



Projeto final

O projeto final vai ter peso dois na nota e deve ser feito no máximo em trios. Ou seja, também pode ser feito em dupla ou de forma individual. O grupo deve acessar o link abaixo, o qual mostra a implementação e treinamento de uma arquitetura *transformer* usando o método semi-supervisionado de treinamento conhecido como MLM (*masked language modeling*).

https://keras.io/examples/nlp/masked_language_modeling/#create-bert-model-pretraining-model-for-masked-language-modeling

Os seguintes passos devem ser realizados:

- Colocar o código para rodar, fazendo todos os comentários necessários para o entendimento do mesmo.
- O código foi refinado para o dataset do IMDB. O grupo deve adaptar o código para outro dataset. Usem a criatividade para encontrar um problema interessante. Lembrando que existem diversos repositórios com dados para machine learning, como o Kaggle.
- Tentar solucionar o mesmo problema usando Markov. Comparar os resultados e discutir a análise.

A entrega do grupo será apenas o notebook com a implementação. Tal notebook deve conter qualquer discussão, comentários ou figuras que ajudem no entendimento do que foi feito.

A apresentação vai ser feita grupo a grupo na sala 206 de acordo com o agendamento a ser feito por sorteio. Para realização do sorteio, o grupo deve responder a tarefa do SIGAA indicando o nome dos componentes do grupo. Apesar do projeto ser feito em grupo, a nova vai ser individual uma vez que será feita uma “chamada oral” com cada membro. A data para o envio do notebook é dia **10 de abril até as 23:59** via SIGAA. Ou seja, vocês têm 20 dias para fazer o projeto!!!