

## iFeel 2.0: A Multilingual Benchmarking System for Sentence-Level Sentiment Analysis

**Matheus Araújo**

CS Dept, UFMG, Brazil  
matheus.araujo@dcc.ufmg.br

**João P. Diniz**

CS Dept, UFMG, Brazil  
jpaulo@dcc.ufmg.br

**Lucas Bastos**

CS Dept, UFMG, Brazil  
lucasmbastos@dcc.ufmg.br

**Elias Soares**

CS Dept, UFMG, Brazil  
eliassoares@dcc.ufmg.br

**Manoel Júnior**

CS Dept, UFMG, Brazil  
manoelrmj@dcc.ufmg.br

**Miller Ferreira**

CS Dept, UFMG, Brazil  
millermarques@dcc.ufmg.br

**Filipe Ribeiro**

CS Dept, UFOP, Brazil  
filiperibeiro@dcc.ufmg.br

**Fabrcio Benevenuto**

CS Dept, UFMG, Brazil  
fabricio@dcc.ufmg.br

### Abstract

Sentiment analysis became a hot topic, specially with the amount of opinions available in social media data. With the increasing interest in this theme, several methods have been proposed in the literature. Recent efforts have showed that there is no single method that always achieves the best prediction performance for different datasets. Additionally, novel methods have not being extensively compared with other methods and across different datasets, specially methods that are not designed to the English language. Consequently, researchers tend to accept any popular method as a valid methodology to measure sentiments, a practice that is usual in science. In this context, we propose iFeel 2.0, an online web system that implements 19 sentence-level sentiment analysis methods and allows users to easily label a dataset with all of them. iFeel aims at easing the comparison of new methods with baseline approaches and can also be helpful for those interested in using sentiment analysis, allowing them to choose an appropriate sentiment analysis method that works fine for a new dataset. We also incorporate a multiple language feature to allow methods designed for specific languages to be easily compared with a baseline approach that simply translates the input data to English and run these 19 methods. We hope this system can represent an important contribution to this field.

### Introduction

Sentiment analysis is a research field that is growing in the same pace as the amount of available data in social media systems. In the last few years thousands of scientific papers have investigated sentiment analysis, several startups that measure opinions on real data have emerged, and a number of innovative products related to this theme have been developed. There are multiple methods for measuring sentiments, including lexical-based approaches and supervised machine learning methods. Recent efforts have showed that there is no single method that always achieves the best prediction performance for different datasets (Ribeiro et al. 2015).

As the state-of-the-art has not been clearly established and a method's performance is highly dependent on the input data, researchers tend to accept any popular method as

a valid methodology to measure sentiments. It is common to see concurrent important papers, sometimes published in the same conference or journal, using completely different methods. This is a very unusual situation from a scientific perspective, in which benchmark comparisons are the rule.

In this context, we propose iFeel 2.0,<sup>1</sup> a benchmark system for sentence-level sentiment analysis. First published at (Araújo et al. 2014), iFeel implemented only 8 methods. Its second version increased the set of methods to 19. More important, iFeel do not only provide its users access to these methods. We also add two important functionalities on iFeel 2.0. The first is a benchmark feature in which researchers can easily compare the performance of a new sentiment analysis method with 19 "state-of-the-practice" methods. We also put together 19 public datasets that include text from forums, twitter, comments on YouTube, and in Blogs, reviews, etc. to the research community so that new methods can be compared across different datasets as well.

Our second feature consists of a multilingual support for 60 different languages. The idea behind this feature comes from another previous work (Araújo et al. 2016) that shows that a number of methods proposed for specific languages are not necessarily better than a simple approach that translates the input text to English and run these 19 methods designed to English implemented on iFeel 2.0. Machine translation should be a simple baseline for novel methods designed for specific languages. For this reason, we incorporate this feature into iFeel 2.0.

iFeel allows users to submit just a simple sentence in the front page or upload a file with up to 1,000 sentences. The result is a *.xml* or *.xlsx* file containing the output of all methods implemented. Additionally, we share our codes and datasets aiming at contributing to this research field and to scientific community working on this topic. Next, we briefly summarize the methods we have implemented.

### Sentiment Analysis Methods

We incorporate new 11 methods as part of iFeel 2.0 in addition to Emoticons, Happiness Index, SentiWordNet, Senticnet, Panas-t, Sentistrength, SASA, implemented on

<sup>1</sup>iFeel is hosted on <http://www.ifeel.dcc.ufmg.br>

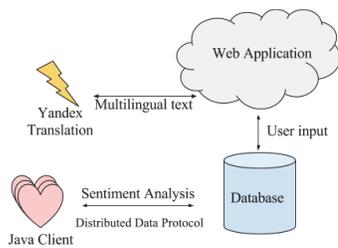


Figure 1: iFeel Architecture

iFeel 1.0: These methods are Opinion Lexicon (Hu and Liu 2004), Opinion Finder (Wilson, Wiebe, and Hoffmann 2005), AFINN (Nielsen 2011), SO-CAL (Taboada et al. ), Emoticons Distant Supervision (Hannak et al. 2012), NRC Hashtag (Mohammad 2012), EmoLex (Mohammad and Turney 2013), SANN (Pappas and Popescu-Belis 2013), Sentiment140 Lexicon (Mohammad, Kiritchenko, and Zhu 2013), Stanford Recursive Deep Model (Socher et al. 2013), Umigon (Levallois 2013), Vader (Hutto and Gilbert 2014). All those methods are deeply discussed in (Ribeiro et al. 2015).

### iFeel Architecture

The iFeel 2.0 Web application has changed from the original iFeel 1.0 as well. Now it uses Meteor <sup>2</sup>, a full stack web framework that uses NodeJS as server-side javascript engine and MongoDB to persist data. The superficial architecture of iFeel 2.0 is shown on Figure 1. It corresponds to a Web Application that has support from the translation API and shares a database to a java client.

The Java client is the sentiment analysis core in our system that puts all 19 methods implemented to run in different threads. It implements a Distributed Data Protocol (DDP) to communicate with the Web Application. When a user uploads a sentence or a file, everything is saved on the database and the java client is alerted to request those sentences. After analysing, everything is submitted back to the database and the access to the results is granted to the user.

### Multilanguage Support

As discussed before, our approach for multilanguage sentiment consists of a basic baseline for novel efforts on sentence-level sentiment analysis. It consists of translating the input text to English and run the 19 methods implemented on iFeel. To translate text we use the Yandex Translation API<sup>3</sup>, as it is a free API that converts user input from 60 different languages when requested.

### Conclusion

In this work, we proposed iFeel 2.0, a system that implements 19 sentiment analysis methods. The system aims at allowing new methods, proposed in English or other languages, to be compared with 19 known sentence level sen-

timent analysis methods. We also hope our system can represent an important contribution for researchers interested in using sentiment analysis as it allows users to investigate which method would perform better in novel datasets.

### Acknowledgments

This work was partially supported by the project FAPEMIG-PRONEX-MASWeb, Models, Algorithms and Systems for the Web, process number APQ-01400-14, and by individual grants from CNPq, CAPES, and Fapemig.

### References

- Araújo, M.; Gonçalves, P.; Benevenuto, F.; and Cha, M. 2014. ifeel: A system that compares and combines sentiment analysis methods. In *Proc. of WWW*.
- Araújo, M.; Reis, J. C.; Pereira, A. C.; and Benevenuto, F. 2016. An evaluation of machine translation for multilingual sentence-level sentiment analysis. In *Proc. of ACM SAC*.
- Hannak, A.; Anderson, E.; Barrett, L. F.; Lehmann, S.; Mislove, A.; and Riedewald, M. 2012. Tweetin' in the rain: Exploring societal-scale effects of weather on mood. In *Proc. of ICWSM*.
- Hu, M., and Liu, B. 2004. Mining and summarizing customer reviews. In *Proc. of KDD*.
- Hutto, C., and Gilbert, E. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proc. of ICWSM*.
- Levallois, C. 2013. Umigon: sentiment analysis for tweets based on terms lists and heuristics. In *Proc. of SemEval*.
- Mohammad, S., and Turney, P. D. 2013. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence* 29(3).
- Mohammad, S. M.; Kiritchenko, S.; and Zhu, X. 2013. Nrc-canada: Building the state-of-the-art in sentiment analysis of tweets. In *Proc. of SemEval*.
- Mohammad, S. 2012. #emotional tweets. In *Proc. of SemEval*.
- Nielsen, F. Å. 2011. A new anew: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*.
- Pappas, N., and Popescu-Belis, A. 2013. Sentiment analysis of user comments for one-class collaborative filtering over ted talks. In *Proc. of SIGIR*.
- Ribeiro, F.; Araújo, M.; Gonçalves, P.; Benevenuto, F.; and Gonçalves, M. 2015. A benchmark comparison of state-of-the-practice sentiment analysis methods. *arXiv preprint arXiv:1512.01818*.
- Socher, R.; Perelygin, A.; Wu, J. Y.; Chuang, J.; Manning, C. D.; Ng, A. Y.; and Potts, C. 2013. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proc. of EMNLP*.
- Taboada, M.; Brooke, J.; Tofiloski, M.; Voll, K.; and Stede, M. Lexicon-based methods for sentiment analysis. *Comput. Linguist.*
- Wilson, T.; Wiebe, J.; and Hoffmann, P. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proc. ACL*.

<sup>2</sup><https://www.meteor.com/>

<sup>3</sup><https://tech.yandex.com/translate/>