Master Degree in Computer Engineering

Natural Language Processing
Final Exam

February 14th, 2025

1. [**2 points**] List some of the most common tasks involved in text normalization. Can you provide examples of tasks that should be applied in some fixed order?

2. [**7 points**] Consider the following term-context matrix, providing co-occurrence counts for the target words newspaper, job, coffee, table, and the context words restaurant, car, book, music.

	restaurant	car	book	music
newspaper	5	0	52	17
job	72	78	16	0
coffee	102	11	41	37
table	117	0	78	19

   (a) For all entries in the main diagonal, indicate how to compute the positive pointwise mutual information (PPMI). Use fractions and logarithms in your answers **without** computing these operators.

   (b) The rows of the PPMI term-context matrix can be used as sparse word embeddings. Discuss a method to obtain dense word embeddings from the PPMI term-context matrix.

3. [**2 points**] In the context of the estimation of the probabilities of an N-gram models, explain what is meant for bias/variance tradeoff.

4. [**5 points**] In the context of large language models (LLMs) answer the following questions.

   (a) Introduce the LLM classification discussed in lecture 6, consisting in encoder-only, decoder-only, and encoder-decoder architectures.

   (b) Provide and discuss examples (LLM or LM) for each of the architecture typologies in (a).

5. [**2 points**] Introduce the linguistic resource called WordNet. How is the notion of synset defined in WordNet? Provide an example.

6. [**6 points**] Consider the English sentence 'The agency permitted the stuff to review their documents under surveillance' along with the projective dependency tree consisting of the following unlabeled dependency relations.

head	agency	permitted	⟨ROOT⟩	stuff	permitted	review	permitted	documents	review	surveillance	review
dependent	The	agency	permitted	the	stuff	to	review	their	documents	under	surveillance

Answer the following questions.

(a) Draw a graphical representation of the dependency tree above, with arcs directed from the head to the dependent.

(b) Apply to the above tree the oracle presented in class to construct a sequence of training instances for the arc-standard parser.

7. [**7 points**] In the context of the task of machine reading, answer the following questions.

(a) Introduce the notions of factoid and non-factoid questions, providing a simple example.

(b) Introduce the notions of query, passage and span and the associated notation.

(c) Define the main equations of the neural model known as Stanford attentive reader, and explain how the model implements the attention mechanism.

8. [**2 points**] In the context of large language models, describe the mixture of experts architecture. What are the advantages of this architecture?