

PREGUNTA N 1

RECONOCIMIENTO DE ESCRITURA

**** Usted se equivocó escribiendo MINST en vez de MNIST**

1. Que es la base de datos MNIST ? Y que relevancia tiene para las técnicas de aprendizaje de maquina?
2. Que se entiende por redes neuronales convolucionales y en que se basa su descubrimiento y aplicación ? Describa como se aplican en el caso de reconocimiento de escritura.
3. Formule y describa los términos de la función de costo adoptada en la solución del problema de reconocimiento de escritura descrito en este estudio.
4. Comente el efecto de la constante de aprendizaje, el numero de iteraciones usadas, así como del tamaño de la muestra (Fig 5 y 6)
5. De acuerdo a lo visto en clase. Que representan en las figuras anteriores los sets de pruebas y el entrenamiento ?

Solucion:

1. La base de datos **MNIST** (Mixto Instituto Nacional de Estándares y Tecnología de base de datos) es una gran base de datos de dígitos escritos a mano que se utiliza comúnmente para la formación de varias de procesamiento de imagenes de sistemas. La base de datos también es ampliamente utilizado para el entrenamiento y las pruebas en el campo de aprendizaje automatico Fue creada por "re-mezclar" las muestras de conjuntos de datos originales del NIST. Los creadores sintieron que ya formación de datos del NIST fue tomada de American Oficina del Censo delos empleados, mientras que el conjunto de datos de prueba se tomó de estadounidenses de secundaria los estudiantes, conjunto completo de datos del NIST era demasiado duro. Además, las imágenes en blanco y negro de NIST se normalizaron a caber en un cuadro de límite de píxeles 20x20 y regulando los niveles de escala de grises.

Su importancia es la normalizacion de la escritura que hace para poder ser identificado de manera mas eficiente.

2. Una red neuronal convolucional es un tipo de red neuronal artificial donde las neuronas corresponden a campos receptivos de una manera muy similar a las neuronas en la corteza visual primaria (V1) de un cerebro biológico. Este tipo de red

es una variación de un perceptron multicapa, pero su funcionamiento las hace mucho más efectivas para tareas de visión artificial, especialmente en la clasificación de imágenes.

Las redes neuronales convolucionales consisten en múltiples capas con distintos propósitos. Al principio se encuentra la fase de extracción de características, compuesta de neuronas convolucionales y de reducción de muestreo. Al final de la red se encuentran neuronas de perceptron sencillas para realizar la clasificación final sobre las características extraídas.

La fase de extracción de características se asemeja al proceso estimulante en las células de la corteza visual. Esta fase se compone de capas alternas de neuronas convolucionales y neuronas de reducción de muestreo. Según progresan los datos a lo largo de esta fase, se disminuye su dimensionalidad, siendo las neuronas en capas lejanas mucho menos sensibles a perturbaciones en los datos de entrada, pero al mismo tiempo siendo estas activadas por características cada vez más complejas.

Al ser este tipo de redes diseñadas para reconocer imágenes, es una manera potente para reconocimiento de escritura y por consiguiente los caracteres de escritura.

3. Se tiene la función de costo es:

$$E(W) = \frac{1}{P} \sum_{p=1}^P y_{D^p}(Z^p, W)$$

Donde:

y_{D^p} es la salida de las D_{p^*} -ene-avo RBF(Radial Basics Functions)
por ejemplo: el uno corresponde a la clase correcta de patrones de ingreso.

Z^p Clase de patrón de ingreso.

El parametro vector de RBF tiene un role en la capa de vectores F6.

Cabe señalar que ser componentes de los vectores son 1 o -1, que está dentro de la gama de la gama del sigmoid de F6, y no antes evita esos sigmoids de quedar saturada.

En efecto +1 y -1 son los puntos de maxima curvatura de las sigmoids.

Esto fuerza a las unidades de F6 a operar en un rango no lineal.

A esta funcion que minimiza se le llama Loss Function, que seria en nuestro problema, la funcion de costo.

En este caso la red felizmente ignora las entradas y todas las salidas de RBF son iguales a 0. Este fenómeno colapso no se produce si los pesos RBF no se les permite adaptarse. El segundo punto a tomarse es que no hay competición entre las classes. En cada competición se puede obtener usando mas entrenamiento discriminativo, basado en el criterio de Maximo a posteriori, similar al criterio Maximum mutual information, a veces usadopara entrenar HMMs (Hidden Markov Model).

Esto corresponde a maximizar la probabilidad posterior de la clase correcta D_{p^*} (o minimizando el logaritmo de la probabiliad de la clase correcta), dando la imagen de entrada desde una de las classes o desde el label de clase rubbish background.

En terminos de pero, significa que tenemos que añadir la manera correcta de la clase, entonces transforma a la Loss Function en:

El negativo del segundo miembro se da como el rol de competitividad.

Y tiene que ser necesariamente menor o igual al primer termino para que asi la Loss Function sea positiva. Se define como la probabilidad posterior de rubbish como el radio de e^{-j} y $e^{-j} + \sum_i e^{-y_i(Z,W)}$, este criterio evita las colisiones.

Y al computar el gradiente de Loss Function, nos da el “rubbish background” que es una propagacion inversa (hacia atras).

Entonces el algoritmo debe ser diseñado de tal manera que, las derivadas parciales de cada conexion, como si la red donde es convencional una red multicapa sin capas compartidas. Y estas derivadas parciales definidas por nodo son añadidas para derivar con respecto al parametro.

4. Se usaron 15000, 30000 y 60000 ejemplos en unas 20 iteraciones, al final añadiendole 540000 datos generados de manera aleatoria.

En la figura 5 se puede notar que el error de entrenamiento se mantiene como que constante cuando se aproxima al sobreaprendizaje. En la figura 6 muestra que a pesar de la nueva arquitectura utilizada (LeNet-5) mientras más data de entrenamiento hay más va a mejorar la aproximación.

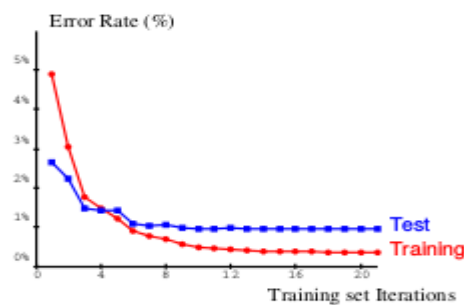


Fig. 5. Training and test error of LeNet-5 as a function of the number of passes through the 60,000 pattern training set (without distortions). The average training error is measured on-the-fly as training proceeds. This explains why the training error appears to be larger than the test error. Convergence is attained after 10 to 12 passes through the training set.



Fig. 6. Training and test errors of LeNet-5 achieved using training sets of various sizes. This graph suggests that a larger training set could improve the performance of LeNet-5. The hollow square show the test error when more training patterns are artificially generated using random distortions. The test patterns are not distorted.

5. Los set prueba son los patrones sin distorsión o distorsión mínima y el set de entrenamiento con tiene patrones de distorsión y algunos poco sin distorsión. En las figuras se ve como varía el error promedio resultante a través de la variación del número de iteraciones y el tamaño del set de entrenamiento.