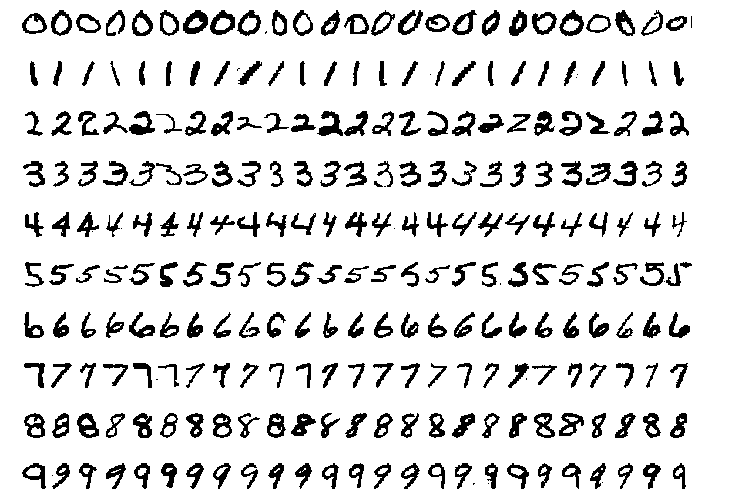
Construction of a Digit Recognition system

****

IN4085

Gianluca Savaia

Guilherme Weishar - 4502043

Wojtek Kleszcz

Index

1 - Introduction

In this report we will provide the results of our digit recognition system that our pattern recognition consulting company developed. Our system will be used by the client to classify indivitual digits in bank account numbers and he monetary amount in cheques. We focused on two scenarios to prove the accuracy of our digit recognition system:

1 - We train the pattern recognition system once with 200 objects per class.

2- We train is trained for each batch of cheques with 10 objects per class.

To develop our digit recognition system we used the **prnist** dataset developed by the US National Institute of Standards and Technology (NIST).

The prnist dataset is a collection of scanned hand written digits from 0 to 9

Each class is a digit and there is a maximum of 1000 objects per class (i.e: 1000 different ways to hand write each digit).

Our first task was to uniformize the dataset. We developed a preprocessing script that resizes every image from the NIST dataset to 68x68 pixels (see **preprocess.m** in the pipeline folder for more details). We tested the classifiers on 3 different ways of representing data: pixels, extracted features and clusters.

For each representation we tested several classifiers:

* Parametric: nmc, ldc, qdc, fisherc, loglc
* Non parametric: knnc, parzenc
* Advanced: neural networks, support vector classifier, one-class classifiers, combining classifiers.

We also tested a series of other techniques like filtering the images (see **getSmoothedDataset.m** in the pipeline folder for more details), kernalizing the dataset (see **raw\_kernelization.m** in the pipeline folder for more details).

In this report we provide graphs and tables that showcase our results. Our objective for scenario 1 is to create a classification error bellow 5% and for scenario 2 bellow 25%.