirements, technical basics and challenges. International Journal of Environmental Research and Public Health, 11, 2580–2607.

[18] Pengfei Yu, Pengcheng Yu and Dan Xu, "Comparison of PCA, LDA and GDA for palmprint verification," 2010 International Conference on Information, Networking and Automation (ICINA), 2010, pp. V1-148-V1-152, doi: 10.1109/ICINA.2010.5636417.

[19] Blei, D. M., Ng, A. Y., Jordan, M. I., Mar. 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3 (4-5), 993-1022. URL http://dx.doi.org/10.1162/jmlr.2003.3.4-5.993

[20] A. M. Martinez and A. C. Kak, "PCA versus LDA," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 2, pp. 228-233, Feb. 2001, doi: 10.1109/34.908974.

[21] P. Ulleri, S. H. Prakash, K. B. Zenith, G. S. Nair and J. M. Kannimoola, "Music Recommendation System Based on Emotion," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021, pp. 1-7, doi: 10.1109/ICCCNT51525.2021.9579689.

[22] P. Devika, K. Jyothisree, P. Rahul, S. Arjun and J. Narayanan, "Book Recommendation System," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021, pp. 1-5, doi: 10.1109/ICCCNT51525.2021.9579647.

[23] S. Singh, R. Rajan, S. Nandini, D. Ramesh and C. P. Prathibhamol, "Friend Recommendation System in a Social Network based on Link Prediction Framework using Deep Neural Network," 2022 2nd International Conference on Intelligent Technologies (CONIT), 2022, pp. 1-7, doi: 10.1109/CONIT55038.2022.9848093.

[24] I. Sri Usha, K. R. Choudary, K. C. R. and T. Sasikala, "Data Mining Techniques used in the Recommendation of E-commerce services," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2018, pp. 379-382, doi: 10.1109/ICECA.2018.8474596.

[25] C. Jayaprakash, Bhushan, B., Vishvanathan, S., and Dr. Soman K. P., "Randomized ICA and LDA Dimensionality Reduction Methods for Hyperspectral Image Classification", arXiv preprint arXiv:1804.07347, 2018.

[26] A. Nawandhar, Dr. Navin Kumar, and Yamujala, L., "Random Subspace Combined LDA Based Machine Learning Model for OSCC Classifier", in Modeling, Machine Learning and Astronomy, Singapore, 2020.