

MNIST Database of Handwritten Digit Images for Machine Learning research

Norman Daniel Vicente Orellana
Universidad Galileo



Introducción

El reconocimiento de dígitos escritos a mano es un problema importante en el área de reconocimiento óptico, y se ha utilizado como casos de prueba en las teorías del reconocimiento de patrones y algoritmos de aprendizaje automático para muchos años. Históricamente, para promover aprendizaje automático y reconocimiento de patrones e investigación, varias bases de datos estándar han emergido en el que los dígitos escritos a mano están preprocesados, incluida la segmentación y normalización, para que los investigadores puedan comparar los resultados del reconocimiento de sus técnicas sobre una base común. La base de datos MNIST de manuscritos se ha convertido en un estándar para los algoritmos de aprendizaje automático de prueba rápida para este propósito.

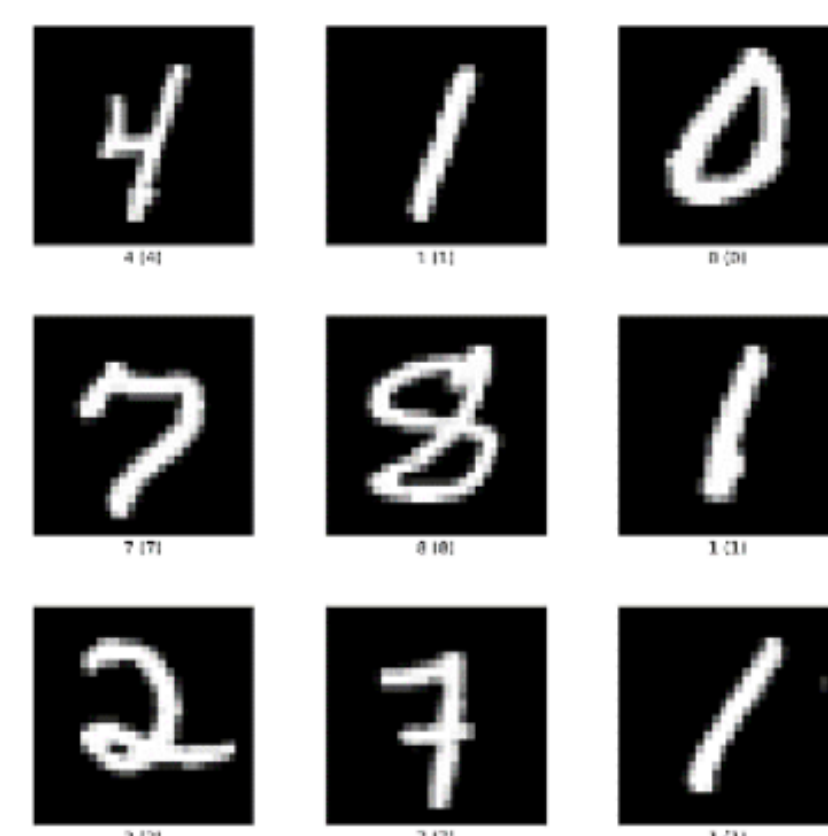
Problema

La importancia del reconocimiento de imágenes en el aprendizaje automático a estado evolucionando por lo tanto se necesita un estándar con el cual poder probar, entrenar y comparar los distintos algoritmos. MNIST provee estas características ya que es estándar.

Metodologia

MNIST fue construida a partir de la base de datos NIST. Consta de 60,000 imágenes de entrenamiento o también para Cross-Validation. Cuenta con 10,000 imágenes de prueba, ambos sets comparten la misma distribución. Cada imagen son dígitos en blanco y negro esta normalizadas y centradas en una imagen de tamaño fijo donde el centro de gravedad de la intensidad se encuentra en el centro de la imagen con 28 x 28 píxeles.

Imágenes de MNIST



Resultados

La dimensionalidad de cada imagen es $28 * 28 = 784$, donde cada imagen esta binarizada.

Un total de 68 distintos algoritmos de clasificación fueron entrenados con MNIST estos algoritmos están basados en las siguientes técnicas:

1. Clasificadores lineales
2. K vecinos cercanos
3. Clasificadores no lineales
4. Redes convolucionales
5. Redes neuronales

En la mayoría de algoritmos no se necesita un preprocesamiento de imágenes ya que MNIST provee esta característica como el normalizado de las imágenes, aplicación de filtros de color etc. En algunos experimentos, el entrenamiento fue alterado con versiones distorsionadas artificialmente de las muestras de entrenamiento originales. El las distorsiones incluyen combinaciones aleatorias de temblores, cambios, escalado, enderezamiento, deslavado, difuminado y compresión.

Los clasificadores de redes neuronales tienden a funcionar significativamente mejor que otros tipos de clasificadores. Específicamente, las redes neuronales convolucionales dan un excelente rendimiento de clasificación, incluso da el rendimiento récord, sobre Tasa de error del 0,27% o 27 errores en total 10,000 test set.

Conclusiones

Esta es una base de datos relativamente simple para las personas que desean probar los algoritmos basados en reconocimiento de imágenes, técnicas de aprendizaje y métodos de reconocimiento de patrones en datos del mundo real y deseen gastar esfuerzos mínimos en preprocesamiento y formateo. MNIST ofrece muchas referencias en su sitio web, los estudiantes y los educadores de aprendizaje automático también pueden beneficiarse de un conjunto bastante completo de literatura sobre aprendizaje automático con comparación de rendimiento disponible.

Debido a los buenos resultados se llevó a la creación de la base de datos TIMIT (es una base de datos creada por MIT) que es familiar para la mayoría de los investigadores de procesamiento de voz y procesamiento de señales.

Bibliografias

1. <https://ieeexplore.ieee.org/abstract/document/6296535>