Recommendation System using Graph Neural Networks

Mughees Asif (180288337)
Supervisor: **Dr Angadh Nanjangud**
MSc. Artificial Intelligence
School of Electronic Engineering and Computer Science
Queen Mary, University of London

Abstract—In recent years, recommendation systems have been developed with graph neural networks (GNNs) to expedite the aggregation process of macro (e.g., topological structure) and micro (e.g., node information) operations to enrich information filtering capabilities. Additionally, using neural networks for recommendation tasks on graph-like data structures has been proven to elicit meaningful representations of user-item relationships. However, the representation learning process is not linear as social relations augmented with item interactions, both need to be considered for optimal results. This research project aims to address this issue by proposing a novel framework that is capable of using multiple preferences for outlining recommendations to users. The framework is also split into three variations that are pitted against published academic recommendation systems. The training of the models is done by using two real-world datasets that contain user-item and user-user information.

Index Terms—Artificial intelligence; Recommendation system; Graph Neural Network; Social Networks

I. INTRODUCTION

Recommendation systems are a vital tool incorporated into every business model to streamline user experiences and mitigate information overload by pinpointing areas of interest. Indeed, the business model of global organisations including Amazon, Google, Netflix, and Spotify relies on propagating items/services of interest to users, not only to increase customer satisfaction and business profitability but to also simplify the addition of new services that would be pertinent to a specific user base.

The basic premise of Graph Neural Networks (GNNs) is to leverage deep learning methodologies on non-Euclidean data structures for node classification and link estimation [1]. Graphs are intrinsic data structures, where each node can be denoted as a set of objects and the edges are defined as the relationship between the nodes and can be used by GNNs to model relationships based on user-to-user and/or user-to-item interactions.

Recently, the representation of these relationships with the expressive power of graphs has been receiving wide attention in research to enable predictive modelling through recommendation systems [2]–[4]. The process of learning information represented in a graph-like structure is underlined by the iterative aggregation of features from local graph neighbourhoods, thus demonstrating a powerful impact on representation learning (discussed further in Section II-A) [5, 6].

This research project aims to develop, explore, and test a GNN framework for recommendation purposes by using social information on two real-world datasets. The remaining project is structured into several sections: Section II details the preliminaries for GNNs techniques and recommendation systems. Section III explains the architecture of the final framework and the datasets used for training and testing. Section IV and V highlight the achieved results with accompanying empirical analysis and suggested future work, and finally, Section VI delivers a conclusion to the research project.

II. BACKGROUND

A. Graph Neural Networks (GNNs)

GNNs have been popularised recently by the advancement made in performances related to many tasks including predicting protein structure [7, 8], analysing dynamic systems [9], and predicting knowledge through base completion [10]. A graph $\mathcal{G}$ is represented by the notation $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is representative of the set of available nodes and $\mathcal{E}$ is the set of edges. Furthermore, $v_i \in \mathcal{V}$ is a node with an edge $e_{ij} = (v_i, v_j) \in \mathcal{E}$ extending from v_j to v_i, and the local neighbourhood of the node v can be denoted as $\mathcal{N}(v) = \{u \in \mathcal{V} | (v, u) \in \mathcal{E}\}$ [6]. There are two main categories of graphs:

1) **Directed/undirected graphs**: A directed graph consists of edges that only point in one direction,

while an undirected graph consists of edges that can point in both directions.

2) **Homogeneous/heterogeneous graphs**: Homogeneous graphs consist of one type of nodes and edges, while heterogeneous graphs can have different types of nodes and edges.

The main intuition behind GNNs is the iterative aggregation of the feature information from a neighbourhood of nodes that are integrated with the information from the current node during the propagation process [11]. The architecture of GNNs is based on stacking several of such layers that allow propagation through aggregation and update operations [6]:

$$\text{Aggregation}: \quad \mathbf{n}_v = \text{Aggregator}_l \left(\left\{ \mathbf{h}_u^l, \forall u \in \mathcal{N}_v \right\} \right) \tag{1}$$

$$\text{Update}: \quad \mathbf{h}_v^{(l+1)} = \text{Updater}_l \left(\left\{ \mathbf{h}_v^{(l)}, \mathbf{n}_v^{(l)} \right\} \right) \tag{2}$$

where:

$\mathbf{h}_u^l$ = represents the node at the lth layer
Aggregator_l = aggregation operation
Updater_l = updating operation
$\mathcal{N}(v)$ = neighbourhood of the node

The aggregation step (Eq. 1) enables the treatment of each neighbour equally with a mean-pooling operation or differentiation of the importance of the neighbours through the attention framework [12, 13]. The update step (Eq. 2) is ". . . the representation of the central node and the aggregated neighbourhood [that] will be integrated into the updated representation of the [current] node" [6]. Various integration techniques exist to alleviate misrepresentation between the aggregator and updater operations including gated mechanisms [14] and applying non-linear transformations [12].

B. Recommendation Systems

Recommendation systems learn the preferences of a user by analysing the user-item interaction or static features that highlight the interest of the user towards a particular item. The task can be defined as [6]:

$$y_{u,i} = f(h_u^*, h_i^*) \tag{3}$$

where:

$i \in I$ = item in the directory
h_u^* = user representation
h_i^* = item representation
$f(\cdot)$ = score function
$y_{u,i}$ = preference score for a user on an item

One of the main breakthroughs was the integration of social information about the user in the recommendation system, where the underlying social relations of the user helped to boost performance [5, 6, 15, 16]. The social recommendation framework assumes that the users' social relationships present a novel approach to enhancing user representations by leveraging the social influence theory that connected people influence each other [17]. The social network can be integrated into a user-item graph as a unified graph where the social information and the aggregation operation are combined during the propagation process.

III. METHODOLOGY

A. Problem formulation

TABLE I
MATHEMATICAL NOTATIONS

Symbol	Description
r_{ij}	Rating values for the user u_i and the item v_j
$\mathbf{p}_i$	Embedding of the user u_i
$\mathbf{q}_j$	Embedding of the item v_j
$\mathbf{e}_r$	Opinion embedding of the rating level r i.e., 5-star rating, $r \in 1, 2, 3, 4, 5$
d	Length of the embedding vector
$C(i)$	Set of items the user interacted with
$N(i)$	Set of social friends that the user connected with
$B(j)$	Set of users that interacted with item v_j
$\mathbf{h}_i^I$	Item-space user latent factor
$\mathbf{h}_i^S$	Social-space user latent factor
$\mathbf{h}_i$	User latent factor
$\mathbf{z}_j$	Item latent factor
$\mathbf{x}_{ia}$	Opinion-aware interaction representation of item v_a for user u_i
$\mathbf{f}_i^I$	Opinion-aware interaction representation of user u_t for item v_j
α_{ia}	Item attention of item v_a
μ_{jt}	User attention of user u_t
β_{io}	Social attention of the neighbouring user u_o
r'_{ij}	Predicting rating of the item v_j by user u_i
$\mathbf{T}$	User-user social graph
$\mathbf{R}$	User-item rating graph
$\oplus$	concatenation of two vectors

Consider a system comprised of sets of users $U = \{u_1, u_2, \ldots, u_n\}$ and items $V = \{v_1, v_2, \ldots, v_m\}$, where $R \in \mathbb{R}^{n \times m}$ is the user-item graph represented as a rating matrix. If a rating is passed from u_i to v_j, r_{ij} equals the rating score, and conversely for an unknown rating $r_{ij} = 0$. Consider also $\mathcal{O} = \{\langle u_i, v_j \rangle \mid r_{ij} \neq 0\}$ as the set

of observed/known ratings, and $\mathcal{T} = \{\langle u_i, v_j \rangle \,|\, r_{ij} = 0\}$ as the set of unknown ratings. Additionally, users can create social relations, where $\mathbf{T} \in \mathbb{R}^{n \times n}$ highlights the user-user social graph and is given a value of 0 or 1 dependent on if u_j has a relation to u_i. Therefore, in combination with a user-item graph $\mathbf{R}$ and $\mathbf{T}$, the aim of the project is to predict the missing value in $\mathbf{R}$.
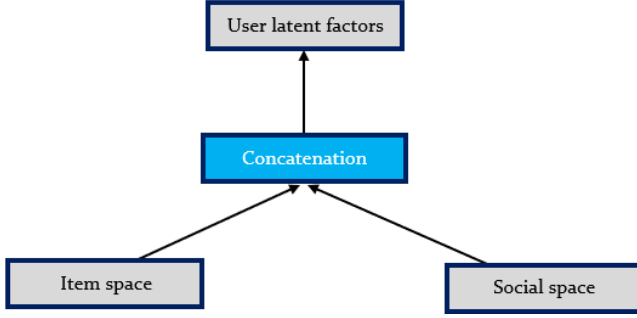
B. User modelling

Fig. 1. User modelling

The user modelling operation aims to learn the latent factors $\mathbf{h}_i \in \mathbb{R}^d$ of users u_i. Highlighted by Fig. 1, this operation requires the concatenation of two latent factors to obtain the user latent factors $\mathbf{h}_i$: item space user latent factor $\mathbf{h}_i^I \in \mathbb{R}^d$ from the user-item graph, and a social space user latent factor $\mathbf{h}_i^S \in \mathbb{R}^d$ from the social graph. A Multi-Layer Perceptron (MLP) is introduced to concatenate the two vectors, resulting in[1]:

$$\mathbf{h}_i = \sigma(W_l \cdot c_{l-1} + b_i) \tag{4}$$

where, $c_l = \left[\mathbf{h}_i^I \oplus \mathbf{h}_i^S \right]$.

1) Item space: The item space operation utilises the interactions between the users and items and also the users' preferences regarding the item, all encoded as a user-item graph. As shown in Fig. 2, the main premise is to learn item-space user latent factor $\mathbf{h}_i^I$. This can be defined in the classic $y = mx + c$ equivalent linear function as:

$$\mathbf{h}_i^I = \sigma \left(\mathbf{W} \cdot A_{\text{item}} \left(\mathbf{x}_{ia}, \forall a \in C(i) \right) + \boldsymbol{b} \right) \tag{5}$$

where:

σ = rectified linear unit function
$\mathbf{W}$ = weights of the network
$\boldsymbol{b}$ = bias of the network
A_{item} = the aggregation operation
$C(i)$ = the items the user interacted with
$\mathbf{x}_{ia}$ = representation vector to include users' opinion

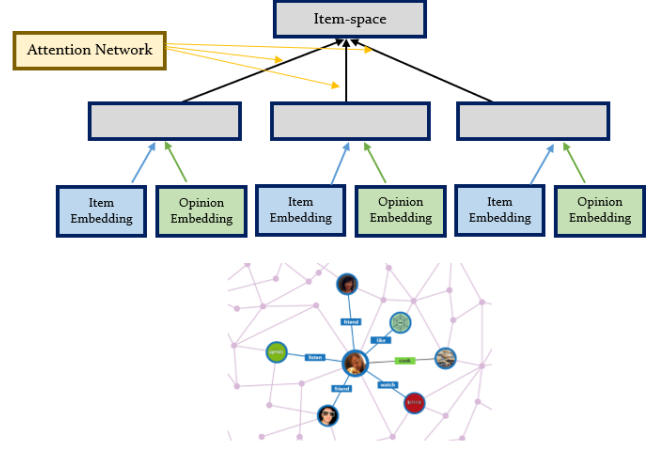[1]An accessible side-note on MLP can be found here.

Fig. 2. Item space operation

The representation vector $\mathbf{x}_{ia}$ contains the rating r a user gives an item from a 5-star rating i.e., $r \in 1, 2, 3, 4, 5$, or simply a dense embedding vector $\mathbf{e}_r$[2]. This mapping can be used to discern the users' preferences and return user-item relationship latent factors. This is achieved by using the direct item embedding $\mathbf{q}_a$, and retrieving the opinion embedding $\mathbf{e}_r$ using an MLP. The output is the representation vector that includes the users' opinion on a certain item $\mathbf{x}_{ia}$:

$$\mathbf{x}_{ia} = g_v \left(\mathbf{q}_a \oplus \mathbf{e}_r \right) \tag{6}$$

where:

g_v = the MLP

The aggregation operation is the element-wise mean of the vectors $\mathbf{x}_{ia}$ and $\forall a \in C(i)$ by using a linear approximation of a localised convolution [6]. However, this produces subpar results as not each interaction should have the same weight due to differing preferences of the user towards the items. To counteract this linearity, a 2-layer attention mechanism intervenes, where each interaction is given an individual weight dependent on the user's interest in the item. Substituting this into Eq. 5 yields:

$$\mathbf{h}_i^I = \sigma \left(\mathbf{W} \cdot \left\{ \sum_{a \in C(i)} \alpha_{ia} \mathbf{x}_{ia} \right\} + \boldsymbol{b} \right) \tag{7}$$

where, α_{ia} is representative of the interaction between the user u_i and the item v_a.
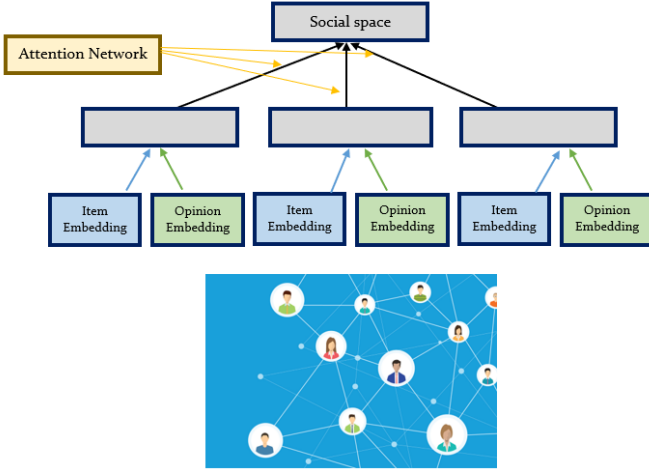
Fig. 3. Social space operation

C. Item modelling

This section highlights the incorporation of the item latent factors $\mathbf{z}_j$ b using information from the user-item graph as shown in Fig. 4. For each item, the users' preferences are aggregated i.e., the mean of the ratings ranging from 1 to 5 for all items ($R = r \in 1, 2, 3, 4, 5$), and using a MLP, the two vectors holding information regarding plain user embedding $\mathbf{p}_t$ and opinion embedding $\mathbf{e}_r$ are used to develop a user representation $\mathbf{f}_{jt}$:

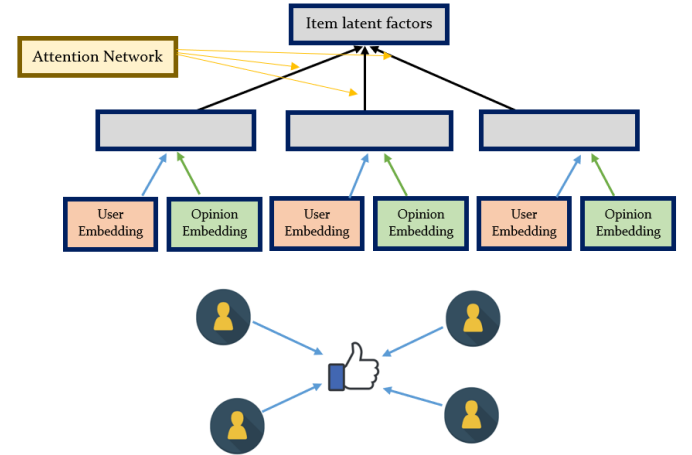$$\mathbf{f}_{jt} = g_u(\mathbf{p}_t \oplus \mathbf{e}_r) \tag{10}$$

Fig. 4. Item modelling

2) Social space: Well-established social correlation theories [18, 19] have highlighted the impact of a user's social circle where a specific category of items appeals to a specific group the user is associated with. This shows the influence of the social circle on a user's decision-making process. To encode heterogeneous strengths of social relations into the GNN, an attention mechanism is introduced "... to select social friends that [can] characterise users' social information and then aggregate [the] information" [6]. This is achieved by incorporating social space latent factors $\mathbf{h}_i^S$ that are an aggregation of the neighbouring users item space $A_{\text{neighbours}}(\cdot)$, as defined in Section III-B1 and displayed in Fig. 3:

$$\mathbf{h}_i^S = \sigma \left(\mathbf{W} \cdot A_{\text{neighbours}} \left(\mathbf{h}_o^I, \forall o \in N(i) \right) + \boldsymbol{b} \right) \tag{8}$$

The aggregation operation on the user's neighbours follows a similar structure to the item space aggregation operation where a mean operator takes the element-wise mean of the vectors [6]. To diminish the impact of assuming that all neighbours contribute equally, an attention mechanism using a 2-layer neural network is introduced that develops the correlation between user-to-user and user-to-item interaction:

$$\mathbf{h}_i^S = \sigma \left(\mathbf{W} \cdot \left\{ \sum_{o \in N(i)} \beta_{io} \mathbf{h}_o^I \right\} + \boldsymbol{b} \right) \tag{9}$$

where, βio is representative of the interaction between the user's social circle u_i and the item v_a.

The latent factors are derived similarly to the previous two sections by the introduction of an attention mechanism:

$$\mathbf{z}_j = \sigma \left(\mathbf{W} \cdot A_{\text{users}} \left(\mathbf{f}_{jt}, \forall t \in B(j) \right) + \boldsymbol{b} \right) \tag{11}$$

$$\mathbf{z}_j = \sigma \left(\mathbf{W} \cdot \left\{ \sum_{t \in B(j)} \mu_{jt} \mathbf{f}_{jt} \right\} + \boldsymbol{b} \right) \tag{12}$$

where, μ_{jt} is "... to capture heterogeneous influence from user-item interactions on learning item latent factor" [6].

D. Ratings

Lastly, the rating predictions are obtained as displayed in Fig. 5. The latent factors gained from the previous sections are concatenated, $\mathbf{h}_i \oplus \mathbf{z}_j$, and passed through the MLP to get the final ratings r'_{ij}:

$$r'_{ij} = \mathbf{w}^T \cdot \boldsymbol{g}_{l-1} \tag{13}$$

---

[2] An informative and succinct explanation of dense layers can be found here.
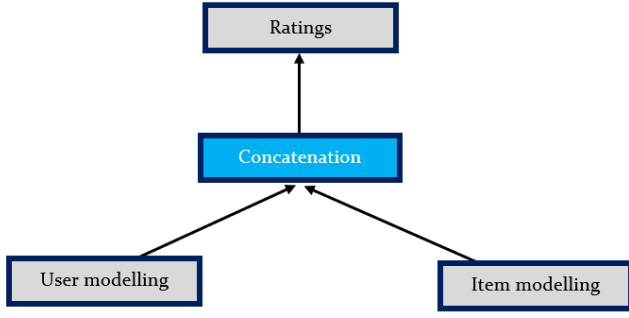
Fig. 5. Ratings

E. Objective function

The objective/loss function was designed to be optimised by an enhanced version of the vanilla gradient descent, in the form of Root Mean Squared Propagation (RMSProp). The premise of the optimisation is to establish an adaptive learning rate that responds to the gradient of the learning function by storing and using the moving average of the squared gradients for each weight initialisation [20]. The mathematical notation is:

$$\text{Loss} = \frac{1}{2|\mathcal{O}|} \sum_{i,j \in \mathcal{O}} \left(r'_{ij} - r_{ij} \right)^2 \qquad (14)$$

where, $\mathcal{O}$ is the observable ratings, and $r_{i,j}$ is the ground truth as expressed by the user.

F. Datasets

Two datasets from prominent product review websites were used: Epinions and Caio[3]. The structure of the datasets follows suit including information on allowing the user to add friends, product names, varieties and the rating score. Table II shows the data points that were used to establish the efficacy of the GNN model in discerning suitable recommendations by using densely embedded user-to-user and user-to-item information.

TABLE II
DATASETS' STATISTICS

Characteristic	Epinions	Caio
Users	$7,000$	$18,000$
Items	$100,000$	$250,000$
Ratings	$7,000$	$18,000$

[3]Available here.

G. Training

The model was built using the `PyTorch` framework and the Python programming language, as per the industry standard. The three embeddings (user, item and opinion) were stochastically initialised to allow the model to learn concurrently. For training purposes and adhering to a finite computational cost, dropout was added to the network that neglected specific weights which did not add sufficient value to the overall result i.e, anomalous and minute updates. Each component of the system contained a maximum of three hidden layers (same size as the embedding size) with the ReLU activation function. The embedding size was chosen from a range of 8 to 256, the batch size was kept to 32, and the variable learning rate ranged from $1e^{-3}$ to 0.1.

H. Testing

The model was tested against industry benchmark recommendation systems sourced from published academic research and against variants of the research framework:

- **Social Recommendation** (SoRec): Hao et al. developed a recommendation system to use a factor analysis approach based on probabilistic matrix factorisation to incorporate users' social network information [21].
- **Social Regularization** (SoReg): Consequently, Hao et al. also suggested constraining the matrix factorisation objective function by employing average-based regularisation (ratings from all the social connections), and an individual-based regularisation (constrains the framework to use only similar social connections) [22].
- **Deep neural network model on Social Relations for recommendation** (DeepSoR): Fan et al. suggested improving the accuracy of a recommender system by integrating a deep neural network to learn complex and non-intrinsic features from social relations [23].
- **Graph Convolutional Matrix Completion** (GC-MC): Berg et al. proposed a matrix completion framework: an auto-encoder that produced latent features of users and items at the node level on a bipartite interaction graph, which is then decoded by reconstructing rating links bi-linearly [24].
- **Variants**:
  1) X_a: Framework as described in this section.
  2) X_b: The item-space operation is disabled to evaluate the elimination of user-item interactions on the predicted ratings.

5

3) X_c: Same as X_b but instead of the item-space operation, the social-space operation is eliminated.

I. Evaluation

The evaluation criteria included comparing the following two negatively-oriented metrics (the lowest numerical value is considered the best) for analysing the classifiers' precision [25]:

- **Mean Absolute Error** (MAE): Measures the average magnitude of the prediction and actual errors without any directional considerations:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - x_i| \qquad (15)$$

where n is the total number of data points, y_i is the prediction value, and x_i is the true value.

- **Root Mean Squared Error** (RMSE): Same as MAE but with the caveat of square-rooting the average of the squared errors which translates into enhanced sensitivity for larger errors:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (|y_i - x_i|)^2} \qquad (16)$$

IV. Results

TABLE III
TEST METRICS

Dataset	Metric	Algorithm			
		SoRec	SoReg	DeepSoR	GC-MC
Ciao	MAE	0.81	0.82	0.85	0.87
	RMSE	1.04	1.06	1.03	1.02
Epinions	MAE	1.09	1.07	0.89	1.01
	RMSE	1.14	1.17	1.09	1.07

TABLE IV
VARIANT METRICS

Dataset	Metric	Algorithm		
		X_a	X_b	X_c
Ciao	MAE	0.97	0.79	0.88
	RMSE	1.00	0.98	1.01
Epinions	MAE	1.01	1.04	1.02
	RMSE	0.99	1.02	0.87

V. Discussion

The aim of this research was to develop a recommendation system using graph neural networks that would incorporate embedded social information about users into the predicted rating of an item. The testing for the research algorithms was completed on the Google `Colaboratory` online workspace with standard GPU allocations, and the results are analysed below.

TABLE V
WEIGHTED AVERAGE OF MAE AND RMSE

Algorithm	Dataset	
	Ciao	Epinions
SoRec	0.925	1.115
SoReg	0.950	1.120
DeepSoR	0.940	0.990
GC-MC	0.945	1.040
X_a	0.985	1.000
X_b	0.885	1.000
X_c	0.945	0.945

A. Analysis

Table III and IV display the MAE and RMSE for the test algorithms as compared to the research algorithm variants. Table V highlights the weighted average of both the metrics for all the algorithms, where the three deep learning methodologies (DeepSoR, GC-MC, and variants) outperform the SoRec and SoReg models in several metrics, thereby highlighting the relevancy of neural networks in social recommendation tasks.

The test models' results displayed in Table III show that the GC-MC framework outperforms the other architectures, yet again emphasising the importance of GNNs in representation learning, as the model is capable of handling fluid topological information i.e., nodes forming or changing and/or edges changing direction or associated connections.

Table IV shows the superiority of using the complete framework, X_a, with the user-user and user-item latent factors both included. X_c scores the second best among the three variations highlighting the importance of using social network information over user-item interactions to develop rating predictions.

Overall, the final weighted results from Table V highlight the X_a and X_b variants as the most precise within this research with a percentage difference of 4.14% and 5.66% for the Ciao and Epinions datasets respectively. This further validates the need to incorporate social information i.e., user-user interaction, to develop precise recommendation systems.

B. Further work

This research can be built upon by considering different types of datasets that contain more than social network information i.e., other preferences. This will require changing the architecture of the models to be capable of working with dynamic graph structures with multiple user-item associated edges.

Recently, combining GNNs and recurrent neural networks using a diffusion-convolutional process has shown

promise in the accuracy of the score matrix whilst being computationally efficient [26]. This technique could be employed to the models from this research, as during training, the GPU did time out several times due to the computational load.

VI. Conclusion

The research project is a clear indicator of the importance of using item interaction information augmented with social network information to increase a recommendation network's accuracy. The original framework was split into three neural network variants to test the effect of removing either information source, where the results have been presented and examined to elucidate representative baselines.

References

[1] Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C. and Sun, M., 2020. Graph neural networks: A review of methods and applications. AI Open, 1, pp.57-81.

[2] Cai, D., Hu, J., Qian, S., Fang, Q., Zhao, Q. and Xu, C., 2021. GRecX: An Efficient and Unified Benchmark for GNN-based Recommendation. arXiv preprint arXiv:2111.10342.

[3] Guo, Z. and Wang, H., 2020. A deep graph neural network-based mechanism for social recommendations. IEEE Transactions on Industrial Informatics, 17(4), pp.2776-2783.

[4] Feng X., Liu Z., Wu W., and Zuo W., 2022. Social recommendation via deep neural network-based multi-task learning. Expert Systems with Applications. 10.1016/j.eswa.2022.117755. 206. (117755). Online publication date: 1-Nov-2022.

[5] Fan, W., Ma, Y., Li, Q., He, Y., Zhao, E., Tang, J. and Yin, D., 2019, May. Graph neural networks for social recommendation. In The world wide web conference (pp. 417-426).

[6] Wu, S., Sun, F., Zhang, W., Xie, X. and Cui, B., 2020. Graph neural networks in recommender systems: a survey. ACM Computing Surveys (CSUR).

[7] Fout, A., Byrd, J., Shariat, B. and Ben-Hur, A., 2017. Protein interface prediction using graph convolutional networks. Advances in neural information processing systems, 30.

[8] Marx, V., 2022. Method of the year: Protein structure prediction. Nature methods, 19(1), pp.5-10.

[9] Battaglia, P., Pascanu, R., Lai, M. and Jimenez Rezende, D., 2016. Interaction networks for learning about objects, relations and physics. Advances in neural information processing systems, 29.

[10] Bi, Z., Zhang, T., Zhou, P. and Li, Y., 2020. Knowledge transfer for out-of-knowledge-base entities: Improving graph-neural-network-based embedding using convolutional layers. IEEE Access, 8, pp.159039-159049.

[11] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C. and Philip, S.Y., 2020. A comprehensive survey on graph neural networks. IEEE transactions on neural networks and learning systems, 32(1), pp.4-24.

[12] Hamilton, W., Ying, Z. and Leskovec, J., 2017. Inductive representation learning on large graphs. Advances in neural information processing systems, 30.

[13] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P. and Bengio, Y., 2017. Graph attention networks. arXiv preprint arXiv:1710.10903.

[14] Li, Y., Tarlow, D., Brockschmidt, M. and Zemel, R., 2015. Gated graph sequence neural networks. arXiv preprint arXiv:1511.05493.

[15] Guo, G., Zhang, J. and Yorke-Smith, N., 2015, February. Trustsvd: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings. In Proceedings of the AAAI conference on artificial intelligence (Vol. 29, No. 1).

[16] Jamali, M. and Ester, M., 2010, September. A matrix factorization technique with trust propagation for recommendation in social networks. In Proceedings of the fourth ACM conference on Recommender systems (pp. 135-142).

[17] Song, C., Wang, B., Jiang, Q., Zhang, Y., He, R. and Hou, Y., 2021, July. Social recommendation with implicit social influence. In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 1788-1792).

[18] Marsden, P.V. and Friedkin, N.E., 1993. Network studies of social influence. Sociological Methods & Research, 22(1), pp.127-151.

[19] McPherson, M., Smith-Lovin, L. and Cook, J.M., 2001. Birds of a feather: Homophily in social networks. Annual review of sociology, pp.415-444.

[20] Tieleman, T., and Hinton, G.,. 2012. Lecture 6.5-RMSProp. Neural Networks for Machine Learning [Coursera online course]. University of Toronto, Technical Report.

[21] Ma, H., Yang, H., Lyu, M.R. and King, I., 2008, October. SoRec: social recommendation using probabilistic matrix factorisation. In Proceedings of the 17th ACM conference on Information and knowledge management (pp. 931-940).

[22] Ma, H., Zhou, D., Liu, C., Lyu, M.R. and King, I., 2011, February. Recommender systems with social regularization. In Proceedings of the fourth ACM international conference on Web search and data mining (pp. 287-296).

[23] Fan, W., Li, Q. and Cheng, M., 2018, April. Deep modelling of social relations for recommendation. In

Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 32, No. 1).

[24] Berg, R.V.D., Kipf, T.N. and Welling, M., 2017. Graph convolutional matrix completion. arXiv preprint arXiv:1706.02263.

[25] Wang, S., Tang, J., Wang, Y. and Liu, H., 2018. Exploring hierarchical structures for recommender systems. IEEE Transactions on Knowledge and Data Engineering, 30(6), pp.1022-1035.

[26] Monti, F., Bronstein, M. and Bresson, X., 2017. Geometric matrix completion with recurrent multi-graph neural networks. Advances in neural information processing systems, 30.

APPENDIX A
CODE

It is preferred that the reader accesses the code on the `NBViewer` extension for rendering and sharing Jupyter Notebooks online: **Link**

The code and associated utilities have been made open-source on the author's GitHub profile ⊙: **Link**