

## Redes Neuronales para Lenguaje Natural

Ensayo de prueba

### Consideraciones generales

- i) La prueba es sin material escrito.
- ii) Escriba nombre y C.I. en todas las hojas.
- iii) Numere todas las hojas.
- iv) En la primera hoja, indique el total de hojas.
- v) Comience cada ejercicio en una hoja nueva.
- vi) Utilice las hojas de un solo lado.
- vii) Entregue los ejercicios en orden.

Este ensayo de prueba contiene ejemplos de ejercicios de dificultad similar a los que se propondrán en la prueba del 27/11.

### Ejercicio 1

Considere la siguiente tarea de PLN: Dado el texto de un mail, se desea obtener una clasificación de si es spam o no. Se cuenta con un corpus de 5000 textos de mails, para los cuales se sabe si son spam o no, como los siguientes:

*Tenemos que planificar la reunión para la semana que viene. → NO SPAM*

*Soy un príncipe africano y necesito mover millones de dólares de mi país → SPAM*

*El informe quedó pronto, pero le falta que le agregues tu firma. Podés? → NO SPAM*

*Imperdible oportunidad de negocio gran negocio venta dinero → SPAM*

- a) Indique por lo menos dos arquitecturas de redes neuronales de las vistas en el curso que serían apropiadas para esta tarea.
- b) Diseñe una red para resolver esta tarea indicando entradas, salidas, capas utilizadas y funciones de activación.
- c) Describa cómo realizaría el entrenamiento de la red de la parte anterior, incluyendo cómo partiría los datos y cómo elegiría los mejores hiperparámetros.

### Ejercicio 2

Responda las siguientes preguntas sobre el modelo de red neuronal Transformer.

- a) ¿Qué ventaja, en cuanto a eficiencia, introduce el Transformer en comparación con una Red Neuronal Recurrente?
- b) Explique brevemente el concepto de self-attention causal.
- c) ¿Cuál es la principal diferencia entre BERT y GPT?

### Ejercicio 3

Elija solo una respuesta correcta para las siguientes preguntas de múltiple opción:

- 1) Skip-gram con negative sampling es:
  - a) Un modelo de lenguaje generativo del estilo de LLaMA2.
  - b) Una técnica para construir representaciones de las palabras mediante la reducción de dimensión de la matriz de conteos de n-gramas.
  - c) Una técnica para construir representaciones de las palabras a partir de ejemplos positivos y negativos de pares (palabra, contexto).
  - d) Una técnica para construir representaciones de palabras a partir de predecir la próxima palabra dadas las n anteriores.

- 2) Las arquitecturas tipo encoder-decoder:
- a) Son arquitecturas que solo permiten procesar una entrada de tamaño fijo.
  - b) Solo se utilizan para obtener una salida diferente para cada palabra procesada.
  - c) Pueden utilizar redes recurrentes o redes tipo transformer.
  - d) No permiten la utilización de word embeddings.
- 3) Los perceptrones multicapa:
- a) No pueden modelar un conjunto que no es separable linealmente (como el XOR)
  - b) Permiten tomar como entrada una secuencia procesando elemento a elemento
  - c) Solo pueden utilizarse en problemas de regresión
  - d) Están compuestos por capas de unidades completamente conectadas

#### **Ejercicio 4**

Indique si las siguientes afirmaciones son verdaderas o falsas. Justifique en ambos casos.

- a) Un modelo de lenguaje permite estimar la probabilidad de una secuencia de tokens.
- b) Los LLMs del estilo de LLaMA2 deben ser entrenados con corpus paralelos para poder realizar tareas de traducción automática.
- c) Una forma de evaluar colecciones de word embeddings es mediante tests de analogías o tests de similitud.