

Information Retrieval Assignment Proposal

Konstantina Ellina - s0230419

Nada Chehbouni - s0233173

Pablo de Vicente - s0233489

Scope:

Our work will focus on the second proposed assignment “Fact verification”. We will try to implement a verification system that, given a claim, predicts whether the claim is Factual, Untrue or Unsubstantiated. These terms are defined as follows:

- **Factual:** Something that is factual is based on facts and can be proven to be true. It is information or statements that are accurate and can be verified. Factual information is objective and not influenced by personal opinions or beliefs.
- **Untrue:** Something that is untrue is not based on facts and is not accurate. It is a statement or information that is false, incorrect, or misleading. Untrue information is not supported by evidence or reality.
- **Unsubstantiated:** Something that is unsubstantiated lacks sufficient evidence or proof to support its validity. It means that a claim or statement has been made, but there is no substantial or verifiable evidence to confirm its truth or accuracy. Unsubstantiated claims may be true or untrue, but they cannot be confirmed due to the lack of supporting evidence.

As suggested in the proposed assignment, we will take a claim as an input and return a prediction along with documents that support or refute that claim.

Resources:

We will be using the **FEVER** dataset [[FEVER Dataset - Fact Extraction and VERification](#)] which consists of 185,445 claims generated by altering sentences extracted from Wikipedia and subsequently verified without knowledge of the sentence they were derived from. The claims are classified as *Supported*, *Refuted* or *NotEnoughInfo*. We are willing to use Python as a language as well as incorporate some language processing libraries tailored for our needs as we're progressing through the assignment.

Literature:

We plan on reading research papers on fact verification and claim detection (like the ones available in the fever website*) to gather more knowledge about the subject. We believe that this can provide us with valuable information about how to proceed with regard to the problem. As we will use machine learning models for the implementation, we are planning to read ML and NLP papers to help us understand these concepts and their application for fact checking and verification. We will also heavily rely on the class material provided to us to

understand the available techniques and decide which ones would be most beneficial for us to use.

*[Fact Extraction and VERification \(fever.ai\)](https://fever.ai)

Evaluation:

We plan to use the same dataset to evaluate our results. By the end of the implementation phase, we'll have settled for the evaluation metrics that would work better for us. But as a first attempt, we plan on comparing true positives and false positives, using ROC curve and AUC score to understand the performance of the system. If these are not enough, we will use some extra evaluation metrics. In case of errors, we will try to analyze them in order to better understand the flaws of our fact checker and work on solving them, or minimizing their impact.