|  |
| --- |
| Trabajo Práctico 6 Filtro de Spam |
| Inteligencia Artificial |

|  |
| --- |
| Pedro Tamargo Allué  2-1-2020 |

Tabla de contenidos

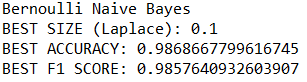
[Comparación Naive Bayes en función de la distribución 2](#_Toc28860527)

[Comparación Naive Bayes en función del parámetro del suavizado de Laplace 4](#_Toc28860528)

## Comparación Naive Bayes en función de la distribución

Para la realización de esta comparación se han realizado pruebas con las distribuciones Bernoulli y Multinomial, ambas aplicando suavizado de Laplace.

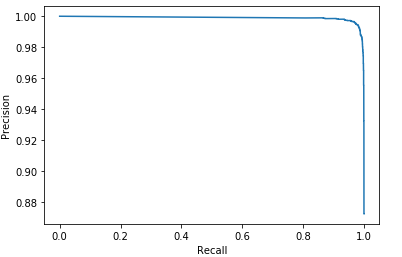
Para la implementación de las pruebas con la distribución Bernoulli se ha utilizado la clase BernoulliNB, implementada en el paquete sklearn.naive\_bayes. Los resultados de la misma en tasa de acierto y f1 son:



Captura de pantalla de los resultados del entrenamiento de la red bayesiana con la distribución Bernoulli.

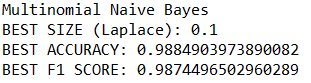


Matriz de confusión para la Red Bayesiana con distribución Bernoulli.



Curva Precisión – Recall de la Red Bayesiana con distribución Bernoulli.

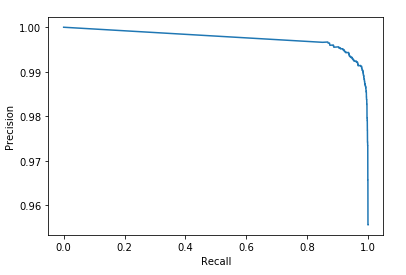
Para las pruebas con la distribución Multinomial se ha realizado un proceso análogo al de la distribución Bernoulli, utilizando la clase MultinomialNB del paquete nombrado anteriormente. Los resultados de la ejecución de la misma son:



Captura de pantalla de los resultados del entrenamiento de la Red Bayesiana con la distribución Multinomial.



Matriz de confusión para la Red Bayesiana con distribución Multinomial.



Curva Precisión – Recall para la Red Bayesiana con una distribución Multinomial.

Podemos concluir que, para ambas distribuciones se ha utilizado el valor 0.1 para el parámetro del suavizado de Laplace, y que ambas proveen buen resultado, un 98% de tasa de acierto en ambos casos, siendo la de la distribución Multinomial mínimamente superior.

## Comparación Naive Bayes en función del parámetro del suavizado de Laplace

El suavizado de Laplace es un parámetro utilizado para dar cierta estabilidad estadística al comportamiento del estimador. En este caso, se utiliza para que cuando aparezcan palabras que no han sido vistas anteriormente, es decir, no tienen una probabilidad asociada, su probabilidad no sea 0 y pueda ocasionar fallos en la predicción.

En este caso, para la realización de las comparaciones se van a muestrear, con las dos distribuciones estudiadas anteriormente, los distintos resultados para distintos valores del parámetro del suavizado de Laplace.

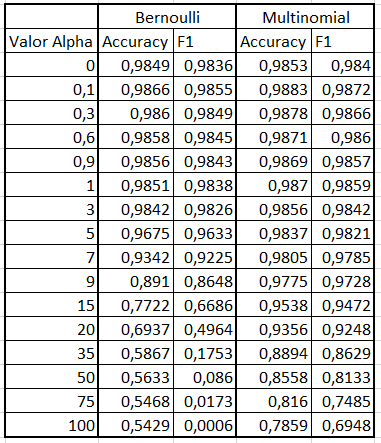


Tabla comparativa variando el valor del parámetro de suavizado de Laplace.

A la vista de los resultados de la tabla anterior, podemos observar que, con se obtienen valores de accuracy y de F1 bastante altos, mientras que con obtenemos valores peores, empeorando más cuanto mayor es . En el caso de usar la distribución Bernoulli los resultados empeoran hasta obtener un valor casi nulo en F1.



Matriz de confusión de una Red Bayesiana con distribución Bernoulli con .

Podemos observar que, la matriz de confusión anterior muestra una gran cantidad de falsos negativos.



Matriz de confusión de una Red Bayesiana con distribución Multinomial con .

En este caso observamos que también existen falsos negativos, pero, a diferencia del caso anterior, se han identificado correctamente más elementos.