Abstract

key words: computational linguistics, machine learning, user generated content, semantic web, argument web vision, argument mining, opinionated claim detection, sentiment.

**Argumentation mining** is the process of extracting **structured argument data** from **raw natural text**.

This thesis aims to explore the **benefits** of incorporating **opinion mining** **and subjectivity** **detection features/techniques** in **claim detection**, one of **argumentation mining sub-tasks**.

We aim to **design** of a **semi-supervised** **argumentation mining framework** that enables feature inspection on **opinion claim detection** from **user comments** in relation to a **debate topic**.

This will be done by exploiting **natural language processing** and **machine learning techniques** starting at the smallest granularity possible to retain enough **argumentative features** and finding leverage to expand both in **pragmatic** and **semantic comprehension**.

In order to bring the most out of both human and machine intelligence interactive methods will be applied so the results of classifier combination can be more transparent from a human perspective.

instead reference the possible datasets to be used in experiences

[**MPQA Opinion Corpus**](http://mpqa.cs.pitt.edu/corpora/mpqa_corpus/)

[**IBM Debating Techbnologies**](https://www.research.ibm.com/haifa/dept/vst/mlta_data.shtml)

[Internet Argument Corpus](https://nlds.soe.ucsc.edu/iac2)

debates

ArgMine

The corpora used is manually annotated and consists in opinion authored articles and comments in the public area from a Portuguese online newspaper.