### **📚 Resumen de la Clase 7 – Transformers y Self-Attention**

#### **🧩 Contenido Técnico**

**1. Self-Attention y Transformers**

* Se introdujo el mecanismo de **self-attention**, eje central de los Transformers.
* A diferencia de las RNNs, la atención no depende de la longitud de la secuencia → **no hay degradación del gradiente**.
* Cada vector de entrada se transforma en tres: **Query (Q), Key (K) y Value (V)** mediante matrices entrenables.
* La atención se calcula como producto interno entre Q y K (escalado y normalizado con softmax), y se usa para ponderar los V.
* El resultado: una nueva secuencia con representación contextual.

**2. Diccionario suave (soft dictionary)**

* Se explicó el mecanismo de atención como una especie de diccionario suave, donde una consulta (query) puede recuperar múltiples valores con distintos pesos, en vez de uno solo como en un diccionario tradicional.

**3. Multi-Head Attention**

* Se aplican varias atenciones en paralelo (multi-head) para capturar distintos tipos de relaciones (e.g. sujeto-verbo, adjetivo-sustantivo).
* Las salidas se concatenan y pasan por una **capa densa** para reducir la dimensionalidad y continuar el procesamiento.

**4. Arquitectura Transformer**

* El bloque básico tiene: Multi-Head Attention + conexión residual + normalización + feed-forward.
* Se introducen **positional embeddings** para que el modelo sea sensible al orden de las palabras.

**5. Encoder y Decoder**

* Transformer completo incluye un **encoder** (procesa la entrada) y un **decoder** (genera salida, e.g. traducción).
* El decoder usa **enmascaramiento** para no mirar información futura durante el entrenamiento.

**6. Segmentación de texto (WordPiece)**

* En lugar de usar caracteres o palabras completas, se introdujo **WordPiece** como técnica para generar tokens intermedios.
* Esto permite construir un vocabulario compacto y eficiente.

#### **📌 Organización del curso**

* **Es la última clase del curso.**
* Se recordaron los **4 desafíos ya entregados**.
* Se plantea un **desafío final**: preparar un *README* o *guía visual* del repositorio con sus proyectos, para presentar el trabajo hecho.
* El plazo límite es antes del **cierre de notas (posiblemente el 26)**.
* Se está evaluando si se suma un mini desafío más.