Phishing email detection through linguistic patterns and sentiment analysis

Student: Daniel Pedrinho Nº 107378

Supervisor: João Rafael Duarte de Almeida

Co-supervisors: Sérgio Guilherme Aleixo de Matos

Summary: With the wide usage of e-mail as a communication tool, phishing attacks have become increasingly common and sophisticated. This dissertation aims to explore **the use of linguistic patterns** and **sentiment analysis to detect phishing emails**. By analyzing the emotional tone and language used in emails, we may be able to identify potential phishing attempts and improve email security.

Objectives

- Objectives:
  - Review State of the Art
  - Research technologies to investigate and use
  - Review existing datasets
  - Use LLM to generate synthetic dataset
  - Use collaborative annotation platform to annotate the dataset
  - Develop initial validation pipeline
  - Validate the annotated dataset

Work done / Results

- Work Done:
  - Generated the synthetic dataset
  - Collaborative annotation process
  - Development of validation pipeline

- Results:
  - Fully annotated synthetic dataset
  - Validation Pipeline
  - Preliminary Results:
    - Phishing Accuracy: 0.815
    - Emotion F1 Score: 0.539

Next Steps / Challenges / Bottlenecks

- Next Steps:
  - Execute work plan proposal
  - Develop Phishing Detection Models
  - Develop Emotion Detection Models