PROJECT PROPOSAL TEXT ANALYTICS

GROUP 1:

BOMBINO, FABBRI, RICCI, MASTRORILLI, SARBACH

GROUP MEMBERS PRESENTATION

- **Vincent Sarbach-Pulicani**:
  - Exchange Students in Digital Humanities between École nationale des chartes of Paris and Pisa
  - Master's degree in Late Modern history of the University of Strasbourg
- **Lucia Fabbri**:
  - Bachelor's Degree in Engineering and Computer Science, University of Bologna
  - Master's Degree in Data Science and Business Informatics, University of Pisa
- **Bianca Bombino**:
  - Bachelor's Degree in Computer Science, University of Pisa
  - Master's Degree in Data Science and Business Informatics, University of Pisa
- **Davide Ricci**:
  - Bachelor's Degree in Computer Science, University of Pisa
  - Master's Degree in Data Science and Business Informatics, University of Pisa
- **Alessandro Mastrorilli**:
  - Bachelor's Degree in Computer Science, University of Bari
  - Master's Degree in Data Science and Business Informatics, University of Pisa

MOTIVATION & ISSUES

- **Toxic Tweets Dataset**

- **WHY?**

  - Social Network Role in people's lives

  - Social haters

- https://www.kaggle.com/datasets/ashwiniyer176/toxic-tweets-dataset

- **54313** records

- ***0: non-toxic, 1: toxic*** : tweets classified by a binary code

- ***Tweets are very short messages.*** For this reason we need enough data to train our models.

- There is no a particular topic about the toxicity of the tweets, ***the corpus is varied***.

DATASET CHOICE

WHAT IS "TOXICITY"?

- **What is a toxic tweet?**

  - Insults, disrespect language..
  - It can also be caught in a specific context : retweets, answers,…

  - *For example*: a Tweeter user can qualify another one as a "cunt", one of the answer can agree to this without using an equivalent lexicon.

DATASET USE

- The **dataset** will be randomly divided into **2 sections**:

  - the **training set**, a set of examples used to train a supervised predictive model capable of determining the target value for new examples.
  - The trained model will be evaluated on a new set of examples, the **test set**, which are not used during training.

- **How to classify a tweet?**
  - Toxicity detector

  - **TOOLS**:

    - **Transformer**: it uses a new form of Neural Network

    - Algorithms for supervised learning contained in **sckitlearn framework** and then compared by k-fold cross validation.

SUPERVISED LEARNING TASK

- **How to classify different topics of tweets?**
  - Sentiment analyzer

  - **MODEL:**

    - **Vector Space Model:** to measure the similarity between the various vectors previously transformed to determine the nature of tweets.

    - A **graphic representation** by grouping the various results obtained, in several cluster. => A **better distinction of categories** to which tweets belong.

UNSUPERVISED LEARNING TASK

TOXICITY TYPES IDENTIFICATION

- **Types of toxicity expected**:
  - Gender inequalities
  - Xenophobia
  - Cyberbullying

How to identify them?

- using topic modeling methods
- *IDEA*: to detect latent topics from our corpus. To go further, we can identify the types of insults used according to a certain keyword ("women", "blacklivesmatter", ..)

IMPLEMENTATION & EVALUATION

Implementation:

Python

Jupyter Notebook

Evaluation:

Extrinsic Evaluation

THANKS FOR YOUR ATTENTION!