Politecnico di Torino
Dipartimento di Automatica e Informatica
Deep NLP (01VIXSM)
Written Exam simulation

January 23rd, 2022

Name: _____

Surname: _____

Student ID: _____

Exam rules:

- The present exam consists of 6 pages (including this cover page) and 7 questions overall. Any inconsistencies/printing errors in the written exam content must be reported to the teacher *at the beginning* of the exam.

- Exam duration: *60 minutes*.

- Withdraw is allowed only *at the end* of the exam.

- The exam is *closed-book*. Electronic devices, mobile phones, smart watches, and extra papers (even blank papers) are *not allowed*.

- Closed-ended questions: cross the right answer (just one) at pag. 2. Wrong or missing answers to closed-ended questions will receive *no penalty*.

- Open questions: write your answers below the text of the question. If you need more space please use the last page (i.e., pag. 6) and/or the back side of the paper.

Evaluation grid

Question:	1	2	3	4	5	6	7	Total
Points:	1	1	1	1	4	6	6	20
Score:								

1. (1 point) What of the following statement holds **True** for the Transformer architecture?

   - The encoder consists of a stack of six layers. Each layer has two sub-layers. The first is a multi-head self-attention module, and the second is a simple, position-wise recurrent neural network.
   - The encoder consists of a stack of six layers. Each layer has two sub-layers. The first is a masked multi-head self-attention module, and the second is a simple, position-wise feed-forward neural network.
   - The encoder consists of a stack of two layers. The first one is based on a multi-head self-attention module, and the second is a simple, position-wise feed-forward neural network.
   - The encoder consists of a stack of two layers. The first one is based on a multi-head self-attention module, and the second is a simple, position-wise recurrent neural network.
   - None of the above.

2. (1 point) The FastText embedding model

   - considers n-grams in place of entire words.
   - considers n-grams beyond entire words.
   - Considers n-grams and syntactical dependencies instead of entire words.
   - Considers n-grams and syntactical dependencies beyond entire words.
   - None of the above.

3. (1 point) In a RASA chatbot architecture a story

   - consists of a sequence of entities, intents, or actions.
   - consists of a sequence of intents and actions.
   - consists in a conversation between a user and an AI assistant expressed in plain text.
   - consists of a set of decision rules applied on top the provided intents.
   - None of the above.

4. (1 point) MultiLingual BERT:

   - do not allow end-users to encode input text snippets longer than 512 tokens.
   - separately processes text snippets written in different languages.
   - does not support finetuning.
   - requires a step of language identification.
   - None of the above.

5. (4 points) Describe the **distributional hypothesis**

    1. State the initial hypothesis.

    2. Provide examples of applications in English.

    3. Enumerate at least two NLP models/techniques that rely on the aforesaid hypothesis.

6. (6 points) Describe the process of **Statistical Machine Translation (SMT)**

    1. State of general problem.

    2. Describe the pipeline.

    3. Compare SMT with the main alternative Machine Translation solutions.

7. (6 points) Describe the LSARank summarization algorithm (*Using Latent Semantic Analysis in Text Summarization and Summary Evaluation* by Josef Steinberger and Karel Jezek).

   1. Define the summarization problem addressed by the paper.

   2. Explain why LSA can be useful for summarization purposes.

   3. Describe the procedure used by LSARank to retrieve the most relevant sentence in the original documents to put into the summary.

This page is intentionally left blank to accommodate work that wouldn't fit elsewhere and/or scratch work.