

Processamento de Linguagem Natural

[Painel do utilizador](#)[As minhas unidades curriculares](#)[Processamento de Linguagem Natural](#)[Assignments](#)[Assignment 1](#)

Assignment 1

In the scope of the [DARGMINTS](#) project, an annotation project was carried out which consisted of annotating argumentation structures in opinion articles published in the [Público](#) newspaper. The annotation included several layers:

1. Selecting text spans that are taken to have an argumentative role (either as premises or conclusions of arguments) -- these are Argumentative Discourse Units (ADU).
2. Connecting such ADUs through support or attack relations.
3. Classifying the propositional content of ADUs as propositions of fact, propositions of value, or propositions of policy; within propositions of value, distinguish between those with a positive (+) or negative (-) connotation.

In a proposition of fact, the content corresponds to a piece of information that can be checked for truthness. This does not usually happen with propositions of value, which denote value judgments with a strong subjective nature; often, they also have a (positive or negative) polarity attached. A proposition of policy prescribes or suggests a certain line of action, often mentioning the agents or entities that are capable of carrying out such policies.

The aim of this assignment is to build a classifier of types of ADUs, thus focusing on the last annotation step described above. For that, you have access to two different files:

- A file containing the content of each annotated ADU span and its 5-class classification: **Value**, **Value(+)**, **Value(-)**, **Fact**, or **Policy**. For each ADU, we also know the annotator and the document from which it has been taken.
- A file containing details for each opinion article that has been annotated, including the full article content.

Besides ADU contents, you can make use of any contextual information provided in the corresponding opinion article.

Each opinion article has been annotated by 3 different annotators. For that reason, you will find in the ADU file an indication of which annotator has obtained the ADU. It may happen that the same ADU has been annotated by more than one annotator. When that is the case, they do not necessarily agree on the type of proposition.

How good a classifier (or set of classifiers) can you get? Don't forget to properly split the dataset in a sensible manner, so that you have a proper test set. Start by obtaining an arbitrary baseline, against which you can then compare your improvements.

Última alteração: quarta, 16 de março de 2022 às 23:48

[◀ Text classification](#)[Opinion Articles dataset ▶](#)

Tecnologias Educativas - 20 anos na U.Porto

[Creative Commons](#)

[Requisitos mínimos utilização](#)

[Portal de e-learning](#)

[Ajuda Moodle](#)

[Inovação Pedagógica](#)

Nome de utilizador: Rúben Filipe Seabra de Almeida (Sair)

[FEUP-M.EIC022-2021/2022-2S](#)

[Português \(pt\)](#)

[Deutsch \(de\)](#)

[English \(en\)](#)

[Français \(fr\)](#)

[Português \(pt\)](#)

[Obter a Aplicação móvel](#)