Final Exam

601.467/667 Introduction to Human Language Technology

Fall 2021
Johns Hopkins University
Co-ordinator: Philipp Koehn

13 December 2021

Complete all questions.

Use additional paper if needed.

Time: 75 minutes.

Name of student: _____

Q1. Question Answering   *25 points*

1. Explain two ways in which the fields of Question Answering (QA) and Information Retrieval (IR) are different. (There are more than two possible answers; any two reasonable explanations will be sufficient.) [8pts]

2. QA systems can be classified by the type of questions they are designed to handle. Please provide an example question for each type listed below. [12pts]

   (a) Factoid question:

   (b) List question:

   (c) Definition question:

   (d) Opinion question:

3. The following is an example from the Winograd Schema Challenge. Please explain why this kind of question is difficult to solve using current technology, and what kind of breakthrough might be needed to tackle these kinds of questions. [5pts]

   Question: The trophy would not fit in the brown suitcase because it was too big. What was too big?

   A. The trophy

   B. The suitcase

Q2. Digital Humanities   *25 points*

1. Describe an advantage of employing unsupervised models for exploratory analysis [5 pts]

2. Explain how graph-convolutional networks extend the principles behind standard CNNs [10 pts]

3. Describe a common issue with humanistic primary sources, and how the use of structured representations (e.g. SQL, RDF) help mitigate it [10 pts]

Q3. Clinical NLP   *25 points*

1. What is the name of the law in the United States that protects sensitive patient health information from being disclosed without the patient's consent or knowledge? [6 pts]

2. Why are rule based methods popular in clinical NLP? (Select all answers that apply.) [6 pts]

    ☐ Rules are easily interpretable

    ☐ Federal regulation does not permit AI methods without FDA approval.

    ☐ Most useful text is contained in images.

    ☐ Doctors can often express medical knowledge as rules.

3. The task of clinical note segmentation refers to: (Select the best answer) [6 pts]

    ☐ Dividing a paragraph into sentences.

    ☐ Decomposing a biomedical term into constituent parts.

    ☐ Splitting a note into sections, such as history, medications, chief complaint.

    ☐ Dividing notes based on the date of the patient encounter.

4. Why is the task of negation important when processing clinical notes? [7 pts]

Q4. Ethical Problems   *25 points*

1. Why do we need to ensure that AI algorithms take into account the principle of non-maleficence? [10 pts]

2. How can federated learning help providing non-maleficent algorithms? [5 pts]

3. Explain why we need to have *Explicability* in AI applications? Which measures would you follow to have more explicable systems? [10 points]