* **Code**

**2 dataset:**

**old : can, random forest, zero-shot. Fine-tune a transformer**

**[new - larger dataset 1.8million, 0—1; train some models, save the models for presentation**

**train some models, save the weights, use for presentation**

**Linear, Random forest, cnn, fine-tune a transformer]**

**fine-tune a transformer**

**Linear**

**Cnn: batch normalization, dropout**

**Lstm**

**transformer : learning rate**

* **Outline** 
  + **Old dataset: models, accuracy, classification report, visuals for the code report**
    - **Model: linear, cnn and zero-shot**
  + **New dataset: Jigsaw**
    - **Models: mse, linear, cnn, fine-tune a transformer on this**
    - **UI -**

**1. Introduction (.0.5 - 1 minute):**

Introduce the main problem(s) that your project addresses. Clearly state the questions or issues of the analysis that you would address in the rest of the presentation.

**2. Analysis and Findings (6-7.5 minutes):**

Present a comprehensive analysis of the data you studied during the project. Describe the NLP methodology and techniques employed. Summarize your findings and highlight the most important results. Use visuals like graphs, charts, and tables to support your conclusions and make your presentation storytelling.

Data: ? 0.5

Preprocessing: 0.5. tokenized, embedding [word2vec, fastest, glove]

Model: 1

old: Linear, Random Forest, CNN, zero-shot - accuracy, classification report

new: Linear, CNN, fine-tuned transformer - mse

Present the UI: 2min

different models: 3 - CNN, Linear, transformer - new dataset

\* texts:

**3. Limitations and Lessons Learned (2):**

Acknowledge your analysis's limitations, such as data restrictions, time constraints, or methodological challenges. Share the key lessons you have learned throughout your project regarding the subject matter and the research process.

* **Presentation**
* **Report**