

Memoria Práctica 2. Clasificación de textos

Marcos Esteve Casademunt, Jose Gómez Gadea

Junio 2018

Índice

1. Multinomial	1
2. Multinomial con suavizado	1

Índice de figuras

1. Evolución del error al variar el parámetro ϵ en el suavizado de Laplace	2
-------------------------------------------------------------------------------------	---

1. Multinomial

Si aplicamos la distribución multinomial sobre los datos sin realizar suavizado obtenemos un error del 55,1 %. Esto es debido al sobre-entrenamiento que se da cuando entrenamos por máxima verosimilitud.

2. Multinomial con suavizado

En la práctica se ha implementado un suavizado utilizando Laplace. Este suavizado consiste en sumarle una constante ϵ (muy pequeña) a cada parámetro y renormalizar. En nuestro experimento hemos variado esta constante en el rango 10^{-1} , 10^{-20} . Como podemos observar, la tasa de error disminuye al disminuir ϵ . Pasando de un error de un 55,1 % al no aplicar suavizado a un 7,6 % al utilizar un ϵ de 0.0001. Estos resultados se aprecian mejor en la tabla inferior.

ϵ	error
10e-1	0.191
10e-2	0.211
10e-3	0.154
10e-3	0.076
10e-4	0.042
10e-5	0.022
10e-6	0.010
10e-7	0.006
10e-8	0.005
10e-9	0.005
10e-10	0.004

Cuadro 1: Evolución del error al disminuir ϵ en el suavizado de LaPlace

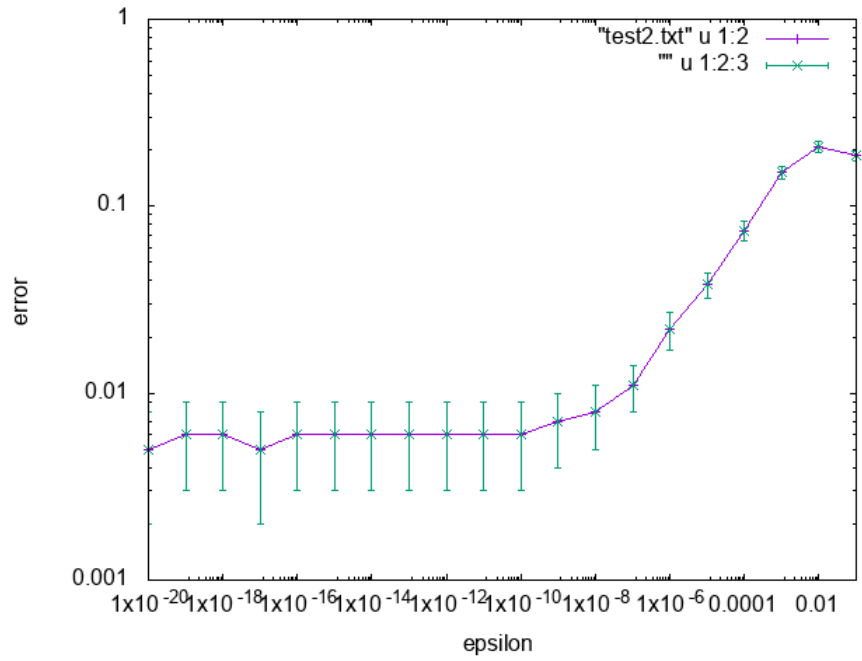


Figura 1: Evolución del error al variar el parámetro ϵ en el suavizado de Laplace