Recommender System for Global Terrorist Database Based on Deep Learning

Rafah Shihab Alhamdani, Mohammed Najm Abdullah, and Ismael Abdul Sattar

***Abstract*—Recommender system represents an effective key solution to overcome information overloaded due to huge in volume, multi levels and autonomous of online information. Recently, deep learning gained significant attention through the revolutionary role advanced in many field likes recognition the speech, analyzing images, as well as natural language processing (NLP). Mixing deep learning context mechanisms with recommender system has been gaining momentum because of its wise performances and valid high quality Recommendations. Compared with known traditional recommendation techniques, one of the deep learning main goals is achieve better understanding of customer demands, the characteristics of an items as well as possible interactions between them.**

The paper aims to provide a general review of most recent research works on deep learning (DL) formed on recommender systems lead to fostering change of recommender system research. Deep learning with recommendation models can be formed a taxonomy levels that attract many researchers in a various fields. Like this paper focus on using recommender system to detect how terrorist are spreading online propaganda using various forms of social media working with global terrorist database.

***Index Terms*—Deep learning, recommender system, terrorist attacks, social networking.**

I. INTRODUCTION

The high availability side by side with big growth of information online making users confuse in various aspects. Finding a service or product fall into user satisfaction from huge scope of valid possible choices recommender will play the role of guiding. Recommender used in many aspects as more essential player for filtering, simplifies decision making, and several information access system for financial purpose [1].

Such system used in several fields not only deals with user interests but also for customer-items past demand and some other extra information like spatial and temporal data as well. Categorization of recommender based on challenges into three types as shown in the Fig. 1. Each type has its set of problems to work on [2]:

All of the above models have weaknesses in working with:

1) Data sparsity
2) Could start problem
3) Balancing recommendation quality in term of different evaluation metric.

In several applications field like speech recognition and computer vision as well deep learning achieve big successes. There is a big race in industry and academic to apply deep learning for solving huge problems by exploiting deep learning capabilities side by side providing useful art results [3]. Most recently, several companies utilized deep learning in recommender systems due to the big change in recommender that deep learning make to chive high levels of user satisfaction.

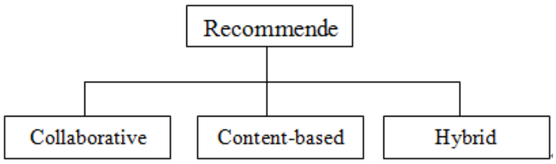

Fig. 1. Recommender model types.

II. TERRORISM

Terrorism is one of complex concept with high uncertainties in user strategy. Such uncertain nature of the terrorism is a main challenge for designing terrorism policy. Many governments how fighting against terrorism makes big systems to deal with the big data available on social media and telecommunications to capture the intentions of the terrorist, but as consequences become very carful in the use of these environments to plan and prepare attacks.

Crime and tendency to violence continue to remain severe threats to the entire world with highly complex criminal activities [17]. Crime is not limited to the streets now because the use of the Internet causes a tendency to violence and more sophisticated behaviors in the modern age. Know day's terrorism besides the advertisements and the seducing people they do over the social media they achieve their agenda by attacking strategic goals secretly that is shows it is tot wise at all to completely depend on the social media to judge and identify the terrorism behavior thus, The work with TDG to train our learning model that will provide a rich data based on the long history of terror events with various attributes attached with this databases.

III. BENEFIT OF USING DEEP LEARNING WITH RECOMMENDER SYSTEM

Many companies like Google and Yahoo start using deep learning in recommender systems to achieve high Revenue through gaining user satisfaction or optimize their decision-making. For example, more than eighties percent of movies

Manuscript received May 25, 2018; revised August 3, 2018.
Rafah Shihab Alhamdani and Ismael Abdul Sattar are with the Informatics Institute for Postgraduate Studies, Iraqi Commission for Computers & Informatics, Baghdad, Iraq (e-mail: ismaelabdul@uomustansiriyah.edu.iq, asmaell_sj@yahoo.com).
Mohammed Najm Abdullah is with the Computer Engineering Dept., University of Technology, Baghdad, Iraq.

doi: 10.18178/ijmlc.2018.8.6.747

watched over Netflix provided by recommendations [4], while sixties percent of video clicks provided through home page recommended in YouTube [5]. Recently, several companies exploit deep learning for achieving further utilization of level quality recommendation [3], [6], and [7]. Deep learning becomes a useful for many reasons some of them are:

1) Work with capturing the nonlinear as well as nontrivial customer demands relationships.
2) Summarize complex data representation as abstraction at higher levels.
3) Discovering the connections in data in various format such as text or visual data.

Will be show in Fig. 2 the importance of the recommender system in term of, load of annual publications, publications place, used datasets, and metrics of evaluation.

From the various field that deep learning can cover with achieving high success through utilizing it with recommender system it's become in need to focus on which area to work with as well as select a useful metric in such area of applications by measuring the strength and weaknesses as well.

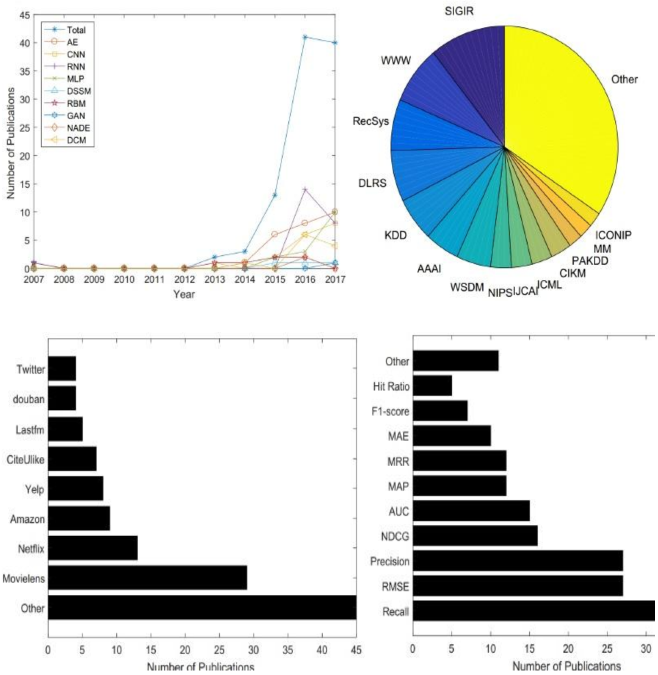

Fig. 2. (I) load of annual publications (II) publications place (III) used datasets (VI) and metrics of evaluation.

IV. Deep Learning Techniques

While a machine learning a big field that's contains deep learning as sub part of its. Deep learning can reach deeply in multi levels forming representation of data to learn. Which is can be used for various kind of learning tasks guided or unguided as well [8]. This section shows deep learning techniques which are related to recommender works.

1) "Perceptron of multilayer (M L P)": kind of neural network with feedforward supported by one or more hidden layer, here such network can work not only with arbitrary function but also exceed binary classifier representation.
2) "Auto-encoder (A E)": is a model trying to build input data in final layer (output) which is a kind of unguided model. It has a drawback in responsible to used salient representation of features. There are various kinds of autoencoder like sparse, denoising, and contractive autoencoder [9], [10].
3) "Convolutional Neural Network [10]": A neural network contains convolution layers as well as operations of pooling. It has the ability to captures general and detailed features as well, that is increasing the efficiency level side by side with accuracy achieving well with grid data topologies.
4) "Recurrent Neural Network [10]:" a well performing neural network for sequential data, such types of networks contains not only loops but also memories achieving computations. This kind of networks works well to overcome the problem of vanishing gradient.
5) "Deep Semantic Similarity Model": is kind of neural network used for semantic representations learning of entities based on the similarity measurements. Which is also describe as deep semantic and structured network [11].
6) "Restricted Boltzmann Machine": is neural network build from visible and hidden layer. This can be formed for deep neural networks. There is no communication layer in visible or hidden layer that is why restricted.
7) "Neural Autoregressive Distribution Estimation" [12], [13]: is unguided neural network model constructed for feedforward and auto-regression model neural network. Its work well as estimator over modeling information, and data sparsity and densities.
8) Generative Adversarial Network [14]: is a model of neural network which is contains not only a discriminator but also generator both of the nets competing with each other and simultaneously trained.

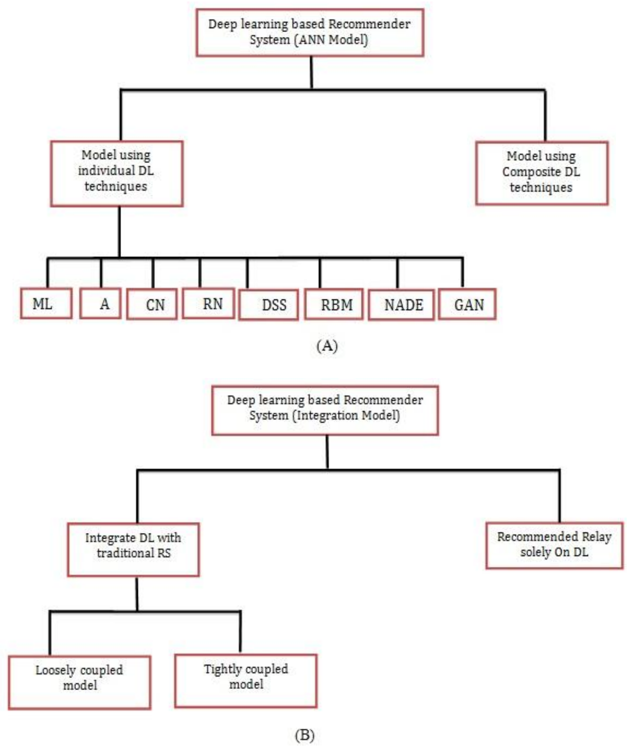

Fig. 3 Division of deep learning (DL) over recommender system (RS). (A) Models based on neural network (B) models based on Integration.

The existing works can be divided into two context schemes, that's give better understanding for the integration and neural network model as well. Fig. 3 shows the divisions for such classification.

V. NEURAL ATTENTION BASED RECOMMENDATION

Human visual attention play a big role for motivates attention mechanism. For example, focussing on specific parts of visual input by human it's sufficient to understand or recognize them. Filtering out uninformative features from raw data input can be done using attention mechanism and reduce the side effects of nosy data. While it is an intuitive but effective technique has considerable attention over the recent years across areas such as computer vision [3], natural language processing [18], [19], and speech recognition [20], [21]. Neural network can't only be used in conjunction with MLP, CNNs and RNN, but also some tasks independently addressed [22]. Integrating attention mechanism into RNNs enables the RNNs to process long and noisy inputs [21]. Although LSTM can solve the long memory problem theoretically, it is still problematic when dealing with long-range dependencies. Attention mechanism provides a better solution and helps the network to better memorize inputs. Attention-based CNNs are capable of capturing the most informative elements of the inputs [23]. By applying attention mechanism to recommender system, one could leverage attention mechanism to filter out uninformative content and select the most representative items [14] while providing good interpretability. Although neural attention mechanism is not exactly a standalone deep neural technique.

Attention model learns to a end to the input with attention scores. Calculating the attention scores lives at the heart of neural attention models. Based on the way for calculating the attention scores, we classify the neural attention models into (1) standard vanilla attention and (2) co-attention. Vanilla attention utilizes a parameterized context vector to learn to attend while co-attention is concerned with learning attention weights from two-sequences. Self-attention is a special case of co-attention. Recent works [14], [23], [24] demonstrate the capability of attention mechanism in enhancing recommendation performance.

Recommendation with Vanilla Attention, Chen *et al.* [14] proposed an attentive collaborative filtering model by introducing a two-level attention mechanism to latent factor model. It consists of item-level and component-level attention. The item-level attention is used to select the most representative items to characterize users. The component-level attention aims to capture the most informative features from multimedia auxiliary information for each user. Two attention networks are used to model user long-term and short-term interests. Introducing attention mechanism to RNNs could significantly improve their performance. Li *et al.* [25] proposed such an attention-based LSTM model for hashtag recommendation. Is work takes the advantages of both RNNs and attention mechanism to capture the sequential property and recognize the informative words from microblog posts. Vanilla attention can also work in conjunction with CNNs for recommender tasks. Gong *et al.* [24] proposed an attention based CNNs system for hashtag recommendation in microblog. It treats hashtag recommendation as a multi-label classification problem. The proposed model consists of a global channel and a local attention channel. The global channel is made up of convolution filters and max-pooling layers.

All words are encoded in the input of global channel. The local attention channel has an attention layer with given window size and threshold to select informative words (known as trigger words in this work). Hence, only trigger words are at play in the subsequent layers.

Recommendation with Co-Attention Zhang *et al.* [27] proposed a combined model, AttRec, which improves the sequential recommendation performance by capitalizing the strength of both self-attention and metric learning. It uses self-attention to learn user short-term intents from her recent interactions and takes the advantages of metric learning to learn more expressive user and item embeddings.

VI. DEEP HYBRID MODELS FOR RECOMMENDATION

With the good flexibility of deep neural networks, many neural building blocks can be integrated to formalize more powerful and expressive models. Despite the abundant possible ways of combination, the suggestion of the hybrid model should be reasonably and carefully designed for the specific tasks. Here, we summarize the existing models that have been proven to be effective in some application fields.

1) **CNNs and Auto-encoder.** Collaborative Knowledge Based Embedding (CKE) [28] combines CNNs with auto-encoder for images feature extraction. CKE can be viewed as a further step of CDL. CDL only considers item text information (e.g. abstracts of articles and plots of movies), while CKE leverages structural content, textual content and visual content with different embedding techniques. Structural information includes the attributes of items and the relationships among items and users. CKE adopts the TransR [29], a heterogeneous network embedding method, for interpreting structural information. Similarly, CKE employs SDAE to learn feature representations from textual information. As for visual information, CKE adopts a stacked convolution auto-encoders (SCAE). SCAE makes efficient use of convolution by replacing the fully-connected layers of SDAE with convolution layers. The recommendation process is done in a probabilistic form similar to CDL.
2) **CNNs and RNNs.** Lee *et al.* [30] proposed a deep hybrid model with RNNs and CNNs for quotes recommendation. Quote recommendation is viewed as a task of generating a ranked list of quotes given the query texts or dialogues (each dialogue contains a sequence of tweets). It applies CNNs to learn significant local semantics from tweets and maps them to distributional vectors. These distributional vectors are further processed by LSTM to compute the relevance of target quotes to the given tweet dialogues. Zhang *et al.* [31] proposed a CNNs and RNNs based hybrid model for hashtag recommendation. Given a tweet with corresponding images, the authors utilized CNNs to extract features from images and LSTM to

learn text features from tweets. Meanwhile, the authors proposed a co-attention mechanism to model the correlation influences and balance the contribution of texts and images. Ebsesu *et al.* [32] presented a neural citation network, which integrates CNNs with RNNs in an encoder-decoder framework for citation recommendation. In this model, CNNs act as the encoder that captures the long-term dependencies from citation context. e RNNs work as a decoder which learns the probability of a word in the cited paper's title given all previous words together with representations a ained by CNNs.

Chen *et al.* [17] proposed an intergrated framework with CNNs and RNNs for personalized key frame (in videos) recommendation, in which CNNs are used to learn feature representations from key frame images and RNNs are used to process the textual features.

1) RNNs and Autoencoder. The collaborative deep learning model mentioned is lack of robustness and incapable of modelling the sequences of text information. Wang *et al.* [33] further exploited integrating RNNs and denoising auto-encoder to overcome this limitation. The authors first designed a generalization of RNNs named robust recurrent network. Based on the robust recurrent network, the authors proposed the hierarchical. Bayesian recommendation model called CRAE. CRAE also consists of encoding and decoding parts, but it replaces feed forward neural layers with RNNs, which enables CRAE to capture the sequential information of item content information. Furthermore, the authors designed a wildcard denoising and a beta-pooling technique to prevent the model from over fitting.
2) **RNNs with DRL.** Wang *et al.* [34] proposed combining supervised deep reinforcement learning with RNNs for treatment recommendation. The framework can learn the prescription policy from the indicator signal and evaluation signal. Experiments demonstrate that this system could infer and discover the optimal treatments automatically. We believe that this valuable topic and benefits the social good.

VII. Global Terrorism Database

The GTD one of the interesting open-source database, which contains information about terrorist events from 1970 to 2017 including more than 170000 cases with the systematic data on the domestic beside international terrorist incidents that have occurred during this time period.

For each GTD incident, information is available on the date and location of the incident, the weapons used and nature of the target, the number of casualties, and-when identifiable-the group or individual responsible.

From the GTD characteristics is currently the most comprehensive unclassified database on terrorist events in the world. Also Includes information on more than 83,000 bombings, 18,000 assassinations, and 11,000 kidnappings since 1970 and information on at least 45 variables for each case, with more recent incidents including information on more than 120 variables, Over 4,000,000 news articles and 25,000 news sources were reviewed to collect incident data from 1998 to 2016 alone. Thus GTD is rich area for doing research in this field.

Modeling approaches, according to Guo, are limited to formulate and test a valid hypothesis in their ability. They build on rigorous statistical or mathematical models, but provide limited understanding of the causes, development, and diffusion of terrorism activities. Moreover, terrorism data is often incomplete or inaccurate and only represents the outcome, not the process [15].

Ziemkiewicz developed an interconnected visual analysis tool that provides three interconnected views of the Global Terrorism Database (GTD) to support the investigative process [16]. Fig. 4 below shows the views from this tool used by investigators to explore the data in an abstract way by examining correlations across multiple dimensions.

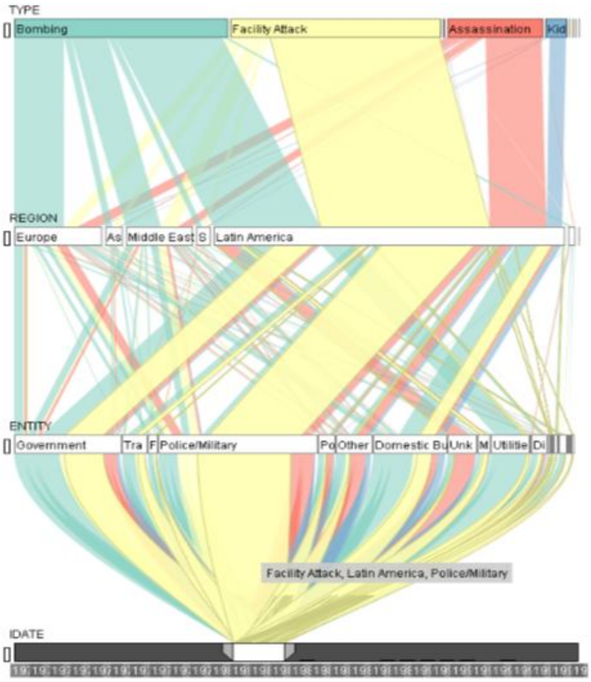

Fig. 4. Visual analysis of correlations across data dimensions [16].

The ribbons between categorical dimensions in Fig. 2 show the proportion of cases from one category that fall under a category from another dimension.

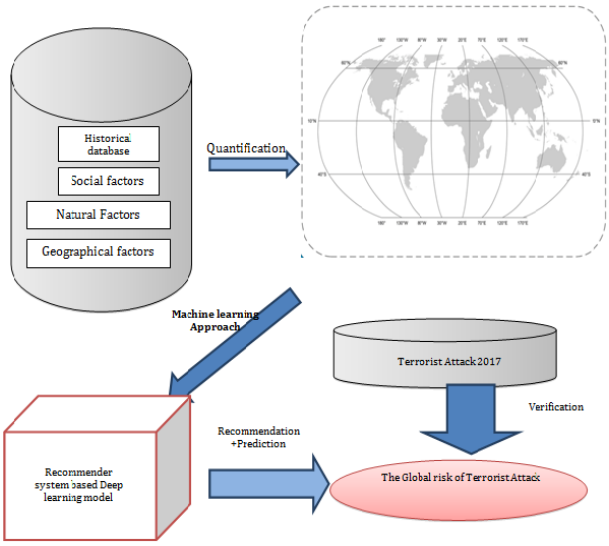

Fig. 5. Workflow of the proposed system.

VIII. Proposed System

Proposed system work on the global terrorist database (GTD), with taking in our consideration social, natural factors and geographical factors as well. Most of the data used for training the recommender system based on deep learning while the rest of the data used for testing and verification the output measurements of the system as shown in the Fig. 5.

IX. Result

Using recommender system over the Global terrorist database will help many security agencies for predicting and determining the risks that could rise for the future due to terrorist behavior, which can be determine by finding the relationships in historical attacks. Some of the visualize result shown in the Fig. 5. To explain the semantic relation, in one specified terrorist sector.

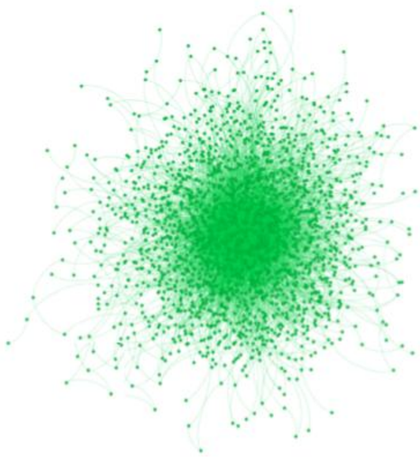

Fig. 6. One terrorist sector in global terrorist database GTD.

X. Conclusion

Most the terrorist attacks happen for the various factors and motivation as well, thus even with using powerful deep learning model over the global terrorist database this not sufficient to determine what are the next attacks instead rising suspicions. To achieve such goal we have to study the behavior of the terrorist over social networking. And assign some values for such kind of decision. That will be powerful social alarm system by fusion multi decisions that contribute in the same issues.

References

[1] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734–749, 2005.

[2] T. Alashkar, S. Y. Jiang, S. Y. Wang, and Y. Fu, "Examples-rules guided deep neural network for makeup recommendation," *AAAI*, pp. 941–947, 2017.

[3] P. Covington, J. Adams, and E. Sargin, "Deep neural networks for YouTube recommendations," in *Proc. the 10th ACM Conference on Recommender Systems*, 2016, pp. 191–198.

[4] C. A. Gomez-Uribe and N. Hunt, "E net ix recommender system: Algorithms, business value, and innovation," *ACM Transactions on Management Information Systems*, vol. 6, no. 4, p. 13, 2016.

[5] J. Davidson, B. Liebald *et al.*, 2010, "The YouTube video recommendation system," in *Proc. the Fourth ACM Conference on Recommender Systems*, New York, NY, USA, pp. 293–296.

[6] H.-T. Cheng, L. Koc *et al.*, "Wide & deep learning for recommender systems," in *Proc. the 1st Workshop on Deep Learning for Recommender Systems*, 2016, pp. 7–10.

[7] S. Okura, Y. Tagami, S. Ono, and A. Tajima, "Embedding-based news recommendation for millions of users," in *Proc. the 23th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017.

[8] L. Deng, D. Yu *et al*., 2014, "Deep learning: Method sand applications," *Foundations and Trends in Signal Processing*, vol. 7, 2014, pp. 197–387.

[9] M. M. Chen, Z. X. Xu, K. Weinberger, and F. Sha, "Marginalized denoisingautoen coders for domain adaptation," arXiv preprint arXiv:1206.4683, 2012.

[10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.

[11] P.-S. Huang, X. D. He, J. F. Gao *et al*., "Learning deep structured semantic models for web search using click through data," in *Proc. the 22nd ACM International Conference on Conference on Information & Knowledge Management*, pp. 2333–2338, 2013.

[12] H. Larochelle and I. Murray, 2011, "The neural autoregressive distribution estimator," in *Proc. the Fourteenth International Conference on Artificial Intelligence and Statistics*, pp. 9–37.

[13] B. Uria, K. Gregor *et al*., "Neural autoregressive distribution estimation," *Journal of Machine Learning Research*, vol. 17, no. 205, 2016, pp. 1–37.

[14] M. Mirza, B. Xu *et al*., "Generative adversarial nets," *Advances in Neural Information Processing Systems*, pp. 2672–2680, 2014.

[15] D. Guo, K. Liao, and M. Morgan, "Visualizing patterns in a global terrorism incident database," Department of Geography, University of South Carolina, Columbia, 2006.

[16] C. Ziemkiewicz *et al.*, "Global terrorism visualization," Southeast Visualization and Analytics Center (SRVAC), University of North Carolina at Charlotte, 2008.

[17] S. Tutun and M. T. Khasawneh, *New Framework That Uses Patterns and Relations to Understand Terrorist Behaviors*, 2017.

[18] M.-T. Luong, H. Pham, and C. D. Manning, 2015, "E effective approaches to attention-based neural machine translation," arXiv preprint arXiv:1508.04025, 2015.

[19] J. McAuley, C. Targe, Q. F. Shi, and A. V. D. Hengel, "Image-based recommendations on styles and substitutes," *SIGIR*, pp. 43–52, 2015.

[20] J. Chorowski, D. Bahdanau, K. H. Cho, and Y. S. Bengio, "End-to-end continuous speech recognition using attention-based recurrent NN: First results," arXiv preprint arXiv:1412.1602, 2014.

[21] J. K. Chorowski, D. Bahdanau, D. Serdyuk *et al.*, "Attention-based models for speech recognition," *Advances in Neural Information Processing Systems*, 2015, pp. 577–585.

[22] A. Vaswani, N. Shazeer, N. Parmar *et al.*, "Attention is all you need," *Advances in Neural Information Processing Systems*, pp. 5998–6008, 2017.

[23] S. Y. Seo, J. Huang, H. Yang, and Y. Liu, "Representation Learning of Users and Items for Review Rating Prediction Using Attention-based Convolutional Neural Network," *MLRec*, 2017.

[24] Y. Y. Gong and Q. Zhang, "Hashtag Recommendation Using Attention-Based Convolutional Neural Network," *IJCAI*, pp. 2782–2788, 2016.

[25] Y. Li, T. Liu, J. Jiang, and L. Zhang, "Hashtag recommendation with topical attention-based LSTM," *COLING*, 2016.

[26] S. Zhang, Y. Tay, L. Yao, and A. X. Sun, "Next item recommendation with self-attention," arXiv preprint arXiv: 1808.06414, 2018.

[27] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," in *Proc. the 25th International Conference on Machine Learning*, 2008, pp. 160–167.

[28] Y. F. Zhang, Q. Y. Ai, X. Chen, and W. B. Croft, "Joint representation learning for top-n recommendation with heterogeneous information sources," *CIKM*, pp. 1449–1458, 2017.

[29] Y. K. Lin, Z. Y. Liu *et al.*, "Learning entity and relation embeddings for knowledge graph completion," *AAAI*, pp. 2181–2187, 2015.

[30] H. Lee, Y. C. Ahn, H. J. Lee *et al.*, "Quote recommendation in dialogue using deep neural network," *SIGIR*, pp. 957–960, 2016.

[31] Q. Zhang, J. W. Wang, H. R. Huang, X. J. Huang, and Y. Y. Gong, "Hashtag Recommendation for Multimodal Microblog Using Co-Attention Network," *IJCAI*.

[32] T. Ebesu and Y. Fang *Neural Citation Network for Context-Aware Citation Recommendation*, 2017.

[33] H. Wang, S.-X. Jian, and D.-Y. Yeung, "Collaborative recurrent autoencoder: Recommend while learning to fill in the blanks," NIPS, pp. 415–423, 2016.

[34] L. Wang, W. Zhang, X. F. He, and H. Y. Zha, "Supervised reinforcement learning with recurrent neural network for dynamic treatment recommendation," *SIGKDD*, pp. 2447–2456, 2018.

Rafah Shihab Alhamdani is the dean of the Informatics Institute for postgraduate studies, Iraqi Commission for Computers & INFORMATICS, Baghdad, Iraq. She has published more than 40 research papers in national or international journals and conferences. She also published 8 books in various fields.

She got the Ph.D degree in economic from Baghdad University in 1997; M.S. degree in agricultural economics from Baghdad University in 1986; B.Sc. degree in agricultural economics from Baghdad University in 1975.

Mohammed Najm Abdullah received the B.Sc. degree in 1983 in electrical engineering from Collage of Engineering, University of Baghdad, Baghdad, Iraq. He is received the M. Sc. degree in electronic and communication engineering from the same collage of Engineering/University of Baghdad in 1989. He is received the Ph.D degree in 2002 in electronic and communication engineering from University of Technology, Baghdad, Iraq. Now, He is an assistant professor in the Department of Computer Engineering in the University of Technology, Baghdad, Iraq. His area of interests is air-borne computer, DSP software and hardware, and PC-interfacing, e-learning, information systems management, wireless sensor networks. He has published 42 research papers in national or international journals and conferences with 13 books.

Ismael Abdul Sattar Jabbar is a Ph.D student in Informatics Institute for postgraduate studies. He gets M.Sc. degree in computer science from Delhi University, Faculty of Mathematical Sciences in 2012; B.Sc. degree in computer science in 2006 from Al-Mustansiriyah University, Collage of Science. He has published more than 15 research papers in national or international journals and conferences with 2 books in information security.