



Universidad Politécnica de Madrid

**Introducción a la inteligencia artificial generativa
a través de los grandes modelos de lenguaje**

Large Language Models (LLM)
Grandes Modelos del Lenguaje

Javier Conde
Pedro Reviriego

¿Qué es un LLM?

- Los LLM son modelos estadísticos que aprenden a predecir la siguiente palabra*
- El modelo estima las probabilidades de que cada palabra sea la siguiente

Al ver al ladrón el perro ...



Predicción

* En realidad el modelo predice tokens no palabras. Se verá más adelante en el curso

A partir de ese modelo podemos completar frases usando el modelo para ir obteniendo cada palabra de forma recursiva

Al ver al ladrón el perro ...

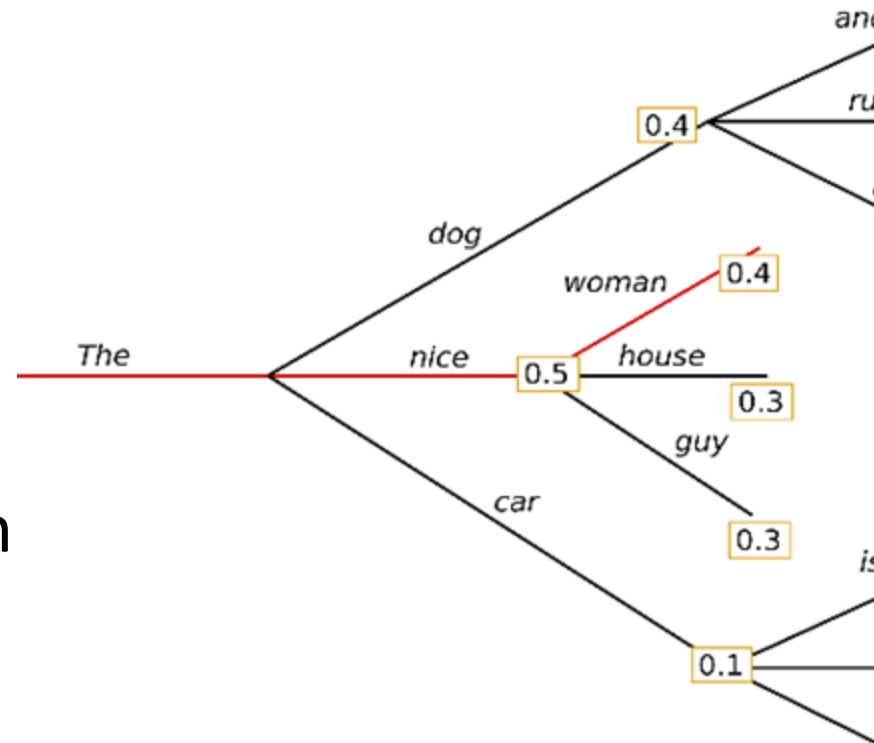
Al ver al ladrón el perro **ladra** ...

Al ver al ladrón el perro ladra **para** ...

Al ver al ladrón el perro ladra para **asustar** ...

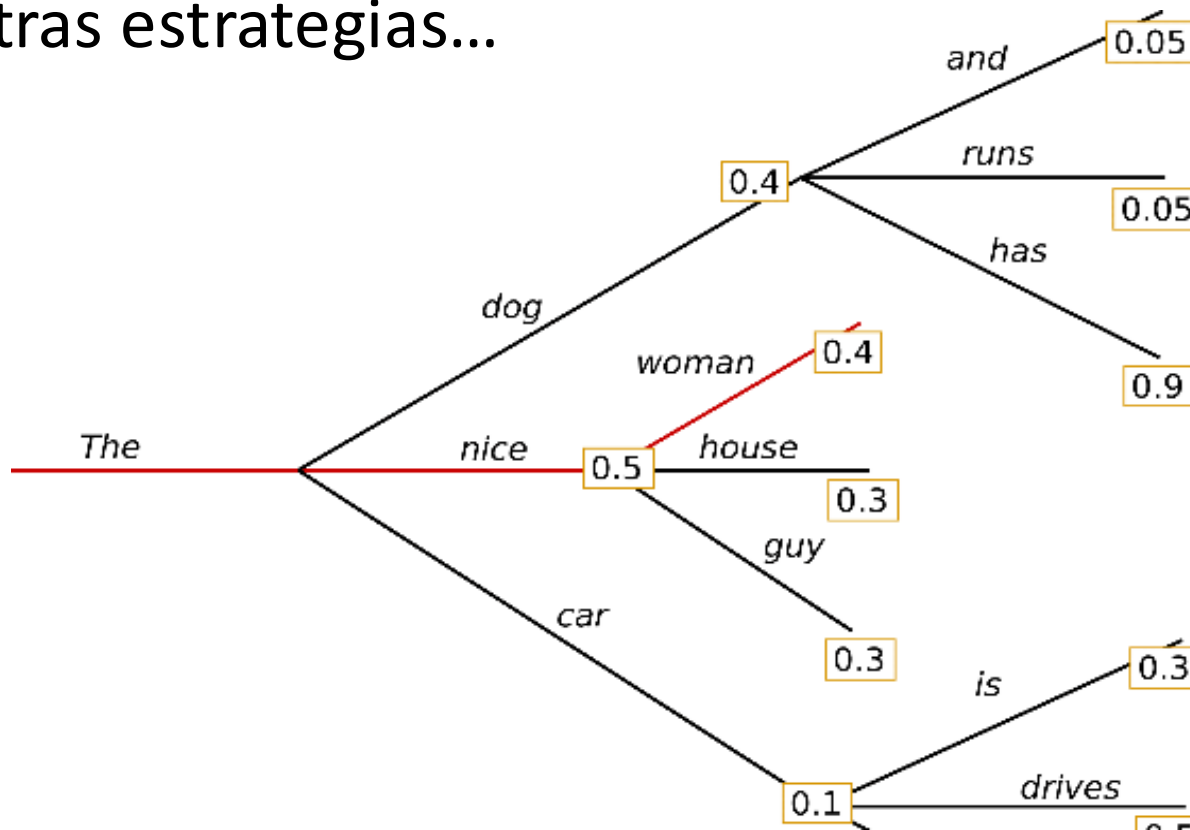
Árbol de probabilidades

- El modelo da una estimación de las probabilidades de que cada palabra sea la siguiente
- Se pueden seguir muchos caminos en función de cómo se elija la siguiente palabra



Posibles caminos

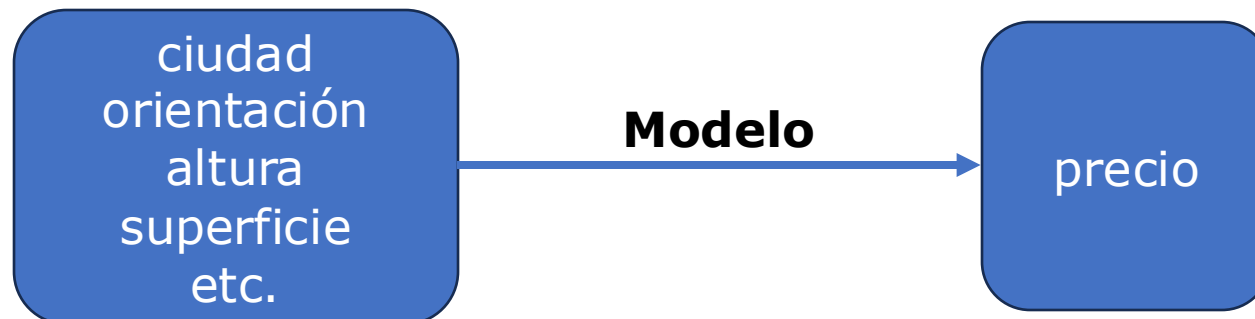
- Tomar siempre la de mayor probabilidad
- Elegir de modo proporcional a las probabilidades
- Otras estrategias...



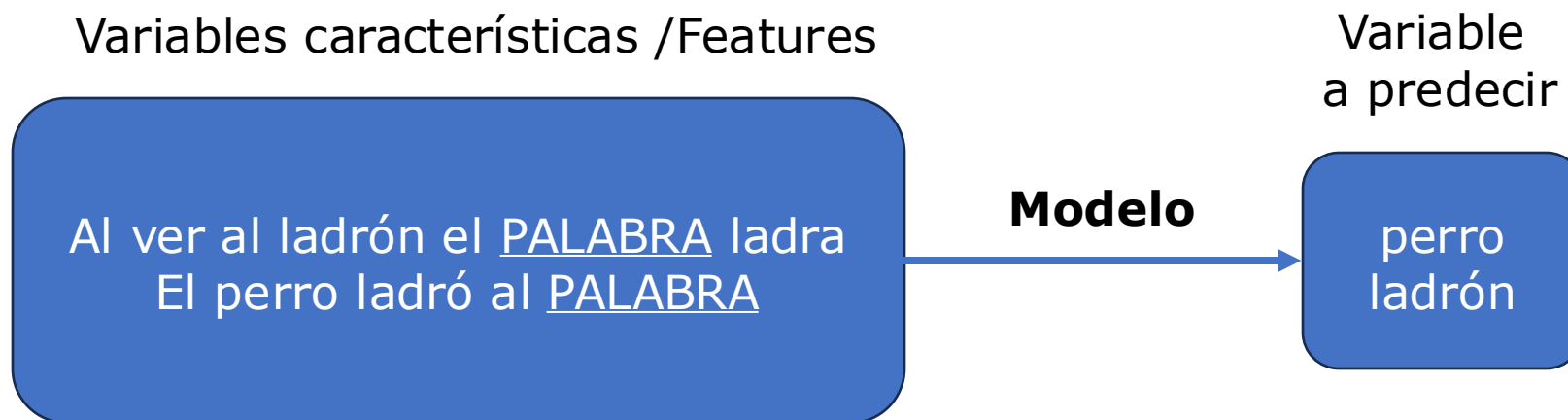
- Los LLM son modelos supervisados, es decir, necesitan un dataset de referencia con el que aprender y ajustar sus parámetros
- Generar un dataset de calidad es una tarea costosa en problemas de Machine Learning ya que se deben recolectar y etiquetar

Variables características /Features

Variable a predecir



- En los LLM es sencillo generar este dataset y automatizar su generación
 1. Toma un texto
 2. Elimina una palabra del texto
 3. La variable a predecir es la palabra eliminada
- Los LLM se entrenan con datasets muy grandes





Universidad Politécnica de Madrid

**Introducción a la inteligencia artificial generativa
a través de los grandes modelos de lenguaje**

Large Language Models (LLM)
Grandes Modelos del Lenguaje

Javier Conde
Pedro Reviriego