

**YΣ19 Artificial Intelligence II (Deep Learning for
Natural Language Processing)
Fall Semester 2021-2022
Homework 4
25% of the course mark
Announced: January 30, 2022 Due: March 6, 2022 at
23:59**

1. Develop a sentiment classifier with 3 classes (pro-vax, anti-vax, neutral) by fine-tuning the pretrained BERT-base model available on Hugging Face.¹ You have to compute precision, recall and F1 for each class.

(3/10 marks)

2. In this and the following question you will develop a model for question answering on the dataset SQuAD 2.0. Start by visiting the Web site <https://rajpurkar.github.io/SQuAD-explorer/>, understanding the dataset and reading the relevant paper available. Your job is to fine-tune BERT-base for this task. For a general introduction to question answering see the slides and YouTube video from the CS224N Stanford class of 2021, lecture “Question Answering”.²

(4/10 marks)

3. In this question you will read the paper “What do Models Learn from Question Answering Datasets?”³ and you will reproduce some of the results of that paper. In particular, you have to reproduce the results of Table 3 (at least the rows and columns for SQuAD and TriviaQA).

(3-6/10 marks depending on how many datasets from Table 3 you will cover)

Note: Your solutions should be implemented in PyTorch. You should hand in: (i) a pdf with a detailed report of your solution for the two questions, including an explanation of the methods you used, an evaluation and comparison of different approaches you tried and citations to relevant literature that you might have used in developing your solutions. (ii) Colab notebooks (ipynb files using <https://colab.research.google.com/>) containing your code. You should use Python 3.6 or a later version.

¹<https://huggingface.co/models>

²<https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1214/>

³<https://arxiv.org/pdf/2004.03490.pdf>