Seq2Seq Neural Architecture for Recommending Short Text Conversations

Anonymous Author(s)
Affiliation
Address
email

Abstract

Even though social media provides massive amounts of data; users can digest only a small set of it. For this reason, recommending relevant content becomes an important task to avoid information overload. This paper explores content-based recommendation models for social media, specifically, for recommending conversations to users that might be interested in participating in the discussion. Previous works recommend content to users based on latent factor models and collaborative filtering to map explicitly from text to factors to enable recommendations that can generalize for new content. This work proposes a neural architecture that encodes the users' profiles history and the conversations' context to learn users' preferences. We employ recurrent neural networks (LSTM), trained end-to-end on microblogging conversational corpus. The empirical results show that neural learning architectures provide higher recall compared to baseline methods for modeling unstructured and noisy short text conversations on Twitter.

1 Introduction

In the field of recommendation systems, previous works addressed several types of recommendation tasks, including hashtags, mentions, news, points-of-interest, profile classification, retweets, tweets, URLs, and whom to follow [7]. In addition to the interaction generated within OSNs, users can consume news or content available, often faster than traditional media [2, 1]. The recommendation of the content is an essential task for companies and organizations that are looking at to reach users, or even individuals looking to attract the attention of the crowd to help in different tasks for instance regarding crisis management [5]. This paper focuses on the latter, recommending conversations to users that may elicit interactions from the crowds.

Previous works focus on analyzing individual tweets for recommendation tasks. In this work, we focus on recommending users to join a set of conversations on social media. Prior work has proposed the use of collaborative filtering (CF), as well in combination with topic modeling [9]. In this paper, we propose a Seq2Seq neural architecture for the task of recommending users to join conversations based on their preferences and conversation context.

2 Model and Experimental Setup

We propose an encoder-decoder neural architecture that uses two siamese sequential networks to encode the conversation context as well the user history (participation in previous conversations) using Recurrent Neural Networks (RNNs). The models presented builds upon on the approaches used for the task of dialog response selection on chatbots [4].

Figure 1 depicts the proposed model based on an encoder-decoder architecture. The proposed models learn to users' preferences and the probability to join a conversation based on previous tweets of their

Submitted to 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). Do not distribute.

timeline, where c_i, p_i are word vectors of the conversation's context and the user' profile. The values c, p correspond to the last hidden states from the sequence models.

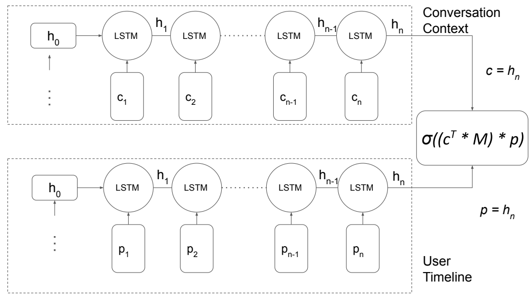


Figure 1: Diagram of the Seq2Seq model.

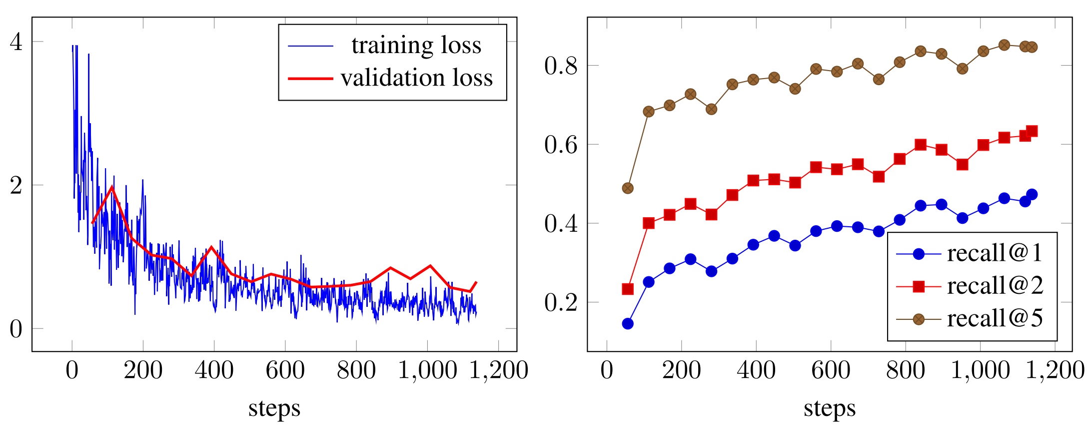
At each step, the siamese RNNs update the hidden states, and the final hidden state represents a *summary* of the input context and profile. Then, the model calculates the probability of having a valid pair of final hidden states from both RNNs, as follows:

$$p(\text{flag} = 1|c, p, M) = \sigma(c^T M p + b), \tag{1}$$

where the bias b and the matrix $M \in \mathbb{R}^{d \times d}$ are parameters of the model. For training, given an input triplet (conversation context, user profile, flag), the model generates a candidate user profile (c') representation as to the product $c' = M * p$ and then measures the similarity to the actual user profile using the dot product. Then, we use the sigmoid function converts it to a probability. The training of the model tries to minimize the cross-entropy [8] of all triples. For a ground truth user profile to a conversation context, the $flag = 1$, and we generate a negative instance (i.e., a conversation context where the user did not participate) with $flag = 0$.

In the validation and test set, for each positive instance (the correct user profile that joined the conversation), we select nine users' profiles that did not join the conversation as negative instances or distractors.

For preliminary experiments, we use the TREC dataset [6] that contains conversations of diverse topics based on the tweets released by microblog track. Figure 2a shows the training and evaluation loss of the best model using LSTM networks. Also, Figure 2b shows the performance of the model on the validation set, using $recall@k$ metric with different values for K. The experimental results show that the proposed neural architecture based on LSTM provides higher accuracy compared to naive models such as TF-IDF and collaborative filtering approaches in previous works.



(a) Training and validation loss of the model.

(b) Evaluation of the model Recall@K.

References

[1] I. Guy. Social recommender systems. In *Recommender Systems Handbook*, pages 511–543. Springer, 2015.

[2] H. Kwak, C. Lee, H. Park, and S. Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pages 591–600. AcM, 2010.

[3] C. Li, Y. Lu, Q. Mei, D. Wang, and S. Pandey. Click-through prediction for advertising in twitter timeline. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1959–1968. ACM, 2015.

[4] R. Lowe, N. Pow, I. Serban, and J. Pineau. The ubuntu dialogue corpus: A large dataset for research in unstructured multi-turn dialogue systems. *arXiv preprint arXiv:1506.08909*, 2015.

[5] L. Palen and A. L. Hughes. Social media in disaster communication. In *Handbook of Disaster Research*, pages 497–518. Springer, 2018.

[6] I. Soboroff, I. Ounis, C. Macdonald, and J. J. Lin. Overview of the trec-2012 microblog track. Citeseer.

[7] L. Terán, A. O. Mensah, and A. Estorelli. A literature review for recommender systems techniques used in microblogs. *Expert Systems with Applications*, 103:63–73, 2018.

[8] L. Yu, K. M. Hermann, P. Blunsom, and S. Pulman. Deep learning for answer sentence selection. *arXiv preprint arXiv:1412.1632*, 2014.

[9] X. Zeng, J. Li, L. Wang, N. Beauchamp, S. Shugars, and K.-F. Wong. Microblog conversation recommendation via joint modeling of topics and discourse. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 375–385, 2018.