UNIVERSITY OF EDINBURGH

COLLEGE OF SCIENCE AND ENGINEERING

SCHOOL OF INFORMATICS

INFR11157 NATURAL LANGUAGE UNDERSTANDING, GENERATION, AND MACHINE TRANSLATION

May 2020

13:00 to 15:00

INSTRUCTIONS TO CANDIDATES

1. **Note that ALL QUESTIONS ARE COMPULSORY.**
2. **DIFFERENT QUESTIONS MAY HAVE DIFFERENT NUMBERS OF TOTAL MARKS. Take note of this in allocating time to questions.**
3. **This is an OPEN BOOK examination.**

MSc Courses

Convener: V.Nagarajan
External Examiners: W.Knottenbelt, M.Dunlop, M.Niranjan, E.Vasilaki

THIS EXAMINATION WILL BE MARKED ANONYMOUSLY

1. **Tranformers**

   (a) Briefly explain the concept of *self-attention* as it is used in transformers. Restrict yourself to the simplest, parameterless case. Use formulas if necessary. [*2 marks*]

   (b) The full concept of self-attention uses parameter matrices called *query*, *key*, and *value*. With reference to your previous answer, explain how these are used to define a more flexible concept of self-attention. [*2 marks*]

   (c) Assume you want to build a simple, transformer-based model that performs *document classification:* It takes a short document as an input (e.g., a newspaper article) and assigns it a topic (e.g., sport, politics, culture). Would you use position embeddings as part of your model? Justify your answer. [*2 marks*]

2. **Contextualized Embeddings**

   Your task is to design a model that performs *semantic role labeling* (SRL). This task takes a sentence and determines the role each phrase plays in the semantics of the sentences. Here is an example:

The	police	officer	quickly	detained	the	suspect	at	the	scene
B-A0	I-A0	I-A0	O	V	B-A1	I-A1	B-AM	I-AM	I-AM

   Here, the semantic roles are A0 (agent), A1 (patient), AM (modifier), and V is the predicate these roles depend on. For simplicity, we will assume that there is exactly one predicate in every sentence. The beginning of the phrase with a given semantic role is indicated by B-, the inner part of that phrase as I-, and O marks other phrases, i.e., ones that are not assigned semantic roles.

   In the lectures, we introduced contextualized embeddings, using BERT as an example. In this question, you will build a model that performs semantic role labeling based on BERT. You have a corpus available that contains sentences annotated with semantic roles in the form of label sequences such as the one given in the example.

   (a) What are the input and the output of your model, and how would you represent them? [*1 mark*]

   (b) Assume you want to train your model by finetuning. Which modifications to BERT's architecture are required for this? How would you implement this modification? [*2 marks*]

   (c) Describe how the finetuning will work. Which parameters do you tune, and what is your training objective? [*2 marks*]

   (d) Now let's say you want to build an SRL model that uses features extracted from BERT, rather than finetuning BERT. What does your architecture look like now? [*2 marks*]

(e) In reality, semantic role labeling is more complicated than the example sentence above implies. In particular, a sentence can have more than one predicate, and each predicate has its own set of semantic roles associated with it. Roles such as A0 and A1 can therefore occur more than once in a single sentence. How would you modify your SRL model to deal with this? [*3 marks*]

3. **Evaluation**

(a) State three differences between recurrent neural networks and convolutional neural networks. [*3 marks*]

(b) Which *automatic metric* would you use to evaluate the performance of a neural network model for the following tasks? Please state for each metric what type of data it assumes as input, describe what it measures, and the rationale behind it.

i. a news summarizer

ii. a data-to-text baseball summary generator

iii. a semantic parser for the travel domain

[*3 marks*]

4. **Question Answering**

Question generation is the task of automatically generating questions for a given answer. The task has many practical applications, such as generating training data for question answering systems, synthesizing frequently asked questions, and automatic tutoring systems.

Assume you are a software developer working for Google, and your manager asks you to build a question generation model. In order to train this model you are given a corpus which consists of instances of text passages and questions. An example is given below. You can also assume that the answer is known and is highlighted in boldface in the passage as shown.

Text Passage
In 1066, duke William II of Normandy conquered England killing king Harold II the battle of Hastings. **The invading Normans and their descendants** replaced the Anglo-Saxons as the ruling class of England
Question
Who conquered king Harold II at the battle of Hastings?

(a) Assume you want to approach question generation as a conditional language modeling problem. How will you express this formally? Provide the appropriate equations and explain why your formulation is a good approximation of the task. [*2 marks*]

(b) Now that you have formulated the generation task, you decide to use neural networks for the estimation of various model quantities following an encoder-decoder modeling paradigm. Please describe how your encoder and decoder would look like providing the appropriate equations where necessary. Bear in mind that you might want to encode the paragraph, and the sentence containing the answer to the question separately. [*4 marks*]

(c) You next decide you wish to incorporate attention in your model? How would you calculate it and why is this a good idea? [*2 marks*]

(d) Can you now enhance your model with a copy mechanism? How would you implement it? Is it a good idea? Justify your answer. [*2 marks*]

(e) What is the training objective of your model? Provide the equation. [*2 marks*]

(f) Which automatic metric would you use to assess the performance of your model? Briefly discuss pros and cons. [*2 marks*]

(g) Finally, you wish to develop a question generation system for German but you have no training data in this language. Can you think of a way to bootstrap a question generation system for German using only English task-specific data? [*2 marks*]

5. **Ethics**

(a) What is dual use? Give an example in NLP. [*2 marks*]

(b) A bank asks you to help automate loan-making in branches across the UK. For training, it provides detailed records from a central London branch, consisting of text-based applications, and the amount of money that the bank loaned to each applicant — which is zero if the loan was denied. Name three sources of bias and explain how each arises in this scenario. [*4 marks*]

(c) A startup wants you to build a system to help automate university admissions, based on an application essay. They tell you: "AI-backed software could immediately score an applicant against a set of factors found to be a signifier of success based on past applicants." What are five ethics questions should you ask about this application of NLP, and how would you answer each of them? [*6 marks*]