**Practical Machine Learning and Deep Learning - Assignment 1**

**Text De-toxification via ConBEGPT model.**

Final Solution Report

Parepko Leon

**Introduction.**

The task of automatic rewriting of offensive content attracted less attention, yet it may find various useful applications such as making online world a better place by suggesting to a user posting a more neutral version of an emotional comment.

The previous works presented a Conditional BERT architecture to solve such problem. It shows quite efficient results but not with cons. It is assumed, that it is possible to make the results better with noting all the disadvantages and negative phenomenas we observed in the behaviour of already existing approach. We define 2 key observations to fix:

* Semantic quality: the problem is that the Conditional BERT model struggles to comprehend indirectly implied phrases, which limits its effectiveness. The model lacks the ability to establish comprehensive multiple-to-one token connections.

This project presents a proposed solution to the problem of text de-toxification by leveraging the capabilities of Conditional BERT (Con. BERT) and GPT-2 models. The ConBEGPT network architecture is utilized to improve the quality and indirectly transform text generated by Con. BERT.

**Data analysis.**

**Model Specification.**

**Training process.**

**Evaluation.**

**Results.**