Eurographics Conference on Visualization (EuroVis), Posters Track (2016)
T. Isenberg and F. Sadlo (Guest Editors)

Visual Analysis of Text Annotations for Stance Classification with ALVA

K. Kucher[1], A. Kerren[1], C. Paradis[2], and M. Sahlgren[3]

[1]Department of Computer Science, Linnaeus University, Sweden
[2]Centre for Languages and Literature, Lund University, Sweden
[3]Gavagai AB, Sweden

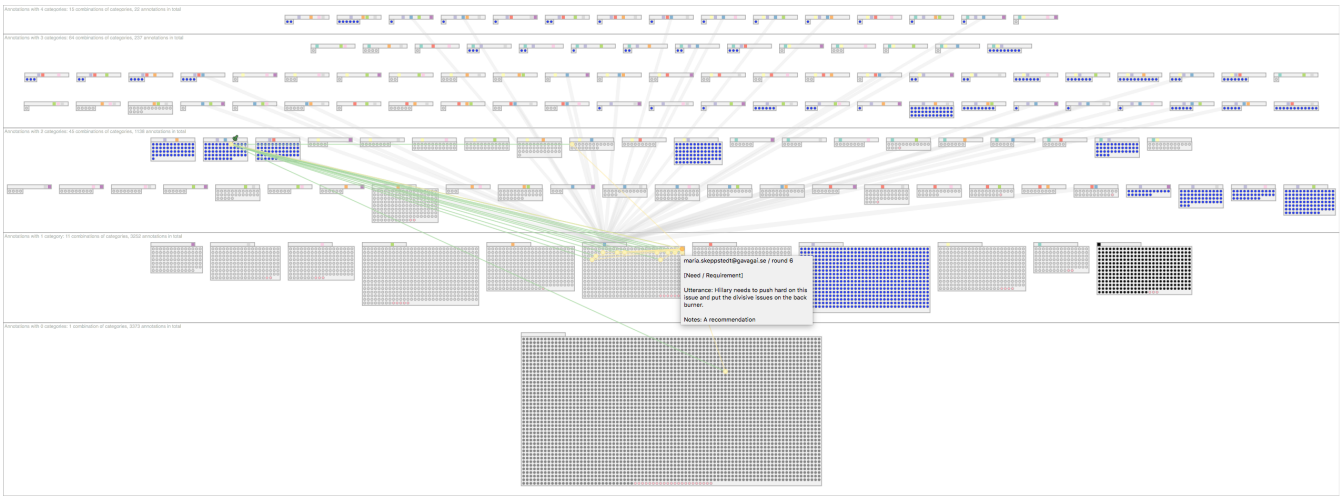

Figure 1: *Visualization of about 8,000 text annotations in our VA system ALVA. Each annotation represented by a colored dot can be labeled with up to ten stance categories in our concrete use case. Annotations are grouped together into rectangular blocks by the combination of categories which occur in the data set. Thus, the block groups form layers: the top layer contains 15 annotation blocks labeled with four categories simultaneously, and the bottom layer with a single block solely contains neutral annotations. Color-coded rectangles in the block headers represent the corresponding sets of categories. Here, all annotations related to the category "Concession and Contrariness" are highlighted in blue.*

Abstract

The automatic detection and classification of stance taking in text data using natural language processing and machine learning methods create an opportunity to gain insight about the writers' feelings and attitudes towards their own and other people's utterances. However, this task presents multiple challenges related to the training data collection as well as the actual classifier training. In order to facilitate the process of training a stance classifier, we propose a visual analytics approach called ALVA for text data annotation and visualization. Our approach supports the annotation process management and supplies annotators with a clean user interface for labeling utterances with several stance categories. The analysts are provided with a visualization of stance annotations which facilitates the analysis of categories used by the annotators. ALVA is already being used by our domain experts in linguistics and computational linguistics in order to improve the understanding of stance phenomena and to build a stance classifier for applications such as social media monitoring.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Information Interfaces and Presentation (e.g., HCI)]: General—Graphical user interfaces (GUI) I.2.7 [Artificial Intelligence]: Natural Language Processing—Text analysis

1. Introduction

The intense development of digital technologies and phenomena such as the social media boom have resulted in the availability of massive amounts of textual data. There already exist methods originating in both computational linguistics (CL) / natural language processing (NLP) and information visualization (InfoVis) / visual analytics (VA) that are successfully used to analyze texts for academic and industrial purposes. While some aspects of text data, such as sentiment or topical structure, are relatively well-studied [PL08], research on stance [Eng07] is actively ongoing in linguistics and CL [SSBS*15, SPKS15]. The analysis of stance in written language can reveal the feelings and attitude of writers (utterers) towards their own and other people's utterances. Analysis results can range from general agreement/disagreement to fine-grained aspects such as the expression of wishes and emotions.

DOI: 10.2312/eurp.20161139

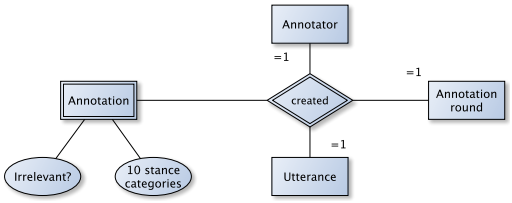

Figure 2: *Entity-Relationship diagram for text annotations.*

However, the existing efforts to build a fine-grained stance classifier are facing multiple challenges related to the collection of training data as well as the actual design and training of a machine learning (ML) classifier. As part of a collaboration with researchers in linguistics and CL, we have designed an integrated visual analytics solution to support all stages of the annotation and classifier training process. In this poster abstract, we introduce our system, called ALVA (Active Learning & Visual Analytics), focusing on aspects related to analysis and visualization of the collected stance annotation data.

2. Related work

Our work is relevant to the visual support of data annotation and classifier training. Thus, previous work includes contributions from the visualization community as well as CL and ML communities. For instance, the approaches proposed by Settles [Set11] and Huang et al. [HTFA13] provide user interfaces for text annotation, however, they do not support visual analysis of the collected data, which makes it difficult to get an overview during/after the training. The related work in visualization includes the approach described by Lu et al. [LKT*14] that focuses on feature selection and training various ML models for predictive tasks involving social media data. While the overall goal of our approach is also to facilitate an ML classifier training, we currently focus on different tasks such as the visual analysis of annotated data. Another approach is discussed by Heimerl et al. [HKBE12] who provide an interactive interface for training a text classifier for two mutually exclusive categories, namely, *relevant* and *irrelevant*. Their approach includes support for data annotation and several visualizations based on dimensionality reduction of feature vectors as well as classification confidence values. The main difference of our approach lies in the choice of categories and corresponding tasks: our stance annotations are effectively vectors with eleven binary dimensions, and scatterplots usually used for the representation of dimensionality reduction results do not support the tasks important for our users.

The specific application in the focus of our work is stance visualization, which has not been described well in existing work in InfoVis or VA. Almutairi [Alm13] discusses the application of the *appraisal* framework to text analysis and visualization. The resulting system supports data annotation, classification of text fragments with regard to *affect*, *judgement*, and *appreciation*, and several visualizations. In contrast, our work uses a different (and much larger) set of categories which requires a different approach to annotated data representation. Finally, our previous work [KSBK*16] introduces the problem of stance visualization and proposes a VA approach based on sentiment analysis—however, it does not focus on the visualization of annotated data.

3. Visualization methodology

The user tasks for visualization in ALVA are related to the explorative analysis of individual annotations as well as distributions of stance categories used in the annotations. Our use case analysts want to be provided with an overview of the overall data set, analyze the co-occurrence of categories, identify interesting cases related to multiple categories, and compare annotations made for the same utterances. The current data set comprises about 8,000 annotations of utterances in English (in most cases, individual sentences) collected from social media on political topics such as the US election. The annotations were performed by several annotators during multiple annotation rounds (see Figure 2). Each annotation can be either marked as *irrelevant*, or labeled—in our case—with up to ten stance categories such as *agreement*, *volition*, etc. This results in a multidimensional data set presenting multiple visualization challenges. For instance, color coding is not feasible for individual annotation items since the total number of possible category combinations is $2^{10}+1$.

After experimenting with scatterplot representations, we have designed a new visual approach called *CatCombos* that combines the semantic substrates principle by Shneiderman and Aris [SA06] with the features of the set visualization technique proposed by Sadana et al. [SMDS14]. Figure 1 shows an ALVA view of the complete annotations data set which realizes the CatCombos approach. Individual annotations are represented by small dots and grouped by the combination of categories used. Each combination is represented by a rectangular block with a header that contains a set of color-coded labels corresponding to the stance categories. Individual blocks are, in turn, grouped and laid out by the number of corresponding categories.

The resulting visualization provides an overview of the stance categories distribution and the characteristics of annotations important for our analysts. For instance, it is easy to estimate that the majority of annotations are contained in blocks at the bottom that correspond to a single selected stance category or no category at all, i.e., *neutral* utterances. The interesting cases where an annotator has used multiple categories simultaneously are located in the top blocks. The visualization supports dynamic queries and highlighting / details on demand for individual blocks and annotations. By hovering over an annotation item, the users can view its details and see links to other annotations for the same utterance. It is also possible to preserve the selection of individual items by clicking. Simultaneously, semi-transparent grey links to blocks with related categories of combinations are displayed for the enclosing block.

4. Conclusions and future work

In this poster abstract, we have briefly introduced our system ALVA designed to facilitate the text annotation process for training a stance classifier by researchers in linguistics and computational linguistics. ALVA supports the visual analysis of annotations by grouping the representations of individual annotations based on stance category combinations. Our future work includes the support for other aspects of the annotation process as well as interactive training and analysis of the stance classifier.

Acknowledgements This research was funded by the framework grant "The Digitized Society – Past, Present, and Future" with No. 2012-5659 from the Swedish Research Council.

References

[Alm13] ALMUTAIRI B. A. A.: Visualizing patterns of appraisal in texts and corpora. *Text & Talk 33*, 4-5 (2013), 691–723. 2

[Eng07] ENGLEBRETSON R. (Ed.): *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*, vol. 164 of *Pragmatics & Beyond New Series*. John Benjamins, 2007. doi:10.1075/pbns.164. 1

[HKBE12] HEIMERL F., KOCH S., BOSCH H., ERTL T.: Visual classifier training for text document retrieval. *IEEE Transactions on Visualization and Computer Graphics 18*, 12 (2012), 2839–2848. doi:10.1109/TVCG.2012.277. 2

[HTFA13] HUANG S.-W., TU P.-F., FU W.-T., AMANZADEH M.: Leveraging the crowd to improve feature-sentiment analysis of user reviews. In *Proceedings of the 2013 International Conference on Intelligent User Interfaces* (2013), IUI '13, ACM, pp. 3–14. doi:10.1145/2449396.2449400. 2

[KSBK*16] KUCHER K., SCHAMP-BJERede T., KERREN A., PARADIS C., SAHLGREN M.: Visual analysis of online social media to open up the investigation of stance phenomena. *Information Visualization 15*, 2 (2016), 93–116. doi:10.1177/1473871615575079. 2

[LKT*14] LU Y., KRÜGER R., THOM D., WANG F., KOCH S., ERTL T., MACIEJEWSKI R.: Integrating predictive analytics and social media. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology* (2014), VAST '14, IEEE, pp. 193–202. doi:10.1109/VAST.2014.7042495. 2

[PL08] PANG B., LEE L.: Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval 2*, 1–2 (2008), 1–135. doi:10.1561/1500000011. 1

[SA06] SHNEIDERMAN B., ARIS A.: Network visualization by semantic substrates. *IEEE Transactions on Visualization and Computer Graphics 12*, 5 (2006), 733–740. doi:10.1109/TVCG.2006.166. 2

[Set11] SETTLES B.: Closing the loop: Fast, interactive semi-supervised annotation with queries on features and instances. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (2011), EMNLP '11, Association for Computational Linguistics, pp. 1467–1478. 2

[SMDS14] SADANA R., MAJOR T., DOVE A., STASKO J.: OnSet: A visualization technique for large-scale binary set data. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1993–2002. doi:10.1109/TVCG.2014.2346249. 2

[SPKS15] SKEPPSTEDT M., PARADIS C., KERREN A., SAHLGREN M.: Finding infrequent phenomena in large corpora using distributional semantics. In *Poster Abstracts of the Symposium on Methods and Linguistic Theories* (2015), MaLT '15. 1

[SSBS*15] SKEPPSTEDT M., SCHAMP-BJERede T., SAHLGREN M., PARADIS C., KERREN A.: Detecting speculations, contrasts and conditionals in consumer reviews. In *Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (2015), WASSA '15, Association for Computational Linguistics, pp. 162–168. 1