Toward Multi-Modal Stance Learning

Sumeet Kumar and Kathleen M. Carley

Carnegie Mellon University
5000 Forbes Ave, Pittsburgh, PA 15213, USA
Email: sumeetku@cs.cmu.edu, kathleen.carley@cs.cmu.edu

Abstract Learning stance of social-media users is an emerging area of research. Recent work on stance mining have considered users' posts, users' connections or both to find their stance. However, besides posting and connecting, social-media platforms like Twitter allow many more types of interactions, e.g. 'sharing others posts' (i.e., retweeting and quoting), 'liking posts', 'replying to others posts', 'grouping users' (i.e., making lists) and more. Despite notable progress in improving the accuracy of stance classifiers using some of the above-listed interactions, it's is unclear which interactions are more useful for stance learning and why. In this article, we explore these interactions and highlight their merits for stance learning. Moreover, we suggest how these interactions could be combined to build a joint stance learning model.

1 Introduction

People express their opinions on blogs and other social media platforms. Automated ways to categorize views of people in such user-generated corpora is of immense value. In particular, the machine learning approaches for *stance learning, which involves learning people's opinion about a topic of interest* is an emerging area of research [4,7]. Most existing studies on stance learning take a simplistic view that assumes a social-media 'post' (like a Tweet) holds a perspective that is independent of the context and the author. This approach to stance learning ignores that group norms and moderators could impact stance by setting an emotional tone [6] and complex interactions exist among social media. For example on Twitter, besides posting Tweets and connecting to others by following, one can engage in 'sharing others posts' (i.e., retweeting and quoting), 'liking posts', 'replying to others posts' and 'grouping users' (i.e., making lists). Despite notable progress in improving the accuracy of stance classifiers using some of the above-listed interactions, it's is unclear which interactions are more useful for training stance learning model and why. There is a need to explore the merits of different types of interactions to understand their utility for stance learning. For example, 'liking' a post that is pro gun-control is a much clear indicator of pro gun-control stance than following a user who is pro gun-control. In this article, we review prior research on using different interactions (modalities) for stance learning and propose future research directions.

2 Modalities For Stance Learning

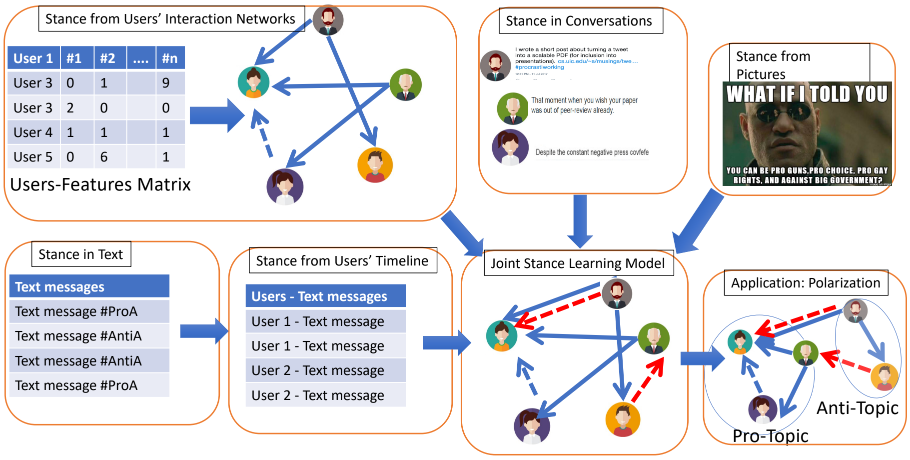

Figure 1. Proposed joint stance learning model: The different colors of lines illustrate different interactions. Solid lines indicate explicit relationships (like follower-followee) and dashed lines indicate inferred relationships. Large blue arrows show one way to combine various modalities and its application in finding polarized communities.

As shown in In Fig. 1, stance of social-media users could be learned from different modalities. Mohammad et al. used text to detect stance [4]. Zubiaga et al. used conversations to learn stance in reply posts [7]. Lu et al. proposed BiasWatch, a bias propagation method to infer opinion bias of Twitter users using their networks [3]. Most existing work use one or two modalities to learn stance. In this article, we propose to use multi-modal interactions. Use of multi-modal interactions for stance learning is non-trivial. Because different interactions result in multiple text or interaction graphs or a combination of both, multi-modal stance learning is challenging. Though there is limited research in combining different modalities for stance learning, there is a quite some work on multi-modal machine learning ([1]). A common technique is to embed all modalities in a continuous representation space. A good stance learner model should be able to use all text and graphs data. For converting text to embedding space, multiple ways to encode sentences (like [2]) are available. To convert graphs to a continuous representation space, several techniques have been proposed recently. For example Perozzi et al. [5] introduced DeepWalk, a model to learn latent vector representations for nodes that encode social relations by using short random walks. Therefore, using a continuous representations space for all types of interactions, we can build a joint model based on inputs of different modalities.

3 Summary of Proposed Work

Display of stance in person-to-person conversation is inherently multi-modal. In addition to speech, we use posture, gesture and eye contact to convey information in a face-to-face communication. The same way, interactions on social media are also multi-modal. In this research, we highlight some of the potential interactions and their value for stance learning. These modalities bring new perspectives to the stance learning task. Therefor, we propose that the machine learning based stance classifiers should try to combine multiple interaction modalities.

To find stance from text, we propose the use of weakly supervised learning methods as it reduces the cost of data collect. In particular, specific hashtags found on social media platforms that carry stance information should be tried. There is also a need to extend the text-based stance learning to use multiple posts from the same user. Conversation threads should also be used to learn the authors' position about an issue based on their posts and the replies to these posts. Users can reveal their alignment with other users based on the stance they take in replying to other's posts. Various interaction networks on Twitter (e.g. follower, mentions, likes) would result in graph structured data which needs networks based machine learning models. After developing independent models for different types of interactions, we suggest to use a joint model that combines the various sources of stance information. To combine different modalities of information (e.g., users' follower graphs and users' likes), one can embed features from all information sources into one continuous representation space. The importance given to a particular modality should be derived from the data itself as this would depend on the topic under consideration. We expect that such multi-modal stance learning techniques could be useful in many areas. For example, we can apply these models to find polarized communities.

References

1. Baltrušaitis, T., Ahuja, C., Morency, L.P.: Multimodal machine learning: A survey and taxonomy. IEEE Transactions on Pattern Analysis and Machine Intelligence (2018)
2. Kiros, R., Zhu, Y., Salakhutdinov, R.R., Zemel, R., Urtasun, R., Torralba, A., Fidler, S.: Skip-thought vectors. In: Advances in neural information processing systems. pp. 3294–3302 (2015)
3. Lu, H., Caverlee, J., Niu, W.: Biaswatch: A lightweight system for discovering and tracking topic-sensitive opinion bias in social media. In: Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. pp. 213–222. ACM (2015)
4. Mohammad, S.M., Sobhani, P., Kiritchenko, S.: Stance and sentiment in tweets. ACM Transactions on Internet Technology (TOIT) 17(3), 26 (2017)
5. Perozzi, B., Al-Rfou, R., Skiena, S.: Deepwalk: Online learning of social representations. In: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 701–710. ACM (2014)
6. Tee, E.Y., Paulsen, N., Ashkanasy, N.M.: Revisiting followership through a social identity perspective: The role of collective follower emotion and action. The Leadership Quarterly 24(6), 902–918 (2013)
7. Zubiaga, A., Kochkina, E., Liakata, M., Procter, R., Lukasik, M.: Stance classification in rumours as a sequential task exploiting the tree structure of social media conversations. arXiv preprint arXiv:1609.09028 (2016)