INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

Collaborative Variational Deep Learning for Healthcare Recommendation

AKULA LALITHA
PG Scholar, Department of M.C.A,
S.K.B.R P.G College,
Amalapuram, E.G.Dt., A.P, India.
E-Mail: lalithaakula351@gmail.com

Mr. NAGA. SRINIVASA RAO*
Asst. Professor, Dept of M.C.A,
S.K.B.R P.G College,
Amalapuram, E.G.Dt., A.P, India.
E-Mail:naagaasrinu@gmail.com

Abstract—
There are a number of large medical recommender systems that show the great potential to become tools to augment providers' and patients' health decision. Healthcare recommender systems may recommend the next item that the patient may need or find informative, but current recommender systems often break down when suffering from data sparsity or cold start problems, or not addressing specific problems like poor exploration of the medical terms and relationships between them. To handle these problems, we devise a novel collaborative variational deep learning system to harness the complementary information gathered from multiple sources, such as medical records, lab results, and doctors' notes. The proposed model adopts a nested variational autoencoder (nVAE) to capture the latent representations of medical entries in latent space, at the same time aggregating complementary information in both observation and latent space. The CVCL team proposes a convolutional variational autoencoder (CVAE) to automatically learn the mental representations behind latent users of alcohol information systems, termed as Patient Features. A VAE neural network model is proposed for learning latent feature representations through incorporating user profile in a VAE to further model the latent user representation behind rated items. CVIDL also makes incorporating ratings from experts a crucial step in target latent user representation.

Index Terms**: **Task analysis,Predictive models,Communication channels,Social network services

I. Introduction

Online social media is one of the defining phenomena in this technology-driven era. Platforms, such as Face book and Twitter, are instrumental in enabling global connectivity. 2.46 billion Users are estimated to be now connected and by the year 2020 one third of the global population will be connected. Users of these platforms freely generate and consume information leading to unprecedented amounts of data. Several domains have already recognized the crucial role of social media analysis in improving productivity and gaining competitive advantage. Information derived from social media has been utilized in health-care to support effective service delivery, in sport to engage with fans, in the entertainment industry to complement intuition and experience in business decisions and in politics to track election processes, promote wider engagement with supporters and predict poll outcomes. However, alongside

ISSN: 2057-5688

the benefits, the rapid increase in social media spam contents questions the credibility of research based on analyzing this data. A report by Nexgate estimates that on average one spam post occurs in every 200 social media posts and a more recent study reports that approximately 15% of active Twitter users are automated bots. The growing volume of spam posts and the use of autonomous accounts (social bots) to generate posts raise many concerns about the credibility and representativeness of the data for research. In this report, focus on Twitter and propose a novel, effective approach to detect and filter unwanted tweets, complementing earlier approaches in this direction. Previous studies rely on historical features of tweets that are often unavailable on Twitter after a short period of time, hence not suitable for real-time use. Our approach utilizes an optimized set of readily available features, independent of historical textual features on Twitter. The employed features are categorized as related to the Twitter account, the user or referring to the pair wise engagement between users. A number of machine learning models have been trained. Recursive feature elimination has been employed in order to ascertain the robustness and the discriminative power of each feature. In comparison to an earlier study, the proposed features exhibit stronger discriminative power with more consistent performance across the different learning models. Spam posting users exhibit some evasive tactics, such as posting on average of 4 tweets per day, and tricks to balance the follower–followee relationship. Our analysis shows that an average automated spam posting account posts at least 12 tweets per day within well-defined activity periods. The activity pattern resembles the staircase function exhibiting surges of intermittent activities. Our study contributes (a) a new set of lightweight features suitable for real-time detection of spammers on Twitter and (b) an additional dataset source1 offering an insight into the behavior of spam users on Twitter to support further studies.

II. RELATED WORK

A mapping scheme for collaborative filtering problem to text analysis

Zhong and Lipropose a unified method combining the latent and external features of users and items for accurate recommendation. A mapping scheme for collaborative filtering problem to text analysis problem is introduced, and the probabilistic latent semantic analysis was used to calculate the latent features based on the historical rating data. The main advantages of this technique over standard memory-based methods are the higher accuracy, constant time prediction, and an explicit and compact model representation.

Collaborative filtering and probabilistic topic modeling

Wang and Bledevelop an algorithm to recommend scientific articles to users of an online community. This approach combines the merits of traditional collaborative filtering and probabilistic topic modeling. It provides

ISSN: 2057-5688

an interpretable latent structure for users and items, and can form recommendations about both existing and newly published articles. Rendle et al., investigate the most common scenario with implicit feedback (e.g. clicks, purchases). There are many methods for item recommendation from implicit feedback like matrix factorization (MF) or adaptive knearest-neighbor (kNN). Even though these methods are designed for the item prediction task of personalized ranking, none of them is directly optimized for ranking. They present a generic optimization criterion BPR-Opt for personalized ranking that is the maximum posterior estimator derived from a Bayesian analysis of the problem. It provides a generic learning algorithm for optimizing models with respect to BPR-Opt. The learning method is based on stochastic gradient descent with bootstrap sampling. It shows how to apply our method to two state-of-the-art recommender models: matrix factorization and adaptive kNN.Pan et al., studied a new recommendation problem called heterogeneous implicit feedbacks (HIF), where the fundamental challenge is the uncertainty of the examination records. They design a novel preference learning algorithm to learn a confidence for each uncertain examination record with the help of transaction records. Specifically, it generalize Bayesian personalized ranking (BPR), a seminal pairwise learning algorithm for homogeneous implicit feedbacks, and learn the confidence adaptively, which is thus called *adaptive Bayesian personalized ranking* (ABPR). ABPR has the merits of uncertainty reduction on examination records and accurate pairwise preference learning on implicit feedbacks.

Collaborative filtering

Collaborative filtering (CF) is a successful approach commonly used by many recommender systems. Conventional CF-based methods use the ratings given to items by users as the sole source of information for learning to make recommendation. However, the ratings are often very sparse in many applications, causing CF-based methods to degrade significantly in their recommendation performance. To address this sparsity problem, auxiliary information such as item content information may be utilized. Collaborative topic regression (CTR) is an appealing recent method taking this approach which tightly couples the two components that learn from two different sources of information. Nevertheless, the latent representation learned by CTR may not be very effective when the auxiliary information is very sparse. To address this problem, Wang et al.,generalize recent advances in deep learning from i.i.d. input to non-i.i.d. (CF-based) input and propose a hierarchical Bayesian model called collaborative deep learning (CDL), which jointly performs deep representation learning for the content information and collaborative filtering for the ratings (feedback) matrix.

Deep learning

Li et al.,learn effective latent representations via deep learning. Deep learning models have emerged as very appealing in learning effective representations in many applications. In particular, they propose a general deep architecture for CF by integrating matrix factorization with deep feature learning. They provide a natural instantiations of our architecture by combining probabilistic matrix factorization with marginalized denoising stacked auto-encoders.

Collaborative deep ranking

Ying et al.,propose collaborative deep ranking (CDR), a hybrid pair-wise approach with implicit feedback, which leverages deep feature representation of item content into Bayesian framework of pair-wise ranking model. Dong et al.,utilize advances of learning effective representations in deep learning, and propose a hybrid model which jointly performs deep users and items' latent factors learning from side information and collaborative filtering from the rating matrix. Purushotham et al.,propose a novel hierarchical Bayesian model which jointly incorporates topic modeling and probabilistic matrix factorization of social networks. A major advantage of our model is to automatically infer useful latent topics and social information as well as their importance to collaborative filtering from the training data. Chen et al., propose a novel context-awarehierarchical Bayesian method. First, they propose theuse of spectral clustering for user-item sub-grouping, sothat users and items in similar contexts are grouped. Then propose a novel hierarchical Bayesian modelthat can make predictions for each user-item subgroup, this model incorporate not only topic modeling to mineitem content but also social matrix factorization to handle ratings and social relationships. Wang et al., develop a novel hierarchical Bayesian model called Relational Collaborative Topic Regression (RCTR), which extends CTR by seamlessly integrating the user-item feedback information, item content information, and network structure among items into the same model. Li and She proposes a Bayesian generative model called collaborative variational autoencoder (CVAE) that considers both rating and content for recommendation in multimedia scenario. The model learns deep latent representations from content data in an unsupervised manner and also learns implicit relationships between items and users from both content and rating. Unlike previous works with denoising criteria, the proposed CVAE learns a latent distribution for content in latent space instead of observation space through an inference network and can be easily extended to other multimedia modalities other than text.

III. EXISTINGSYSTEM

Collaborativetopic regression (CTR) is a probabilistic graphical model,which seamlessly integrates the conventional MF model withprobabilistic topic modeling, and can generate more accuraterecommendations based on item contents and other user'sratings. CTR does not exploit user information andcannot learn reliable latent user representations. CTRSMF and C-CTR-SMF2 integratedCTR with SMF model using a strategy that is similar toSoRec, where the social relationships are simultaneouslyfactorized with the rating matrix. However, they do not revealthe underlying relations among users due to the lack of physicalexplanation.These twomethods assume that the social interactions of users usuallyfollow topically similar contents, so they are very sensitiveto different type of datasets and the prediction accuracy mayvary with the distributions of datasets.For social recommendation,CTRSTE integrates user ratings, item contentsand trust ensemble into CTR, which is simple in algorithmicprinciple, but its representation capability is limited

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688

due toLDA model, and the latent representation learned is not effectiveenough when the side information is very sparse. Several works utilize deep learning models tohelp perform the CF task in CTR, due to the non-linearnature of neural networks, such as CDL, CVAE,CAVAE and CTRDAE. All of CDL, CVAE andCAVAE combine stacked DAE (SDAE) or VAE with CTR,and enable themselves to balance the influences of user ratingsand item content, but the auxiliary information of user profile is not considered at al.

IV. PROPOSED SYSTEM

Proposesnovel collaborative variational deep learningmodel (CVDL) to exploit multi-sourced information for providing appropriate for recommendation. This project constructs the generative processesof users and items through a neural variational framework,which enables our model to capture non-linear latent representationsof both users and items.In CVDL, the items contents and user profiles/relations aregenerated by their latent variables, and rating predictions aregenerated jointly through both latent item and user variables.Latent item variables are incorporated with item contentsinformation through latent content variables by employingan additional VAE model, due to the variety of Latent uservariables are linked with user profiles/relations via latent trustvariables by a standard VAE model. Then, the users' anditems' latent vectors are fed into the PMF model to learn theuser-item relations, and finally predict the ratings.

V. MODULES

1. Feature Extraction
2. Optimization
3. Prediction

Feature Extraction:

- Most MF-based models assume that the prior distributions of user and item latent factors are standard Gaussian distributions, and predict rating only through rating data.
- To extract more effective latent vectors from side information, CVDL incorporates both user's and item's side information into feature extraction, which can make positive contributions to the further rating prediction.
- To learn better user and item features, two variation neural networks are built. The generative process of CVDL is similar to the deep latent Gaussian model.
- The inference model is an encoder network corresponding to the one in the generative model. For user, the inference process is to approximate the intractable posterior distribution which is determined by the generative network.
- Using the Stochastic Gradient Variational Bayes (SGVB) estimator, the posterior of latent user profile variable can be approximated by a tractable variational distribution

Optimization:

- Through the CVDL model, this work utilize maximum a poster probability estimator to learn parameters of our model.

INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

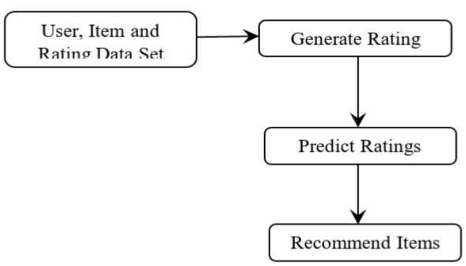
ISSN: 2057-5688

- The objective function includes three parts: the latent loss, the regularization loss and the KL loss.

Prediction:

- After the optimal parameters are learned, CVDL can be employed for in-matrix (non cold-start) and out-matrix (cold-start) prediction.
- Assumed that O is the observed rating data, and both types of predictions can be evaluated by point estimation.

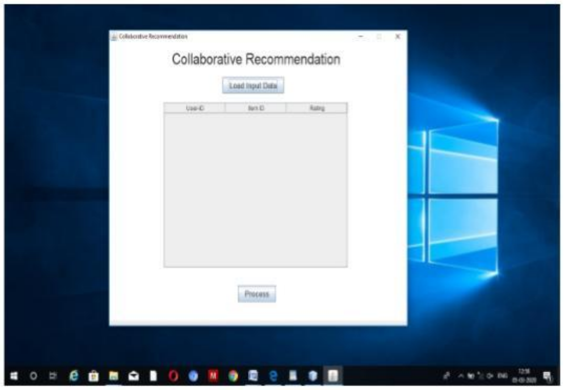
VI. Architecture:

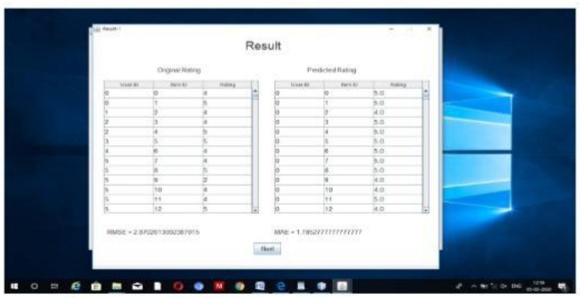


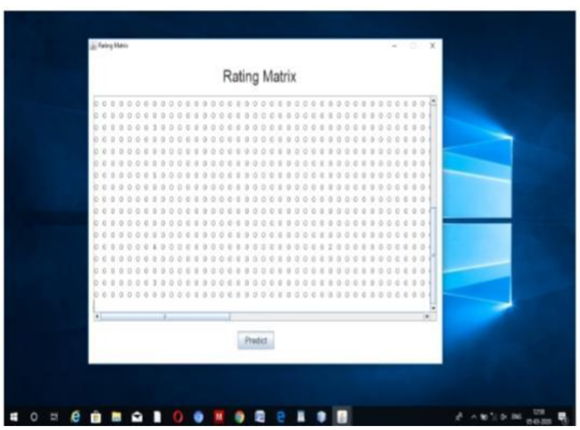
VII.OUTPUT

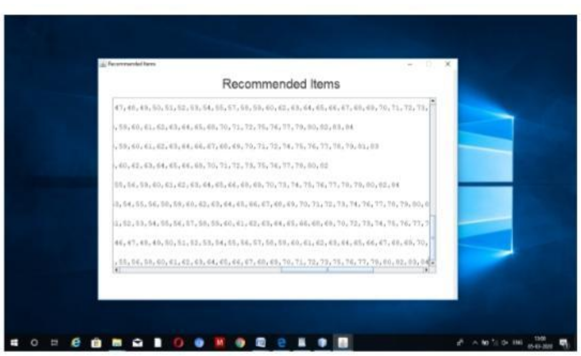
INTERNATIONAL JOURNAL OF TECHNO-ENGINEERING
IJTE

ISSN: 2057-5688









VIII. CONCLUSION

This project propose a hybrid collaborative deep learningmodel (CVDL) for healthcare recommendation, which jointlymodels the generation of item content and user profile whileextracting the implicit relationships between items and userscollaboratively. On the one hand, the proposed CVDL can beconsidered as a Bayesian probabilistic generative model, andits variational inference is deduced from a stochastic gradientvariational Bayesian model. CVDL unifies the collaborativeinformation, item content and user profile through deeplearning model and graphical model, which leads to robustrecommending performance. On the other hand, our CVDLcan unify multimedia in different forms for recommendation,due to its inference of stochastic distribution in latentspace instead of observation space. Experimental results haveshown the proposed CVDL can significantly outperform thecurrent CTR approaches for recommendation jointly withitem content and user profile, with more robust performance,especially on the healthcare dataset. CVDL is proposed by utilizing MLP as the inference andgeneration models, which can also fit into other deep learningmodels, depending on the data type of additional information.

IX. REFERENCES

1. Improving product marketing by predicting early reviewers on E-Commerce websites

S. Kodati, M. Dhasaratham, V. V. S. S. Srikanth, and K. M. Reddy, “Improving product marketing by predicting early reviewers on E-Commerce websites,” Deleted Journal, no. 43, pp. 17–25, Apr. 2024, doi: 10.55529/ijrise.43.17.25.

2. Kodati, Dr Sarangam, et al. “Classification of SARS Cov-2 and Non-SARS Cov-2 Pneumonia Using CNN.” Journal of Prevention, Diagnosis and Management of Human Diseases (JPDMHD) 2799-1202, vol. 3, no. 06, 23 Nov. 2023, pp. 32–40, journal.hmjournals.com/index.php/JPDMHD/article/view/3406/2798, https://doi.org/10.55529/jpdmhd.36.32.40. Accessed 2 May 2024.

3. V. Srikanth, “CHRONIC KIDNEY DISEASE PREDICTION USING MACHINE LEARNING ALGORITHMS,” IJTE, pp. 106–109, Jan. 2023, [Online]. Available: http://ijte.uk/archive/2023/CHRONIC-KIDNEY-DISEASE-PREDICTION-USING-MACHINE-LEARNING-ALGORITHMS.pdf

4. V. SRIKANTH, “DETECTION OF PLAGIARISM USING ARTIFICIAL NEURAL NETWORKS,” International Journal of Technology and Engineering, vol. XV, no. I, pp. 201–204, Feb. 2023, [Online]. Available: http://ijte.uk/archive/2023/DETECTION-OF-PLAGIARISM-USING-ARTIFICIAL-NEURAL-NETWORKS.pdf

5. V. SRIKANTH, “A REVIEW ON MODELING AND PREDICTING OF CYBER HACKING BREACHES,” IJTE, vol. XV, no. I, pp. 300–302, Mar. 2023, [Online]. Available: http://ijte.uk/archive/2023/A-REVIEW-ON-MODELING-AND-PREDICTING-OF-CYBER-HACKING-BREACHES.pdf

6. S. Kodati, M. Dhasaratham, V. V. S. S. Srikanth, and K. M. Reddy, “Detection of fake currency using machine learning models,” Deleted Journal, no. 41, pp. 31–38, Dec. 2023, doi: 10.55529/ijrise.41.31.38.

7. “Cyberspace and the Law: Cyber Security.” IOK STORE, iokstore.inkofknowledge.com/product-page/cyberspace-and-the-law. Accessed 2 May 2024.

8. “Data Structures Laboratory Manual.” IOK STORE, www.iokstore.inkofknowledge.com/product-page/data-structures-laboratory-manual. Accessed 2 May 2024.

9. Data Analytics Using R Programming Lab.” IOK STORE, www.iokstore.inkofknowledge.com/product-page/data-analytics-using-r-programming-lab. Accessed 2 May 2024.

10. V. Srikanth, Dr. I. Reddy, and Department of Information Technology, Sreenidhi Institute of Science and Technology, Hyderabad, 501301, India, “WIRELESS SECURITY PROTOCOLS (WEP,WPA,WPA2 & WPA3),” journal-

ISSN: 2057-5688

article, 2019. [Online]. Available: https://www.jetir.org/papers/JETIRDA06001.pdf

10. V. SRIKANTH, "Secured ranked keyword search over encrypted data on cloud," IJIEMR Transactions, vol. 07, no. 02, pp. 111–119, Feb. 2018, [Online]. Available: https://www.ijiemr.org/public/uploads/paper/1121_approvedpaper.pdf

11. V. SRIKANTH, "A NOVEL METHOD FOR BUG DETECTION TECHNIQUES USING INSTANCE SELECTION AND FEATURE SELECTION," IJIEMR Transactions, vol. 06, no. 12, pp. 337–344, Dec. 2017, [Online]. Available: https://www.ijiemr.org/public/uploads/paper/976_approvedpaper.pdf

12 . SRIKANTH MCA, MTECH, MBA, "ANALYZING THE TWEETS AND DETECT TRAFFIC FROM TWITTER ANALYSIS," Feb. 2017. [Online]. Available: http://ijmtarc.in/Papers/Current%20Papers/IJMTARC-170309.pdf

14 Srikanth, V. 2018. "Secret Sharing Algorithm Implementation on Single to Multi Cloud." International Journal of Research 5 (01): 1036–41. https://journals.pen2print.org/index.php/ijr/article/view/11641/11021.

5. K. Meenendranath Reddy, et al. Design and Implementation of Robotic Arm for Pick and Place by Using Bluetooth Technology. No. 34, 16 June 2023, pp. 16–21, https://doi.org/10.55529/jeet.34.16.21. Accessed 20 Aug. 2023.

16. Babu, Dr P. Sankar, et al. "Intelligents Traffic Light Controller for Ambulance." Journal of Image Processing and Intelligent Remote Sensing(JIPIRS) ISSN 2815-0953, vol. 3, no. 04, 19 July 2023, pp. 19–26, journal.hmjournals.com/index.php/JIPIRS/article/view/2425/2316, https://doi.org/10.55529/jipirs.34.19.26. Accessed 24 Aug. 2023.

17. S. Maddilety, et al. "Grid Synchronization Failure Detection on Sensing the Frequency and Voltage beyond the Ranges." Journal of Energy Engineering and Thermodynamics, no. 35, 4 Aug. 2023, pp. 1–7, https://doi.org/10.55529/jeet.35.1.7. Accessed 2 May 2024.

18. K. Meenendranath Reddy, et al. Design and Implementation of Robotic Arm for Pick and Place by Using Bluetooth Technology. No. 34, 16 June 2023, pp. 16–21, https://doi.org/10.55529/jeet.34.16.21. Accessed 20 Aug. 2023