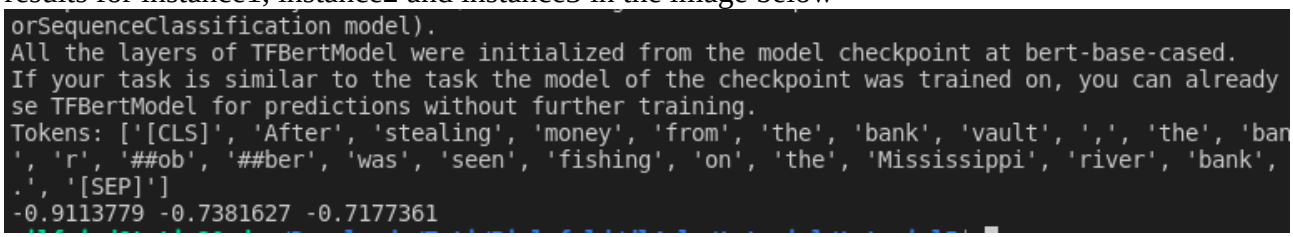
# DL4NLP 2022-Exercise 5

Task 1: Masked Language Modeling

a) Bert thinks that the DL4NLP lecture is pretty good as this sentence show a highest score (0.27)
b) Gender bias: we can see that talking about a nurser will show 0.81 score a women and when it comes to a president it will show 0.96 score a man
c) We have to instantiate the pre-trained BertTokenizer to manage all the preprocessing the pretrained model will expect and for this specific case it will retrieve the mask token for the model to decode

Task 2 : Contextual Word Embeddings

We determine how similar the three instances of the word ‘bank’ are to each other by getting the results for instance1, instance2 and instance3 in the image below




Task 3: Multilingual Bert (Code)

Task 4 :m Bert for sentiment classification (code)

# DL4NLP 2022-Exercise 8

Task 1: RNN Extensions

a) The benefit of bidirectional RNNs over unidirectional RNNs is that the input flows in both directions, and it’s capable of utilizing information from both sides. It’s also a powerful tool for modeling the sequential dependencies between words and phrases in both directions of the sequence.
   Example: Bidirectional LSTM (BiLSTM)
b) The benefit of adding “output connections” to RNNs is that its correspond to the last time step, containing information about the entire input sequence.

Task 2: Named-entity recognition with LSTMs (code)

Task 3: Named-entity recognition with Transformers

a) As a results here we are getting very low score (0.03 f1 score) which can be explain due to the fact that the training loop is incomplete . The datasets have been not fitted to the model and the model has been not evaluated with the evaluation datasets.
b) After fitting the datasets to the model we can observe a huge change on the results (0.97 of accuracy) which can be explain with the fact that after the building the model, it has been trained and evaluated with the transformed datasets and the required parameters.