Recurrent Neural Network Based Visual Analytics Framework For Social Media

Authors
Department of Computer Science, Virginia Tech
Blacksburg, VA 24060

Abstract

Twitter is an emerging microblog service provider from 2006. Twitter users can add hashtags to label and classify their posts. But, only about 11% tweets contain at least one hashtag. Thus, hashtag prediction for any tweet texts is a necessary task for tweets analysis.

From 2012, neural network and deep learning models have been successfully applied into different areas of computer vision, speech recognition and natural language processing (NLP). Recurrent Neural Network (RNN) is one type of neural network model which utilizes the sequential information during training process. RNN models can classify the data in high performance. But they are black-boxes. It is challenge to explore what information a RNN model learns and stores in its hidden states.

Visual analytics and relevant interaction techniques are potential to assist to understand RNN models. However, how to design efficient and usable visual representations of RNN hidden states with necessary interactions to aid sensemaking is still challenging.

To address this, my thesis work highlights three contributions:

(1) We designed a RNN based visual analytics system for rumor tweets analysis in hashtag prediction, tweets clustering, and interactive visualization.

(2) We proposed an interactive heatmap visual representation for RNN hidden states understanding. Our design supports users to find latent patterns of RNN hidden states and explore them in visual context.

(3) We will conduct user studies to evaluate whether our design support users effectively gain more insights in the sensemaking process of RNN models understanding.

Tweets and Hashtags

Twitter is an emerging microblog service provider from 2006. It can allow users to post, publish, share and communicate by tweets, which is a 140 characters limited short message. Because of this characters limited feature, the languages in tweets is including a lot of abbreviations and emotion icons. Another important feature is that tweets allow users to use "@" for reply to specific another users and "#" for the self-tagging and self-categories.

The hashtags in tweets are the words that are preceded by a hash symbol (#). And there is no space in one hashtag. The hashtags can be used in the beginning, the middle and the end of tweets for the operation of tagging or annotation. In Twitter, if you click the hashtag, it will be linked to the page displaying other tweets that contains the same hashtag on Twitter. Hashtags do not just appear in Twitter. They are also used in Flicker, Pinterest, Instagram, and Facebook.

Social media users add hashtags in their posts for different goals, including identification label (#VT or #Hokie), sentiment label (#love, #like, or #hate), topic label (#hiking), event label (#uselection2016) and etc.. Thus, hashtags provide labels for their texts.

Recurrent Neural Network

Neural networks is a machine learning approach to map the features of data into an abstract and high-dimensional representation. From 2012, neural network and deep learning models have been successfully applied into different areas of computer vision, speech recognition and natural language processing (NLP) (LeCun, Bengio, and Hinton 2015).

Recurrent Neural Network (RNN) is one type of neural network models. RNN model utilizes the sequential information during training process. It processes the same computation for each sequential input, and the output relies on its previous sequential result. In another word, RNN model can be thought as a memory to store the sequential inputs which have been processed earlier. RNN model has been proved as Turing-complete by Siegelmann et al.(Siegelmann and Sontag 1995) in 1995. It means that just like Turing Machines, any algorithm can be encoded via a RNN model with parameters tuning.

In NLP research and application area, RNN model has remarkable achievement recently in language modeling (word embedding (Mikolov et al. 2013), sentence embedding (Kiros et al. 2015)), machine translation (Sutskever, Vinyals, and Le 2014), sentiment analysis (Socher et al. 2013), question answering (Iyyer et al. 2014), and etc..

RNN for Hashtag Prediction

Hashtag can label and classify tweets. It is useful in many scenarios, including tweets search and retrieval (Efron 2010), sentiment analysis (Davidov, Tsur, and Rappoport

2010), and etc.. But, not all the tweets contain hashtag. Based on a survey on a collection of $62,556,331$ tweets, Hong et al. find that there are only about 11% of tweets in their collection contain at least one hashtag (Hong, Convertino, and Chi 2011). So, hashtag prediction for any tweet texts is a necessary task for tweets analysis.

Hashtag prediction is a multi-class classification task to assign one or several hashtags to the corresponding tweet texts. Hashtag prediction has multiple applications in tweets analysis. For example, predicted hashtags of tweets would help users to view tweets via different hashtag categories, especially improve the social media journalists to explore tweets efficiently.

In this thesis, we utilize recurrent neural network (RNN) to train a classification model for hashtag prediction of tweet texts. The model can achieve 51.86% in prediction rate, which is 2x higher than the traditional Bag-Of-Word (BOW) prediction model.

Hidden States Visual Analytics for RNN

After training process, RNN models can classify the data in high performance. However, RNN models are black-boxes. Even model creators cannot interpret why their models achieves in high performance. They also do not have hints about what features the model has learned from the data. Thus, except high performance, model creators do not have evidences from model itself to support their decision making. For example, our RNN hashtag prediction model can achieve 51.86% in prediction rate. But from model itself, we cannot find any evidence to let us understand how the model achieve this high prediction rate.

Under the hood of RNN models, they have tons of parameters and repeatedly compute non-linear activation functions for large number of their hidden states. Due to these factors, how to interpret and understand RNN model is a challenge work and an active research topic in deep learning area. Few researchers tried to interpret RNN models via studying the changes in hidden states over time (Strobelt et al. 2016) (Li et al. 2016), . But they found some interpretable patterns with large noise and interruptions.

Visual analytics and relevant interaction techniques are potential to aid sensemaking process (Pirolli and Card 2005) of RNN models understanding and interpretion. It can support users to find latent patterns of RNN hidden states and reveal them in visual context. However, how to design efficient and usable visual representations of RNN hidden states with necessary interactions to assist sensemaking is still challenging.

Research Questions

In this thesis, we will explore:

- How to utilize the Recurrent Neural Network (RNN) techniques for tweets analysis, such as hashtag prediction, tweets clustering, and interactive visualization?
- How to use interactive visual analytics techniques to understand RNN models and form hypotheses about RNN hidden states patterns?

In order to state them, three key research questions (RQs) are listed as following:

RQ 1: How to build a Recurrent Neural Network (RNN) based tweets visual analytics system?

- How we can build a RNN model to predict the hashtag of tweets?
- How to design a tweets visualization system with the results from our RNN model?
- What is the key design trade off among our RNN model, visualization representations, and interactive strategy?

To answer Research Question 1, firstly, we utilize recurrent neural network (RNN) to train a classification model for hashtag prediction of any texts. Our RNN model can achieve 51.86% in prediction rate, which is 2x higher than the traditional Bag-Of-Word (BOW) prediction model. Secondly, we built a RNN we designed a rumor tweets visualization system with the sentence embeddings from our RNN model. Thirdly, we will conduct a user study to find the key design trade off among our RNN model, visualization representations, and interactive strategy.

RQ 2: How can a visual analytics system be designed to understand RNN models?

- How can we explore the RNN models hidden states and their parameters in an interactive visual way?
- How to visualize the probability of activation and probability of contribution of RNN hidden states?
- How to use visualization build the connections between RNN hidden states and inputs/outputs of models? How to answer the typical question, like what information does a RNN model learn and store in its hidden states?

The design of our interactive heatmap visualization for RNN hidden states and relevant matrix reordering technique can answer these questions.

RQ 3: How effective are our visual analytic system designed to support sensemaking process of RNN models understanding?

- Would our proposed design visualization in RQ2 be able to help users more effectively gain insights about how the black-box RNN model works?
- How does our proposed RNN hidden states visualization impact users understanding and interpretation of RNN model?
- Can users answer what information a RNN model learns and stores in its hidden states via our design?
- Can users find the abstract, sentiment information, concepts of inputs with the hidden states pattern in our visualization design?
- What are the strategies of users to use in our design to find latent patterns and form hypotheses about RNN hidden states?

- Compare to other RNN visualization approaches(Strobelt et al. 2016), what differences can be observed in the users workflow, process, and, insight?
- Can our design be deployed on crowdsourcing platform to scale up the sensemaking process of RNN model understanding?

Our future user studies could answer Research Question 3 and its sub questions.

Significance

Exploring these research questions above would help researchers build tools with fully leverage of RNN models. It would lead them to gain more insight of RNN models understanding and to build more usable visual analytics system with RNN models.

References

Davidov, D.; Tsur, O.; and Rappoport, A. 2010. Enhanced Sentiment Learning Using Twitter Hashtags and Smileys. *Proceedings of the 23rd International Conference on Computational Linguistics: Posters* (August):241–249.

Efron, M. 2010. Hashtag retrieval in a microblogging environment. *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval - SIGIR '10* 787.

Hong, L.; Convertino, G.; and Chi, E. H. 2011. Language Matters in Twitter : A Large Scale Study. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media* (1):518–521.

Iyyer, M.; Boyd-Graber, J.; Claudino, L.; Socher, R.; Daumé III, H.; Daum, H.; and III, H. 2014. A Neural Network for Factoid Question Answering over Paragraphs. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* 633–644.

Kiros, R.; Zhu, Y.; Salakhutdinov, R.; Zemel, R. S.; Torralba, A.; Urtasun, R.; and Fidler, S. 2015. Skip-Thought Vectors. *NIPS* (786):1–11.

LeCun, Y.; Bengio, Y.; and Hinton, G. 2015. Deep learning. *Nature* 521(7553):436–444.

Li, J.; Chen, X.; Hovy, E.; and Jurafsky, D. 2016. Visualizing and Understanding Neural Models in NLP. *Naacl* 1–10.

Mikolov, T.; Chen, K.; Corrado, G.; and Dean, J. 2013. Distributed Representations of Words and Phrases and their Compositionality. *Nips* 1–9.

Pirolli, P., and Card, S. 2005. The Sensemaking Process and Leverage Points for Analyst Technology as Identified Through Cognitive Task Analysis. In *International Conference on Intelligence Analysis*.

Siegelmann, H., and Sontag, E. 1995. On the Computational Power of Neural Nets.

Socher, R.; Perelygin, A.; Wu, J.; Chuang, J.; Manning, C.; Ng, A.; and Potts, C. 2013. Recursive deep models for semantic compositionality over a sentiment treebank. In *NIPS*, 1631–1642.

Strobelt, H.; Gehrmann, S.; Huber, B.; Pfister, H.; and Rush, A. M. 2016. Visual Analysis of Hidden State Dynamics in Recurrent Neural Networks. *Arxiv*.

Sutskever, I.; Vinyals, O.; and Le, Q. V. 2014. Sequence to Sequence Learning with Neural Networks. *Advances in Neural Information Processing Systems (NIPS)* 3104–3112.