

Introducción a la Inteligencia Artificial

Actividad de la Clase 13: El Procesamiento de Lenguaje Natural y Prompting a modelos de IA Generativa.

Actividad Práctica: Ingeniería de *Prompts* con Fundamentos de PLN

Objetivo General

Demostrar cómo los conceptos de **Expresiones Regulares (ER)** y **Análisis Sintáctico/Lingüístico (spaCy)** son fundamentales para escribir *prompts* efectivos y controlar las salidas de los modelos GPT.

Ejercicio 1: Control de Formato mediante Patrones (ER en el *Prompt*)

Este ejercicio se centra en obligar a la IA a seguir una estructura de salida rígida, el dominio de las Expresiones Regulares.

- **Tarea:** Actuando como un **analista de datos**, debe procesar una lista de nombres completos y fechas, y devolver la información en un formato de identificador único.
- **Prompt Modelo Base (a completar):**

"Eres un Analista de Datos experto. Genera un identificador único para cada persona de la lista. El identificador debe seguir un **patrón estricto**: Las primeras 3 letras del apellido (mayúsculas), seguidas de un guion bajo, y luego el año de nacimiento (4 dígitos). La salida debe ser **solo** la lista de identificadores."

- **Lista de Entrada:**

"Ana García (1985), Pedro Sánchez (1992), Marta López (2001)"

- **Salida Esperada (Ejemplo):** GAR_1985, SAN_1992, LOP_2001

Ayuda para la Resolución (Foco en ER)

1. **Identifique el Patrón de Salida:** Piense en la estructura del patrón usando la sintaxis de ER, aunque no la escriba en el *prompt*. El patrón es `[A-Z]{3}_\d{4}`.

2. **Use Restricciones Claras:** Las frases clave para la IA son: "**patrón estricto**," "**solo** la lista," y la definición precisa de las partes (3 letras, guion bajo, 4 dígitos). Esto minimiza el texto conversacional de la IA.

Ejercicio 2: Definición de Rol y Dependencias (Sintaxis en el *Prompt*)

Este ejercicio utiliza la estructura sintáctica para definir el **rol** (sujeto) y la **función/relación** (dependencia) de la tarea, inspirándose en los principios de spaCy.

- **Tarea:** Asignar un rol con una voz específica y pedir una tarea que dependa de esa voz.
- **Prompt Modelo Base (a completar):**

"Actúa como un **crítico de cine estricto** con un tono formal y académico (Rol/Tono). **Evalúa y critica** la siguiente sinopsis de película (Tarea y Verbo Principal). Tu crítica debe enfocarse **exclusivamente en la coherencia de la trama y la verosimilitud de los personajes** (Dependencias)."

- **Sinopsis de Entrada:**

"Un cartero jubilado de 80 años descubre que su perro es un extraterrestre con la capacidad de teletransportar edificios enteros. El cartero decide usar este poder para robar bancos en Tokio y financiar su jubilación."

Ayuda para la Resolución (Foco en Análisis Sintáctico/Lingüístico)

1. **Defina el Sujeto:** El *prompt* define claramente el **sujeto/agente** de la acción (el Rol: *Crítico de cine estricto*).
2. **Use Modificadores/Dependencias Clave:** Las frases como "**enfocarse exclusivamente en...**" son los modificadores que indican las **dependencias** y restricciones. Esto es análogo a decirle a spaCy que se fije solo en el POS Tag de "sustantivos" o la dependencia "sujeto" en un análisis.
3. **Ajuste la Voz:** La instrucción del tono (**formal y académico**) guía el estilo lingüístico de la respuesta.

Ejercicio 3: Combinación de Coherencia y Estructura (*Chain-of-Thought* + ER)

Este ejercicio avanzado combina la necesidad de **razonamiento coherente** (propio de GPT, mejorado con CoT) con la necesidad de **estructura formal** (legado de ER y spaCy).

- **Tarea:** La IA debe actuar como un agente de *fact-checking*, primero razonando su verificación, y luego dando la respuesta final en un formato *estricto*.
- **Prompt Modelo Base (a completar):**

"Eres un Agente de Verificación de Datos de alta precisión. **(PASO 1) Piensa paso a paso** para analizar la verdad de la siguiente declaración. Muestra tu razonamiento identificando las fuentes o datos conocidos. **(PASO 2) Luego, y solo al final**, dame el resultado en el **formato de bandera** estricto: [VERDADERO/FALSO/ENGAÑOSO] - Fuente principal del error."

- **Declaración a Verificar:**

"La Torre Eiffel es el monumento más alto de Francia."

Ayuda para la Resolución (Foco en CoT + ER/Sintaxis)

1. **Imponer la Coherencia (Sintaxis/CoT):** La instrucción "**Piensa paso a paso**" es la técnica *Chain-of-Thought* (CoT). Esto obliga a la IA a seguir una estructura lógica (argumento-evidencia-conclusión) similar a cómo spaCy mapea las dependencias.
2. **Imponer el Patrón de Salida (ER):** La instrucción "dame el resultado en el **formato de bandera estricto**" seguido del patrón exacto, fuerza la estructura final. La mezcla de corchetes y guiones es un patrón claro de ER. (La respuesta correcta, por cierto, es **FALSO**, la más alta es la antena de la Torre Eiffel, o estructuras como el Viaducto de Millau o la Torre Montparnasse, dependiendo de la definición de "monumento")