



# WasChatted

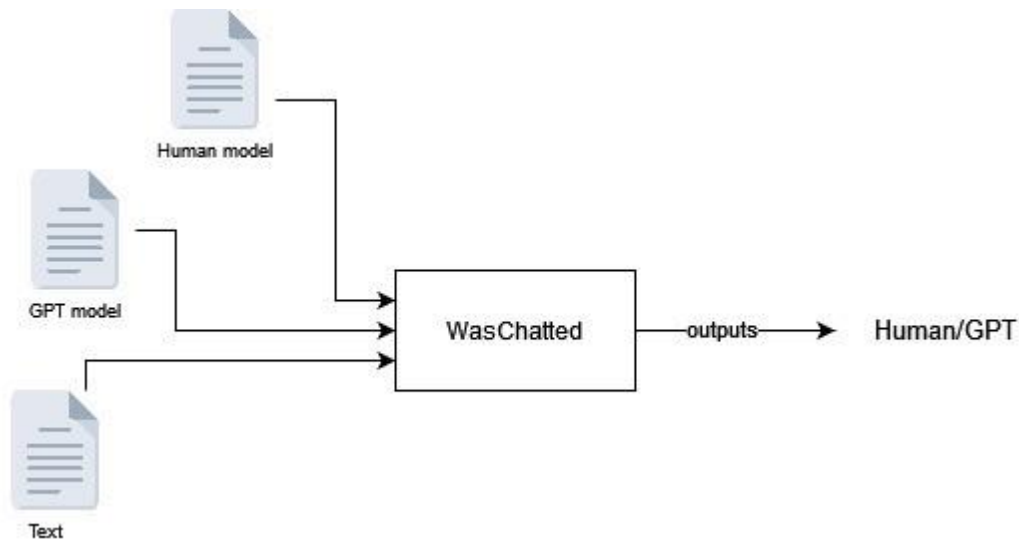
10 May

Pompeu Costa, 103294  
Ricardo Antunes, 98275  
Rafael Pinto, 103379

TAI 23/24  
Group 4

# What is WasChatted

WasChatted is a program to determine if a text was written by a human or ChatGPT



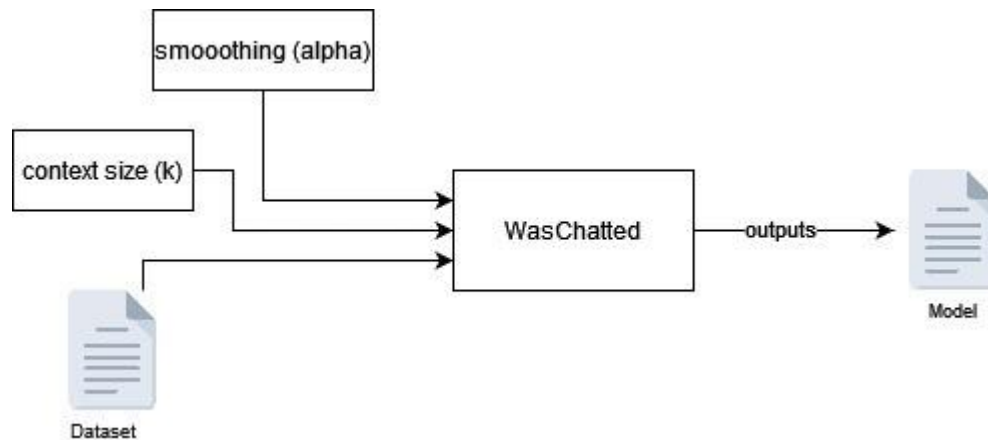


# Functionalities

# Train

Takes a dataset, context size and smoothing to train the model.

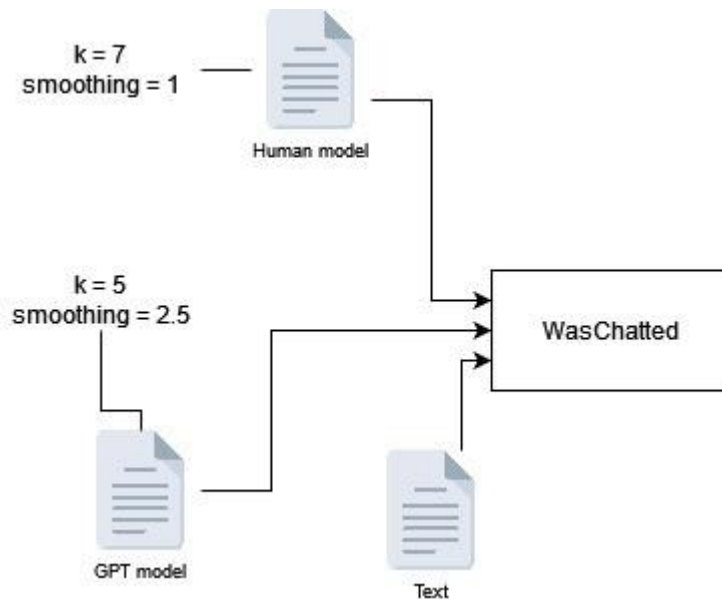
The model is then written on a binary file.



# Analyse

Use the models to predict who wrote the provided text.

Models don't need to have been trained with the same configurations.





# Dataset



# Dataset

- From huggingface.co
- Source of texts: Wikipedia introductions
- Transformation process: Rewritten by GPT-3 model
- 90% for training and 10% for testing



# Results



# Results

F1 Score

Window Size (K)

K = 3 => 0.82

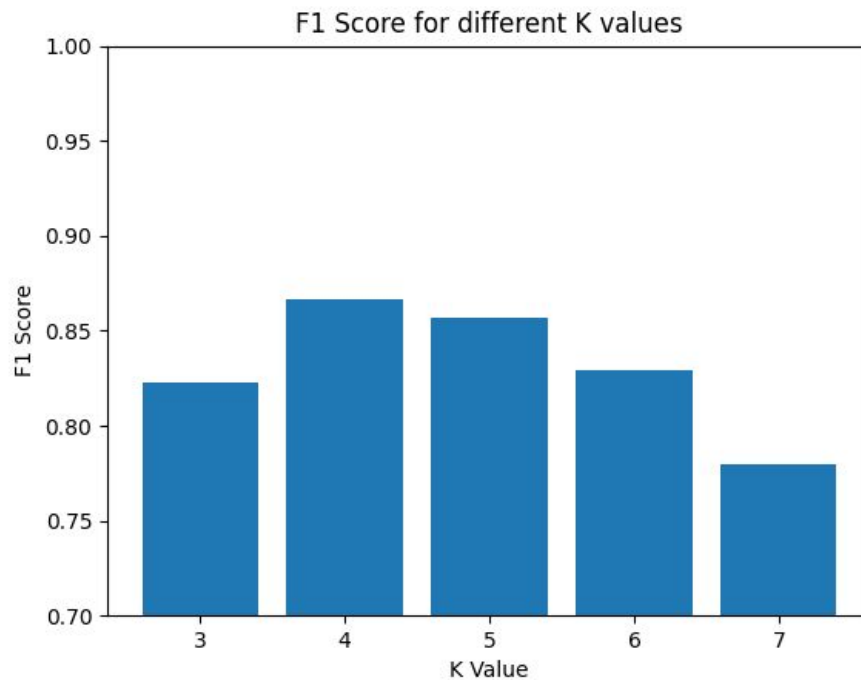
K = 4 => 0.87

K = 5 => 0.86

K = 6 => 0.82

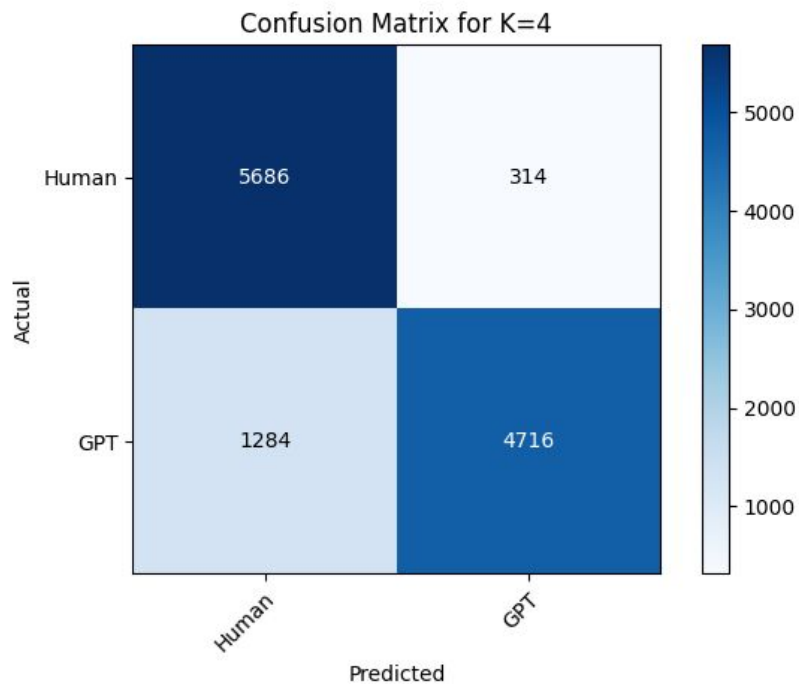
K = 7 => 0.78

Tested for 800 texts



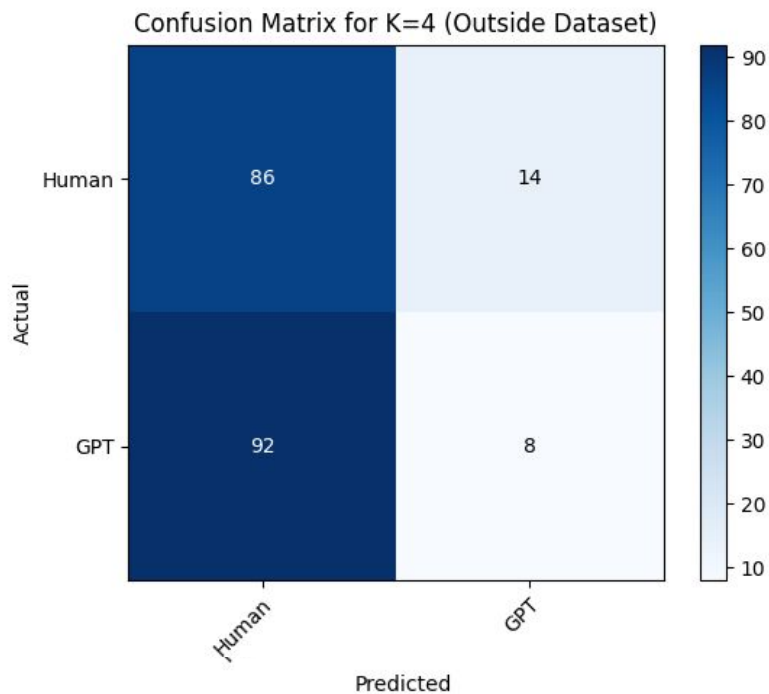
# Results

F1 Score = 0.87



# Results

F1 Score = 0.47





## Conclusion

- The program developed works as expected and implements all the proposed features
- Relatively good performance identifying class of a text (F1 Score: 0.87)
- Could be improved with training data from GPT3.5