

# Deep Learning for NLP

Student name: *<first-name last-name>*

*sdi: <sdiYYZZZZZ>*

---

Course: *Artificial Intelligence II (M138, M226, M262, M325)*

Semester: *Fall Semester 2023*

## Contents

## 1. Abstract

<Briefly describe what's the task and how you will tackle it.>

## 2. Data processing and analysis

### 2.1. Pre-processing

<In this step, you should describe and comment on the methods that you used for data cleaning and pre-processing. In ML and AI applications, this is the initial and really important step.

For example some data cleaning techniques are: Dropping small sentences; Remove links; Remove list symbols and other uni-codes.>

### 2.2. Analysis

<In this step, you should also try to visualize the data and their statistics (e.g., word clouds, tokens frequency, etc). So the processing should be done in parallel with an analysis. >

### 2.3. Data partitioning for train, test and validation

<Describe how you partitioned the dataset and why you selected these ratios>

### 2.4. Vectorization

<Explain the technique used for vectorization>

## 3. Algorithms and Experiments

### 3.1. Experiments

<Describe how you faced this problem. For example, you can start by describing a first brute-force run and afterwards showcase techniques that you experimented with. **Caution:** we want/need to see your experiments here either they increased or decreased scores. At the same time you should comment and try to explain why an experiment failed or succeeded. You can also provide plots (e.g., ROC curves, Learning-curves, Confusion matrices, etc) showing the results of your experiment. Some techniques you can try for experiments are cross-validation, data regularization, dimension reduction, batch/partition size configuration, data pre-processing from 2.1, gradient descent>

#### 3.1.1. Table of trials.

Trial				Score

Table 1: Trials

### 3.2. Hyper-parameter tuning

<Describe the results and how you configured the model. What happens with under-over-fitting??>

### 3.3. Optimization techniques

<Describe the optimization techniques you tried. Like optimization frameworks you used.>

### 3.4. Evaluation

<How will you evaluate the predictions? Detail and explain the scores used (what's fscore?). Provide the results in a matrix/plots>

<Provide and comment diagrams and curves>

3.4.1. *ROC curve.*

3.4.2. *Learning Curve.*

3.4.3. *Confusion matrix.*

## 4. Results and Overall Analysis

### 4.1. Results Analysis

<Comment your results so far. Is this a good/bad performance? What was expected? Could you do more experiments? And if yes what would you try?>

<Provide and comment diagrams and curves>

4.1.1. *Best trial.* <Showcase best trial>

### 4.2. Comparison with the first project

<Use only for projects 2,3,4>

<Comment the results. Why the results are better/worse/the same?>

### 4.3. Comparison with the second project

<Use only for projects 3,4>

<Comment the results. Why the results are better/worse/the same?>

#### 4.4. Comparison with the third project

<Use only for project 4>

<Comment the results. Why the results are better/worse/the same?>

### 5. Bibliography

<You should link and cite whatever you used or "got inspired" from the web. Use the / cite command and add the paper/website/tutorial in refs.bib>

<Example of citing a source is like this:> [?] <More about bibtex>