Sentiment Polarity Detection for Software Development

Fabio Calefato, Filippo Lanubile, Federico Maiorano, Nicole Novielli

University of Bari Aldo Moro

Bari, Italy

{fabio.calefato, filippo.lanubile, nicole.novielli}@uniba.it, f.maiorano2@studenti.uniba.it

CCS CONCEPTS

• **Software creation and management** → **Collaboration in software development**; *Programming teams*; • **Information retrieval** → Retrieval tasks and goals; *Sentiment analysis[1]*

KEYWORDS

Sentiment Analysis, Communication Channels, Stack Overflow, Word Embedding, Social Software Engineering

EXTENDED ABSTRACT

The role of sentiment analysis is increasingly emerging to study software developers' emotions by mining crowd-generated content within software repositories and information sources. With a few notable exceptions [1][5], empirical software engineering studies have exploited off-the-shelf sentiment analysis tools. However, such tools have been trained on non-technical domains and general-purpose social media, thus resulting in misclassifications of technical jargon and problem reports [2][4]. In particular, Jongeling et al. [2] show how the choice of the sentiment analysis tool may impact the conclusion validity of empirical studies because not only these tools do not agree with human annotation of developers' communication channels, but they also disagree among themselves.

Our goal is to move beyond the limitations of off-the-shelf sentiment analysis tools when applied in the software engineering domain. Accordingly, we present Senti4SD, a sentiment polarity classifier for software developers' communication channels. Senti4SD exploits a suite of lexicon-based, keyword-based, and semantic features for appropriately dealing with the domain-dependent use of a lexicon. We built a Distributional Semantic Model (DSM) to derive the semantic features exploited by Senti4SD. Specifically, we ran word2vec [3] on a collection of over 20 million documents from Stack Overflow, thus obtaining word vectors that are representative of developers' communication style. The classifier is trained and validated using a gold standard of 4,423 Stack Overflow posts, including questions, answers, and comments, which were manually annotated for sentiment polarity.

We release the full lab package[2], which includes both the gold standard and the emotion annotation guidelines, to ease the execution of replications as well as new studies on emotion awareness in software engineering. To inform future research on word embedding for text categorization and information retrieval in software engineering, the replication kit also includes the DSM.

Results. The contribution of the lexicon-based, keyword-based, and semantic features is assessed by our empirical evaluation leveraging different feature settings. With respect to SentiStrength [6], a mainstream off-the-shelf tool that we use as a baseline, Senti4SD reduces the misclassifications of neutral and positive posts as emotionally negative. Furthermore, we provide empirical evidence of better performance also in presence of a minimal set of training documents.

ACKNOWLEDGMENTS

REFERENCES

[1] A. Blaz and K. Becker. 2016. Sentiment analysis in tickets for IT support. In Proc. of MSR 2016, 235-246.

[2] R. Jongeling, P. Sarkar, S. Datta, and A. Serebrenik. 2017. On negative results when using sentiment analysis tools for software engineering research. Empirical Software Engineering. 22, 5, 2543-2584.

[3] T. Mikolov, K. Chen, G. Corrado, J. Dean. 2013. Efficient estimation of word representations in vector space. CoRR abs/1301.3781.

[4] N. Novielli, F. Calefato, and F. Lanubile. 2015. The challenges of sentiment detection in the social programmer ecosystem. In Proc. of SSE 2015, 33-40.

[5] S. Panichella, A. Di Sorbo, E. Guzman, C. A. Visaggio, G. Canfora, and H. C. Gall. 2015. How can i improve my app? Classifying user reviews for software maintenance and evolution. In Proc of. ICSME '15, 281-290.

[6] M. Thelwall, K. Buckley, and G. Paltoglou. 2012. Sentiment strength detection for the social web. J. Am. Soc. Inf. Sci. Technol. 63, 1, 163-173.

[2] https://github.com/collab-uniba/Senti4SD