# Assignment 2

Mirco Lescart, Freddy Fernandes, Parsa Mastouri Kashani
Arina Sadeghi Khiabanian
Master's Degree in Artificial Intelligence, University of Bologna
{ mirco.lescart, freddy.fernandes, parsa.mastouri, arina.sadeghi }@studio.unibo.it

January 13, 2024

## Abstract

This report addresses the Human Value Detection challenge, classifying arguments based on conveyed human values in terms of stance, premise, and conclusion. Tasks involve analyzing structured arguments and categorizing them based on stance, like 'in favor of' or 'against.' The approach includes baseline models (random uniform, majority classifiers) and advanced BERT-based classifiers. These classifiers are tested in three configurations: using only the conclusion (BERT w/ C), the conclusion and premise (BERT w/ CP) and adding stance to conclusion and premise .Models are evaluated using F1-score providing a thorough comparison of BERT-based models against baseline classifiers in human value detection.

## 1 Introduction

Our primary task is the classification of arguments based on stance, premise, and conclusion, categorizing them into four Level 3 human value categories: Openness to Change, Self-Enhancement, Conservation, and Self-Transcendence. We focus on analyzing arguments. We employ both basic classifiers, including random uniform and majority classifiers, and advanced techniques using the BERT model. One of the models is uniquely adapted to analyze not just the conclusion, but also the premise, and to incorporate the stance of the argument. Through this comprehensive approach, we aim to effectively capture the complexities of language and the underlying values in argumentative texts.

## 2 System description

In this assignment, Google Colab was utilized as the primary platform for both programming and collaborative work, enabling efficient teamwork. The coding and architectural design of the system were developed through a combined effort of the team, with each member working on and improving various tasks within the assignment. We started with basic

classifiers: a random uniform classifier that randomly predicts labels, and a majority classifier that always predicts the most frequent label in the training set. Building upon these, we implemented advanced models using the 'roberta-base' BERT architecture. Our custom BERT-based classifiers, enhanced with additional dropout and dense layers, were adapted to analyze different components of arguments — their conclusions, premises and stances.

## 3 Experimental setup and results

In our project, we used sklearn's DummyClassifier as a baseline to establish fundamental performance metrics. We used uniform stratergy, and also most frequent stratergy. This simple approach offered a benchmark for comparing more advanced techniques. Further, We utilized the BERT model from the transformers library, a sophisticated tool capable of understanding complex language nuances. For evaluation metrics, we integrated functions from the sklearn library, focusing on precision, recall, and the F1-score.we Tried three different models and the one with conclusion premise and stance work better.

## 4 Discussion

for our problem just running two epochs works well which was surprising although understandable given that BERT is already pre-trained and also it doesn't have that many parameters compared to other transformer models. we also made an effort to make a rather unconventional model by feeding one transformer the conclusion and another one for premise and getting 768 heads from each of them and concatenating the stance with both transformer heads and fed it to multiple layers of dense layer with RELU which gave us good results but was too expensive computationally and memory wise..

## 5 Conclusion

To conclude, our study successfully employed BERT-based classifiers to classify arguments based on stance, premise, and conclusion, outperforming baseline models. The comprehensive approach showcased the effectiveness of leveraging advanced models for nuanced human value detection in argumentative texts.

## 6 References

https://huggingface.co/
https://www.coursera.org/
https://towardsdatascience.com/bert-roberta-distilbert-xlnet-which-one-to-use-3d5ab82ba5f8