

Los modelos N-gramas nos ayudan a predecir la probabilidad de una secuencia de palabras, pero tienen el reto de asignar probabilidades a combinaciones que no aparecen en el conjunto de datos de entrenamiento. Estos modelos se utilizan como solución:

### - Absolute Discounting

Se trata de restar un valor fijo a las cuentas de los N-gramas observados y asignar esa probabilidad a los N-gramas que no hemos visto, en lugar de darles una probabilidad de cero. Esto hace que el modelo sea más preciso.

### - Kneser - Ney Smoothing

Este método no solo se fija en cuantas veces aparece un N-grama, sino también en cuantos contextos diferentes lo preceden. Así, le da más importancia a las palabras que se encuentran en más contextos variados.

Ambas técnicas ayudan a mejorar la estimación de probabilidades cuando tenemos N-gramas poco frecuentes o que no hemos visto antes.