

Pattern Recognition

Moreno Colombo, Egzon Syka, Lin Bai,
Labinot Jakupi & Sukanya Nath



Contents

- Challenges
- Language and libraries
- Solution Outlines



Challenges

Keyword Spotting

Find Keywords in
handwritten
documents using DTW

Signature

Verify signatures to be
genuine using DTW



Language and libraries

Language

- Python3

Libraries

- Scipy (distance functions)
- Fastdtw (dtw)
- Sklearn(MLP)
- Numpy (array processing and math functions)

Solution Outlines





1

Keyword Spotting

Solution

- ❖ Pre processing:
 - Document images were cut out into separate words
 - Words binarized with Gauss method
 - Features extracted from single words scaled to 50x100 px, window width 1px

- ❖ Recognition:
 - Using the the feature vectors from pre processing
 - Distance between given word (train) and all words of test set with DTW
 - Mean average precision on validation set ~ **54.5%**



1

Keyword Spotting

Solution

- ❖ Feature vectors selection was performed yielding features
 - Upper contour
 - Lower contour
 - Number of black-white transitions
 - Original ratio width/height of each word

W/H ratio = 1.2

W/H ratio = 6.9



2

Signature Verification

Solution

- ❖ Mean distance of genuine signatures of each author was computed using DTW (and stored in a dictionary)
- ❖ For each Validation set instance, mean distance was computed against genuine signatures of that author
- ❖ A threshold of 10.000 was used to compare the Validation instance mean distance with the author's genuine signature mean distance
- ❖ Mean average precision on validation set ~ **87%**



THANK YOU!

Any questions?

Solutions at

<https://github.com/colombmo/pattern-recognition>



1

MNIST

Solution

- ❖ Best possible MLP Parameters found by using grid search
- ❖ solver='sgd',
activation='relu',
learning rate=0.3,
neurons = 100,
alpha=0.01
max_iter=200
- ❖ Accuracy >97% on the test set

