



DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Práctico II

Reconocimiento de Dígitos

Métodos Numéricos
Primer Cuatrimestre - 2015

Integrante	LU	Correo electrónico
Christian Cuneo	755/13	chriscuneo93@gmail.com
Ignacio Lebrero	751/13	ignaciolebrero@gmail.com



Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Ciudad Universitaria - (Pabellón I/Planta Baja)

Intendente Güiraldes 2160 - C1428EGA

Ciudad Autónoma de Buenos Aires - Rep. Argentina

Tel/Fax: (54 11) 4576-3359

<http://www.fcen.uba.ar>

Índice

1. Introducción	1
2. Desarrollo	2
2.1. Implementación	2
2.1.1. Knn	2
3. Experimentaciones	3

1. Introducción

En el presente informe se presenta la implementación, desarrollo y análisis de resultados usados en la resolución de un problema práctico, con las herramientas estudiadas y provistas por la materia. El problema principal planteado consiste en lograr reconocer un dígito manuscrito guardado en una imagen, quiere decir, poder predecir a que dígito corresponde una imagen dada.

Para resolver este problema se te da una base de datos imágenes de dígitos manuscritos que ya se saben a que dígito corresponde cada imagen (grupo de imágenes "train"), y por supuesto un grupo de imágenes a las que deberás asignar un dígito utilizando alguna forma de reconocimiento (grupo de imágenes "test"). Para hacer el proceso mas simple todas las imágenes son cuadradas de 28x28 pixeles y en escala de grises.

El problema se aborda desde dos puntos de vista, el probabilista y el del álgebra lineal. El álgebra lineal se utiliza para llegar a las soluciones, y la probabilidad se utiliza para hacer este proceso de una forma mas eficiente y rendidora.

A través de la generalización uno puede ver a una imagen como una coordenada en un espacio \mathbf{R}^n , siendo n la cantidad de pixeles de la imagen (en este caso 784), y como cada imagen de la base de datos de imágenes conocidas (train) es entonces una coordenada con una clasificación conocida (su dígito); luego, una aproximación a la solución valida seria, dada una imagen desconocida, buscar que puntos mas cercanos tiene en el espacio \mathbf{R}^{784} , y asignarle un resultado acorde. Ya existe un método que realiza toma estos pasos, y se llama K-vecinos más cercanos (Knn por su ingles "K nearest neighbors") y consiste en buscar, utilizando la distancia euclidiana, los K puntos mas cercanos a cierto punto, en nuestro caso, la imagen desconocida.

Luego de tener estos K puntos cercanos, uno puede clasificar a que grupo pertenece la imagen desconocida de distintas formas, en nuestro caso se utiliza la votación, básicamente teniendo estos K vecinos, buscamos al grupo dominante (al grupo mayor de puntos con la misma clasificación) y asignamos esta clasificación a nuestra imagen desconocida.

Al trabajar en \mathbf{R}^{784} este proceso se vuelve muy lento, ya que este método se ve afectado directamente por la dimensión de los puntos que procesa, específicamente al calcular la distancia euclidiana, por esto lo que se va a buscar es reducir la dimensión de los puntos, tratando de perder la menor cantidad de características posibles de las imágenes.

Aquí es donde entra en juego la probabilidad y estadística.

Vamos a buscar que pixeles de las imágenes son mas importantes y descartar los menos importantes, reduciendo así la dimensión en la que trabajamos. Para esto vamos a acudir al termino de la covarianza, que, aplicada a dos variables, muestra la tendencia de estas dos variables a variar juntas. Esto, aplicado a estas imágenes, vamos a buscar que puntos de las imágenes son los que mas varían juntos entre todas las imágenes de train, obteniendo así información para saber que características (pixeles) de las imágenes son las mas importantes a analizar, las que destacan a una de otra.

Luego de obtener la matriz de covarianzas de los datos, que va a ser una matriz en $\mathbf{R}^{784 \times 784}$, uno va a usar esta información para reducir las dimensiones de las imágenes, descartando información menos importante. Y luego se reducirá la dimensión de la imagen desconocida usando la misma información y se procederá con el método Knn para clasificar la imagen, esta vez de una forma mas eficiente, ya que las dimensiones son menores; y con mas rendimiento, ya que ahora solo quedo la información mas importante de cada imagen.

El proceso donde se calcula la información mas importante de las muestras se llama "Análisis de Componentes Principales".

2. Desarrollo

2.1. Implementación

2.1.1. Knn

La implementación del método de K-vecinos más cercanos es muy simple, consiste en tomar a la muestra desconocida, que es un vector V en \mathbf{R}^n y la matriz train en $\mathbf{R}^{m \times n}$, con m imágenes ya clasificadas, y tomar la distancia euclidiana de cada fila de la matriz train con V , cada distancia la insertamos en una cola de prioridad inversa. Luego de calcular la distancia con todas las filas de train, se desencolan K elementos de esta fila (al ser de prioridad inversa serán los k mas chicos).

A continuación se realiza el proceso de votación para ver cual es la clasificación dominante en este grupo de k elementos, y se clasifica la imagen desconocida.

Vease que el "n" de las dimensiones sera 784 si se trabaja sin reducción de dimensiones de los datos, o un valor menor α .

3. Experimentaciones

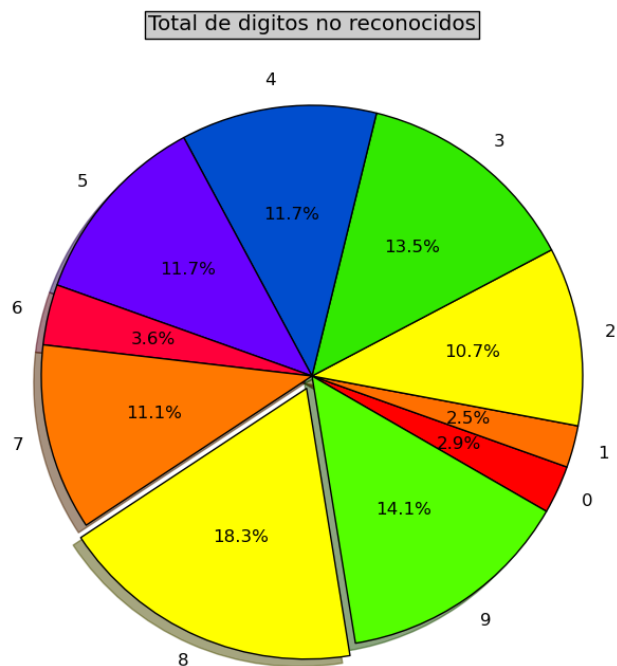


Figura 1: Porcentajes de dígitos clasificados incorrectamente.

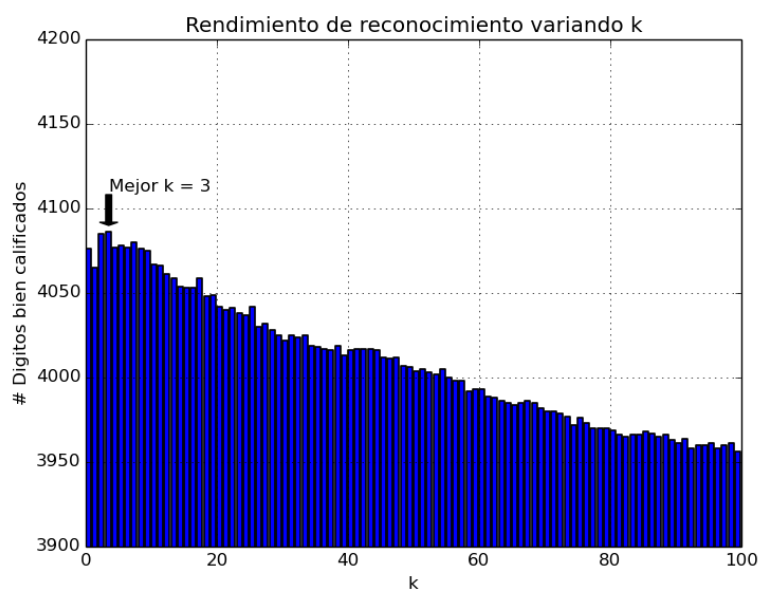


Figura 2: Rendimiento de clasificación al variar k de knn.