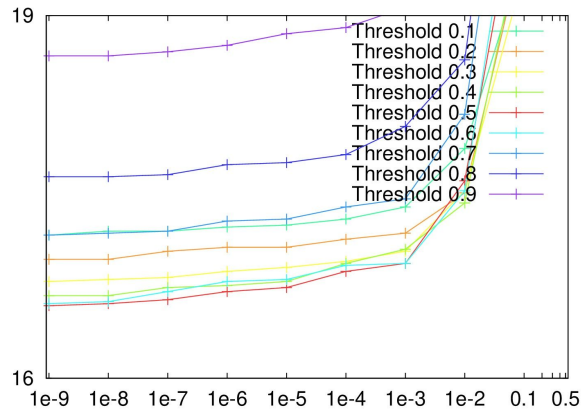
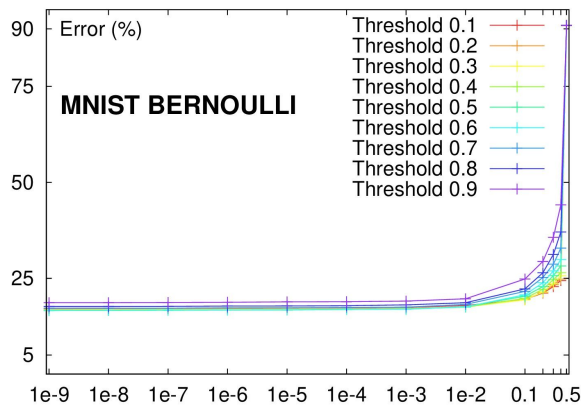


MEMORIA PRÁCTICA 2

Reconocimiento de dígitos manuscritos MNIST

Ejercicio opcional: clasificador de Bernoulli

1. GRÁFICA EPSILON-ERROR PARA CADA VALOR UMBRAL



| Eps/ Thresh | 5e-1 | 4e-1 | 3e-1 | 2e-1 | 1e-1 | 1e-2 | 1e-3 | 1e-4 | 1e-5 | 1e-6 | 1e-7 | 1e-8 | 1e-9 | 0 |
|----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------------|-------|
| 0.1 | 90.883 | 24.383 | 22.833 | 21.15 | 19.5 | 17.9 | 17.416 | 17.316 | 17.266 | 17.25 | 17.216 | 17.216 | 17.183 | 90.25 |
| 0.2 | 90.883 | 24.783 | 23.333 | 21.233 | 19.633 | 17.533 | 17.2 | 17.15 | 17.083 | 17.083 | 17.05 | 16.983 | 16.983 | 90.25 |
| 0.3 | 90.883 | 25.65 | 23.583 | 21.433 | 19.366 | 17.55 | 17.05 | 16.966 | 16.916 | 16.883 | 16.833 | 16.816 | 16.8 | 90.25 |
| 0.4 | 90.883 | 26.433 | 24.666 | 22.133 | 19.766 | 17.45 | 17.066 | 16.95 | 16.8 | 16.766 | 16.75 | 16.683 | 16.683 | 90.25 |
| 0.5 | 90.883 | 28.216 | 25.7 | 22.833 | 20.316 | 17.633 | 16.95 | 16.883 | 16.75 | 16.716 | 16.65 | 16.616 | 16.6 | 90.25 |
| 0.6 | 90.883 | 29.883 | 26.933 | 23.833 | 20.7 | 17.55 | 16.95 | 16.933 | 16.816 | 16.8 | 16.716 | 16.633 | 16.616 | 90.25 |
| 0.7 | 90.883 | 32.866 | 28.683 | 25.216 | 21.583 | 18.183 | 17.483 | 17.416 | 17.316 | 17.3 | 17.216 | 17.2 | 17.183 | 90.25 |
| 0.8 | 90.883 | 37.066 | 31.216 | 26.433 | 22.233 | 18.633 | 18.083 | 17.85 | 17.783 | 17.766 | 17.683 | 17.666 | 17.666 | 90.25 |
| 0.9 | 90.883 | 44.133 | 35.683 | 29.4 | 24.816 | 19.65 | 19.066 | 18.9 | 18.85 | 18.783 | 18.7 | 18.666 | 18.666 | 90.25 |

En esta parte de la práctica primero obtuvimos el error del clasificador sin aplicar la técnica de suavizado por truncamiento simple (90.25) utilizando como conjunto de entrenamiento, el 90% del set de entrenamiento de MNIST y como conjunto de test, el otro 10% restante. Después de esto procedimos a probar diferentes valores de Epsilon y umbral (threshold), valores que se pueden ver tanto en la tabla como la gráfica. El valor umbral sirve par binarizar los datos.

Los mejores resultados se obtienen con Epsilon=1e-9, y umbral=0.5, que es justo cuando por encima de la mitad los valores toman el valor 1, y por debajo, el valor 0.

2. RESULTADOS DEL CLASIFICADOR GAUSSIANO+PCA

En la segunda parte de esta práctica, lo que se nos requería era comprobar el error del clasificador utilizando la técnica de suavizado por truncamiento simple con los valores de Epsilon y umbral que mejores resultados ofrecían, utilizando al completo los sets de entrenamiento y test de MNIST.

Los resultados obtenidos con y sin la técnica de suavizado ($\epsilon=0$) son los siguientes:

| Epsilon (threshold=0.5) | Error |
|-------------------------|-------|
| 0 | 15.62 |
| 1e-9 | 90.2 |

Al aplicar la técnica de suavizado y la binarización con $\epsilon=1e-9$ y $\text{umbral}=0.5$, obtenemos una tasa de error de 15.62%, es decir, hemos reducido el error en un 74.58% respecto del 90.2% que proporciona sin ninguna de éstas.

Finalmente, comentar que el clasificador de Bernoulli sin la técnica de suavizado es incapaz de clasificar correctamente las clases, ya que elige de forma aleatoria, como podemos observar en el porcentaje de error, 9/10 o 90%. También podemos observar que ocurre lo mismo con el valor de $\epsilon=0.5$, esto se debe, a que todas las probabilidades toman ese valor, por lo que no hay forma de discriminar las clases.