# The Best of Both Worlds: Sentiment Analysis in Call Center Conversations using Audio Search Techniques

Naomi Zeichner       Inna Achlow       Victor Shafran

NICE Systems, Zarhin 13, Ra'anana

## Abstract

Sentiment analysis (SA) is a central component in industrial text analytic solutions enabling the automatic detection of customers opinions regarding products and services. However, even though state-of-the-art techniques achieve a very high level of accuracy on written texts, these techniques perform significantly worse on spoken conversation which is still the most prominent way for customers to interact with their service providers.

A common approach to SA, and the one adopted in our in-house solution, is to use a lexicon of sentiment terms extended with negation, intensifying or weakening terms[4]. When it comes to performing SA over audio most works[2] use automatically generated transcriptions to run conventional SA techniques. Alternatively, in the field of speech analytics Emotion Detection uses acoustic and prosodic cues such as pitch and tone and recently a number of works used multi-modal elements to enrich SA[3].

In this work we address the problem of SA in call center conversations focusing on two perspectives. The first is a study into the language register used for sentiment in call-center conversation. Such data differs greatly not only from textual but also from other speech corpora featuring multiple dynamic and spontaneous speech characteristics[1]. Additionally, interactions are much longer and contain fewer sentiment events. The uniqueness of the medium requires a refinement of the sentiment definition and a linguistic adaptation of the lexicon to the data. The definition should account for differences and correlations of politeness vs sentiment; agent vs customer

sentiment and chronological sentiment intensity. The second perspective we explore is the technological one. Two main technologies used in speech processing are phrase spotting and transcription based. The first boasts high accuracy and detection but stumbles on short or similar sounding words and doesn't allow for part-of-speech tagging, stemming or search for non-sequential phrases. The latter uses a language model to address such issues but suffers from a limited vocabulary and, especially in spontaneous conversation, tends to have a high word-error-rate.

Our approach in this work is to tailor a lexicon based on in-depth analysis of the way people express sentiment in call-center conversations and to combine the two types of speech technologies using an underlying traditional SA approach. We start by searching sentiment terms using phrase spotting methods and then resolve short negation and non-sequential terms using automatic transcriptions. Evaluation is performed by comparing results to manually transcribed calls annotated with sentiment events and overall document sentiment. We measure performance using different lexicons, varying degrees of phonetic search speed and certainty, and different ways of combining between the technologies.

We conclude with an analysis of our findings and describe future development directions we aim to pursue. For example, we would like to explore the multi-modal approach by adding acoustic cues, a natural extension given that the analysis receives audio as input rather than text. In addition we would like to extend the analysis to be more fine-grained differentiating between such emotions as anger, surprise and joy providing a more comprehensive picture of the sentiment expressed in the calls.

**Keywords:** Text Analytics, Text mining, Sentiment Analysis, Speech Analytics.

# References

[1] C Clavel, G Adda, F Cailliau, M Garnier-Rizet, A Cavet, G Chapuis, S Courcinous, C Danesi, A Daquo, M Deldossi, S Guillemin-Lanne, M Seizou, and P Suignard. Spontaneous speech and opinion detection: mining call-centre transcripts. *Language Resources and Evaluation*, 47(4):1089–1125, 2013.

[2] L Kaushik, A Sangwan, and J Hansen. *Sentiment extraction from natural audio streams*, pages 8485–8489. 2013.

[3] V Rosas, R Mihalcea, and L.P. Morency. Multimodal Sentiment Analysis of Spanish Online Videos. In *IEEE Intelligent Systems*, 2013.

[4] M Taboada, J Brooke, M Tofiloski, K Voll, and M Stede. Lexicon-based methods for sentiment analysis. *Comput. Linguist.*, 37:267–307.