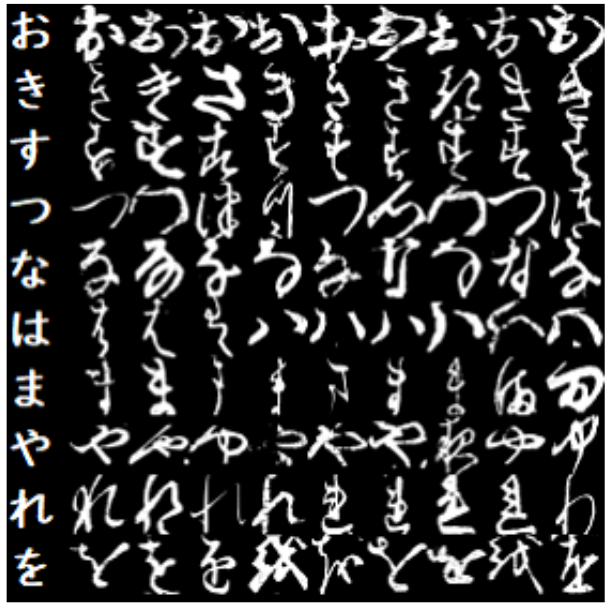
Lab2: Classification with SVM

Machine Learning 2022

F. Chiariotti, A. A. Deshpande

LAB2: Classification with SVM

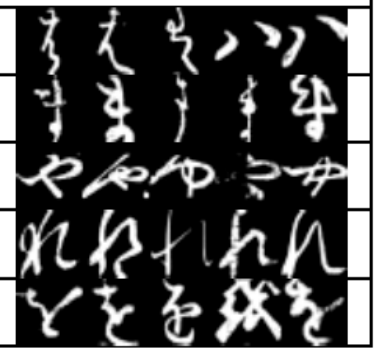


- Classify ancient cursive Japanese (Kuzushiji)writing
- Use Support Vector Machines (SVM)

Lab 2 on 30/11
Delivery on 13/12

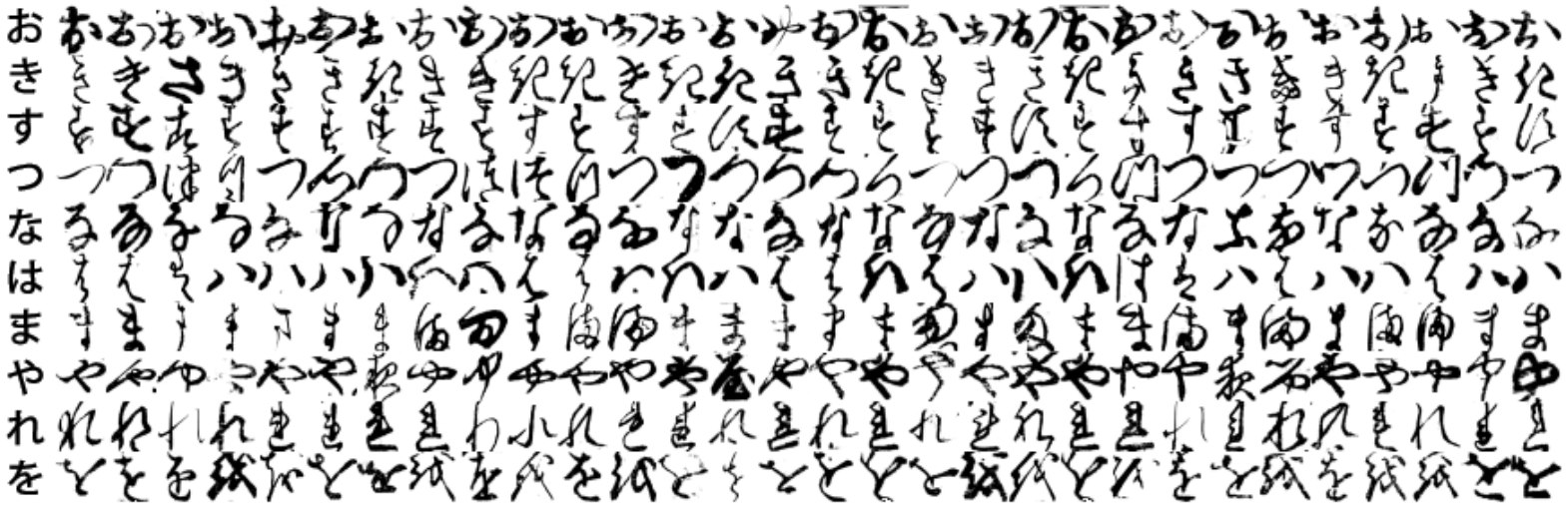
The KMNIST Dataset

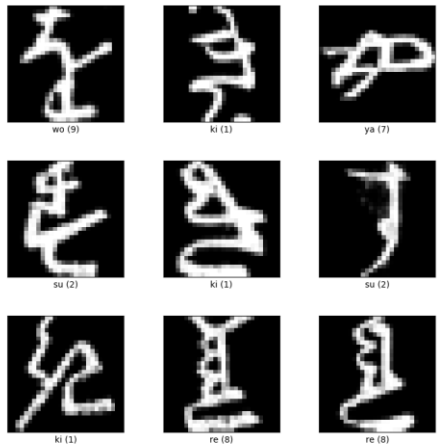
Hiragana	Unicode	Samples	Sample Images
お (o)	U+304A	7000	
き (ki)	U+304D	7000	
す (su)	U+3059	7000	
つ (tsu)	U+3064	7000	
な (na)	U+306A	7000	

Hiragana	Unicode	Samples	Sample Images
は (ha)	U+306F	7000	
ま (ma)	U+307E	7000	
や (ya)	U+3084	7000	
れ (re)	U+308C	7000	
を (wo)	U+3092	7000	

- 10 classes corresponding to 10 different characters
- 70’000 samples (7’000 for each class)
- Divided into 60’000 for training and 10’000 for testing
- Recent deep learning schemes can reach an accuracy of 99%
- Expect an accuracy around 80% for a «baseline» SVM classification

LAB2: Classification of Kuzushiji characters with SVM





- Classify images of characters
- Use Support Vector Machines (SVM)
- Dataset of small pictures of characters: multi-class classification problem
- Use Support Vector Machines
- Try different Kernels
- Estimate parameters with cross validation
- Visualize the results with confusion matrices