

PER - Práctica 2

Alemany Ibor, Sergio
Galindo Jiménez, Carlos S.

1 Gráficas

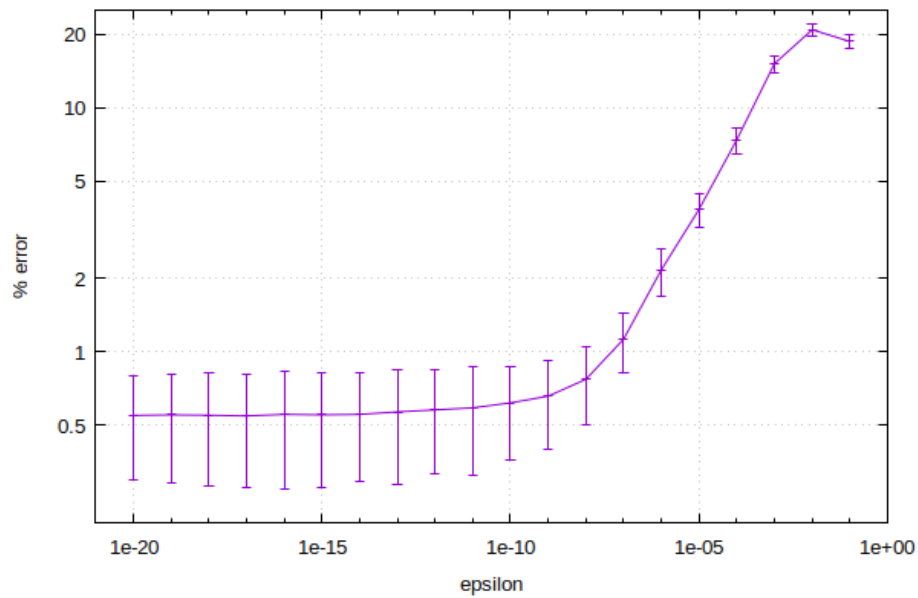


Figure 1: Porcentaje de error en función del valor de epsilon empleado en el suavizado de Laplace. Las barras de error son el valor 2σ para 30 particiones aleatorias de los datos (90% entrenamiento).

2 Comentario de los resultados

Esta práctica consiste en la implementación de un clasificador multinomial por máxima verosimilitud. Se ha obtenido la tasa de error para un clasificador sin suavizado, y es 55.14%. Empleando un suavizado de Laplace con ϵ variable, se obtiene la gráfica de la sección anterior.

Los resultados obtenidos indican que la tasa de error disminuye con el valor de ϵ , bajando hasta un mínimo de 0.55% para $\epsilon = 1e-17$, 100 veces menos que el error original. Esto se debe a que existen muchos elementos en los vectores de pesos que tienen probabilidad 0, lo que genera errores de clasificación. El hecho de que los resultados mejoren conforme disminuye ϵ se debe a que cuanto menos se modifique la distribución de palabras en los vectores de peso mejores resultados se obtendrán.