

Sentiment Classification for Bengali News Articles

How to run this

Environment

Dictionary

Classifier

What is Available

(Unannotated) Dataset

Bengali-English word-level Dictionary

Classifiers

Positive-Negative English Wordlists

Notes

Paper References

Sentiment Classification for Bengali News Articles

How to run this

Environment

The project is implemented as a Jupyter Notebook. Developed on `Anaconda 4.6.12`.

`conda list` output is given in [this file](#)

Dictionary

The dictionary is already saved as a `.pickle` object. If you want to add to it and improve it, then run [this notebook](#)

Classifier

Run [this notebook](#) for the classifier(s)

What is Available

(Unannotated) Dataset

Data from [Kaggle](#)

Formatted properly as JSON. Files available for

1. [Anandabazar](#)
2. [Ebala](#)

Bengali-English word-level Dictionary

Combining scraped results from Google Translate, and [an existing project by Minhas Kamal](#)

Classifiers

The classifier is available [here](#). It is a notebook, and the report is included.

Positive-Negative English Wordlists

Using wordlists compiled [here](#) and [here](#). I added some words to the lists for completeness.

Notes

1. The [test directory](#) is a compilation of code I was writing with a different directory structure. *None of that code is guaranteed to work.*
2. [Slides used for presentation are available.](#)

Paper References

1. Das, Amitava, and Sivaji Bandyopadhyay. "Subjectivity detection in english and bengali: A crf-based approach." *Proceeding of ICON* (2009).
2. Liu, Bing, Minqing Hu, and Junsheng Cheng. "Opinion observer: analyzing and comparing opinions on the web." *Proceedings of the 14th international conference on World Wide Web*. ACM, 2005.