

Recurrent Neural Network Based Visual Analytics Framework For Social Media

Ji Wang

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.

At the same time, the tweets also become a 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:

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

Opportunities and Motivation

There is an abundance of information available from social media, like Twitter. Today, more and more sources of news are from social media. Newman et al. (Newman 2011) find that the social media is more important than mainstream media, like TV, newspapers and magazine, to play a role to broadcast news and information to general audience. Due to the large participants and diversity of information, social media could have breaking news earlier than that in mainstream media (Kwak et al. 2010). The journalists in news agency highly focus on the news feeds and resources from social media, such as Twitter, Facebook and Instagram. However, even if the journalist professionals, verifying false

rumors and misinformation are still time-consuming and difficult work. In most of case, journalist professionals filter them via the information credibility (Silverman 2014).

Twitter 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 Flickr, 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.

Twitter platform is valuable to professional journalists in developing news stories.

Problem: Twitter information is disorganized, working with it creates high cognitive load and physical demands on users, discover of relevant tweets is difficult because they are not clustered by similarity

Existing Solutions (How is it done today, second step in Heilmair, including limits of existing solutions): Existing tools to help organize tweets include Tweet Deck, among others; elaborate on how

Proposed new solution is the tool to organize tweets efficiently.

The innovation part of my approach is creating hashtags for tweets of similar topic through RNN model. At the same time, our model also bring the tweet embedding into the tweets organizing process. User can efficiently find similar tweets within vector space of tweets embedding from RNN model.

Why will it be successful? (performance tests of resulting hashtags from RNN model indicate higher accuracy of hash-

tags assigned for clustering than other procedures - LDA?)

who cares? what difference will it make? (journalists should care; your tools should reduce cognitive load, improve discovery of information, etc.

I don't know about risks and costs or how long it will take Heilmier steps these may not be so pertinent for the overview.

What are the mid-term and final exams' to check for success? These would be the user evaluation studies you should describe in the overview that link back to the steps about why you think your proposed re-design of Tweetdeck will be successful (how will the user evaluation measure this?) and what difference will your re-design make (such as, cognitive and physical loads, news discovery; how will you measure these differences between TweetDeck and your re-design of TweetDeck?).

social media is surging as a new resource for the news discovery and production.

However, traditional tweets tool, like tweet deck, has a lot of drawbacks. not good

Users need to organize the tweets better

how to organize tweets better

predict hashtags via RNN: machine learning can help

solution: redesign tweet deck based on rnn model result for journalist for social media.

how to design a model. how to design

organize tweets better for news discovery. design tasks to fit the goal.

tool did this well, another well, and conclude claims.

better model will achieve the better user experience.

- Based on our knowledge, this is the first visual analytics tool to track, detect and analyze the rumor on social media for journalists society. We propose the first real time rumor visual analytics system including debunking algorithm for Twitter.
- We built this tool based on the pre-survey of the professional journalists. Thus, we design our visual analytics tool for rumor detection and analysis for journalists.
- We use the multiple "newsworthy tweets for our research. However, other study is most based on the single event, like Boston bombing. Thus, our research conclusions can be more general to apply to the normal rumor prediction and application.

Research Questions

RQ 1: - Can we use the results from RNN to develop for journalists? how to utilize rnn

- opportunity / motivation
- problem statement
- solution: organize tweets
- existing solution
- however, the failure of existing one
- one organize is hashtag, but not all have hashtag
- solution to organize tweets better with hashtag and embedding
- RNN model for hashtag and embedding.
- solution

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 (Pirulli and Card 2005) of RNN models understanding and interpretation. 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.

Design a UI for news discovery application

Usability of news discovery application

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?

In order to state them, two 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 we can build a RNN model to search
- How to design a system to organize tweets with the results from our RNN model?

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 effective are our tweets organization system designed to support sensemaking process of tweets understanding?

- What is the key design trade off among our RNN model, tweets representations, and interactive strategy?

- Would our proposed design visualization in RQ1 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?

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.

General Explanation of Tweet Embedding

How the tweet embedding works?

RNN is more like a dynamic programming. The distance of

of
 rnnvector incremental + vector \ll

News Discovery from Tweets

Traditional News Discovery Process from Tweets

Today, more and more sources of news are from social media. Newman et al. (Newman 2011) find that the social media is more important than mainstream media, like TV, newspapers and magazine, to play a role to broadcast news and information to general audience. Due to the large participants and diversity of information, social media could have breaking news earlier than that in mainstream media (Kwak et al. 2010). The journalists in news agency highly focus on the news feeds and resources from social media, such as Twitter, Facebook and Instagram. However, even if the journalist professionals, verifying false rumors and misinformation are still time-consuming and difficult work. In most of case, journalist professionals filter them via the information credibility (Silverman 2014).

According to the guideline in Verification Handbook (Silverman 2014), amid the community and practices of journalism, journalists and reporters spend large amount of time and attention to verify news source, contacts with the relevant

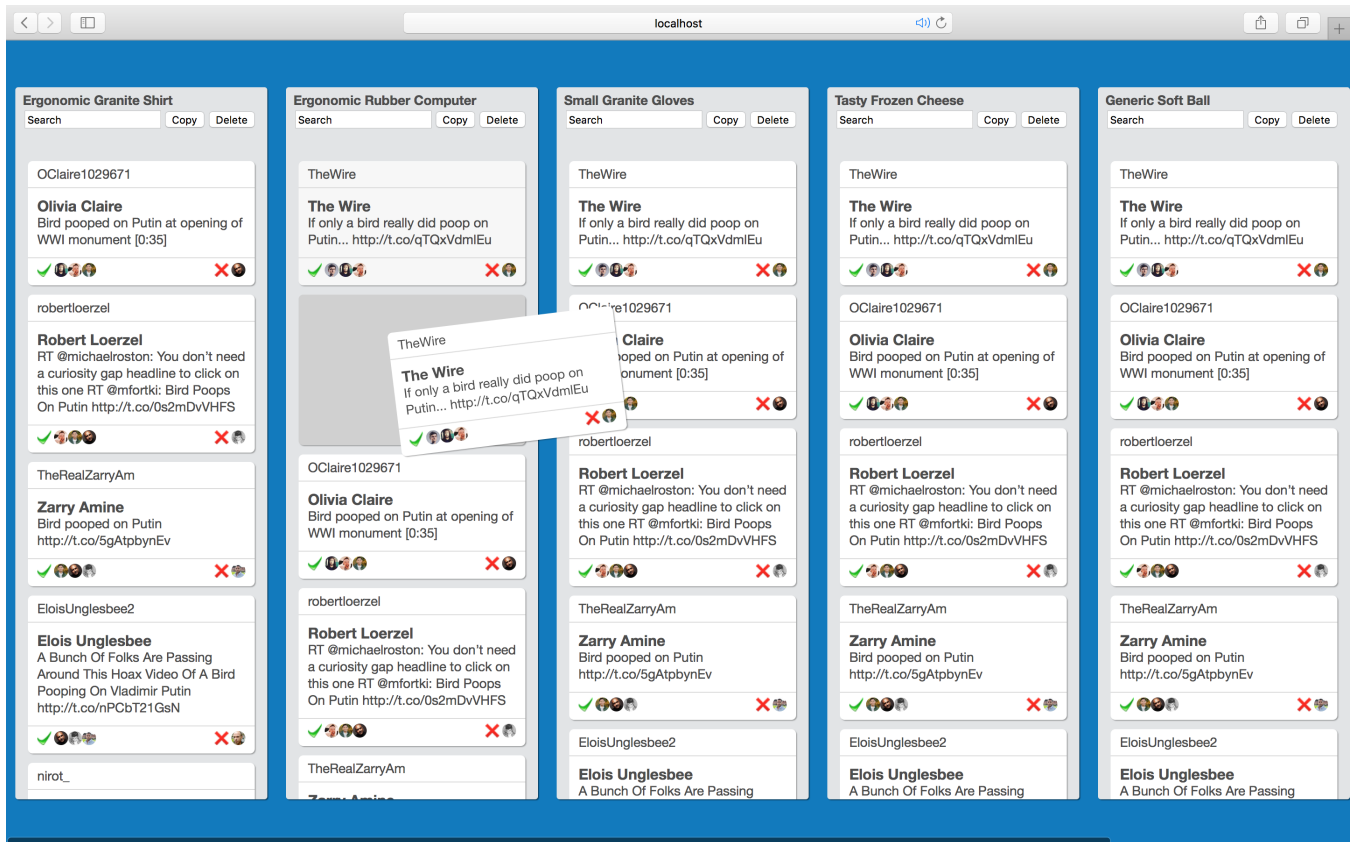


Figure 1: The UI view of vector deck for organizing tweets.

witnesses and build the evidence chain for the storylines. And finally make a report for the press.

The world leading news agency developed their own process to help their journalists and reporters to verify the rumor and misinformation from social media. For example, the Associated Press (AP) created a longstanding guidance on verification, rather than to new tools and technology. In AP, the verification process of user generated content challenged the traditional new agency, and web push them to directly contact with people to verify the information.

At the same time, the result of our preliminary survey for the journalists in Thomson Reuters is also similar as the guideline above. Their daily process to deal with the news resource from social media includes Discover and Verify steps:

- **Discover Step:** Most of social media journalists are using a tool, called TweetDeck, to monitor various channels in social media. It is a “de facto” industrial standard tool for many journalists from different news agencies.
- **Verify Step:** It is same as suggested in Verification Handbook (Silverman 2014):
 - Find the original source;
 - Evaluate the credibility of the source;
 - Try to manually verify by contacting the news source.

Need more information to verify

Cognitive psychology studies found that the tendency of individuals to estimate the likelihood of an event “by the ease with which instances or associations come to mind (Tversky and Kahneman 1973) (also known as availability heuristics). It means that people would like to assign more value to information they can easily recall, such as information that is more emotionally resonant or more recent (Schwarz et al. 1991). This tendency might lead to biases during the judgement and verification.

Avoiding these biases of journalists for credibility judgement is important in the news production process. Providing more information to journalists for their analysis is a solution to this problem. However, the overwhelming of social media information would require heavy mental demands and physical demands of journalists in their credibility assessment process.

For current journalists and news reporters for the social media, they firstly have to find the person who is responsible for the content. As explained in (Silverman 2014), “Researching the history of an uploader can mimic the characteristics of an old-fashioned police investigation, and perhaps also make you feel more like a stalker rather than a journalist or researcher.” At the same time, since people spend more of their lives on different social networks, they

do not realize their footprints can be identify as clues to build a meaningful news story.

User Interface Design

Hashtag

User Study

This section will introduce the research questions we would explore and detail information about our participants, experimental setting, and the study methodology we would deploy.

We will conduct a controlled experiment to evaluate the effectiveness of predicted hashtags and tweet embedding from RNN model for the tweets news discovery task.

Our study would compare the TwitterDeck like app with predicted hashtags and tweet embedding from RNN model to the other normal TwitterDeck like app without RNN model support.

In this study, we will recruit the undergraduate students from Department of Communication, Virginia Tech for this study. We suppose the participants would have experience with news discovery and news article writing with social media.

Compared Techniques

We will compare four different techniques in our study:

- **TwitterDeck Like App (T)** is implemented with the basic keywords search and tweets drag-and-drop functions of TwitterDeck. Similar with the keywords search normally, the keywords search in this technique is to match keywords exactly in the text of tweets. This widely adopted technique will be served as a baseline condition in our study.
- **TwitterDeck Like App with Predicted Hashtag only (H)** adds the predicted hashtag from RNN model for each tweets in TwitterDeck Like App. Users can organize tweets via the predicted hashtags in different columns. At the same time, users can also apply all the functions in the **TwitterDeck Like App (T)** above. The detail of hashtag predictions from RNN model has been introduced in Section 3.
- **TwitterDeck Like App with Tweet Embedding only (E)** plugs the tweet embedding from RNN model to each tweets in TwitterDeck Like App. In this technique, users can search the keywords for the relevant tweets in vector space. In this technique, we believe users would reach more relevant tweets without exactly matching in keywords. For the detail of tweet embedding from RNN model and search in vector space model, we have discussed in Section 3.
- **TwitterDeck Like App with both Predicted Hashtags and Tweet Embedding (H&E)** is the combination of third **H** and fourth **E** techniques. Users would organize and search tweets with the ultimate supports from RNN model.

The Figure 5 shows an user interface example of the four different techniques applied to the same tweets data.

Tasks and Apparatus

Users would be asked to perform a news discovery task from tweets via user interface techniques we provided. In the typical news discovery process, social media journalists would look for

Users were asked to perform a modied visual search task. In a conventional visual search task, the user looks for a specific target item within multiple distractors; a single instance of the target is usually present in about 50% of the presented images [30]. In our experiment, users were required to identify multiple targets within complex images that were visually enhanced by one of the highlighting techniques described above. This task represents a typical visual search task of knowledge workers, who need to scan related elements in multiple views or items related to their current selection in a single visualization. To assess whether users identified all the highlighted elements in the image, we asked them to count as accurately and quickly as possible the number of elements found. Users were instructed that accuracy was considered more important than task completion time.

Because we did not wish to limit our technique to a single scenario, we presented users with 16 different use case images for each technique. These images included scatterplot matrices, biochemical pathways, satellite images, and treemaps as well as multiple application windows containing text, images, and maps (Figure 10). In addition, we varied the number of highlighted elements across these images from 5 to 12, with two occurrences each.

Design and Procedure

Hypotheses

Results

Research Questions

In this paper, the goal we want to evaluate are two things:

Predicted Hashtags By RNN

Will the predicted hashtag can help the users to organize the tweets? How it works? Would the predicted hashtag would reduce the mental demand and physical demands of users?

Compared self-defined hashtag by users, what's the advantages of predicted hashtags and other previous techniques?

Vector Space Search by RNN

Will the vector space search feature can help users

A	B	C	D
Without Predicted Hashtag; Without Tweet Embedding	With Predicted Hashtag; Without Tweet Embedding	Without Predicted Hashtag; With Tweet Embedding	With Predicted Hashtag; With Tweet Embedding

Figure 2: The UI view of vector deck for organizing tweets.

Proposed Study Methodology

Evaluation Results

Contributions

The first contribution of this paper is to use the RNN model to predict the hashtag and generate RNN based tweet embedding to support the vector space model. These two highlight spots

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.

The second contribution is that we conducted a controlled experiment that compared the use of four different techniques to support typical tweets news discovery tasks of social media journalists: simple twitter deck like approach, approach with predicted hashtag only, approach with tweet embedding search only, and approach with both predicted hashtags and tweet embedding search.

We expect the results of this experiment would show that predicted hashtags and tweet embedding search from RNN model could improve performance on tasks for social media news discovery related mounts of tweets. We also expect the results would show that the approach with both predicted hashtags and tweet embedding search is the most effective of the four techniques we tested.

This research will increase the understanding of interaction within visual analytic applications through exploring spatial interaction, a new design space of interaction within visual analytic applications. Designing applications for spatial interaction should lead to more usable and powerful visualizations capable of increasing users insights into complex datasets. Specific contributions of this work include: An interaction design space (spatial interaction) for exploratory analysis of textual information within a spatialization Identification of challenges and changes to mathematical models and visualization frameworks to support spatial interaction Evaluation of effectiveness of spatial interaction for intelligence analysis of textual information Overall, these contributions are focused on advancing the goal of visual analytics of enabling users to gain insight into complex datasets by combining human intuition, statistical computation, and visualization.

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.

Kwak, H.; Lee, C.; Park, H.; and Moon, S. 2010. What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web - WWW '10*, 591. New York, New York, USA: ACM Press.

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.

Newman, N. 2011. Mainstream media and the distribution of news in the age of social discovery. *Reuters Institute for the Study of Journalism* (September):58.

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*.

Schwarz, N.; Bless, H.; Strack, F.; Klumpp, G.; Rittenauer-Schatka, H.; and Simons, A. 1991. Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology* 61(2):195–202.

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

Silverman, C. 2014. *Verification Handbook: A definitive guide to verifying digital content for emergency coverage*. First edition.

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.

Tversky, A., and Kahneman, D. 1973. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* 5(2):207–232.