Character Recognition System Mini-Project

Prepared By

C11-04 Pavan Agarwal

C11-15 Ankita Bhutra

C11-19 Milan Chandiramani

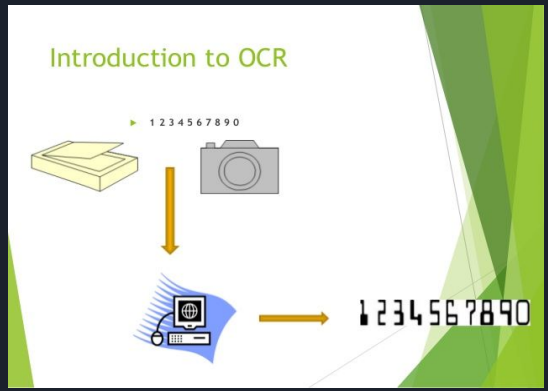
Guide

Mrs. SHILPA VERMA
(Associate Professor,
Department of Computer Engineering,
TSEC)

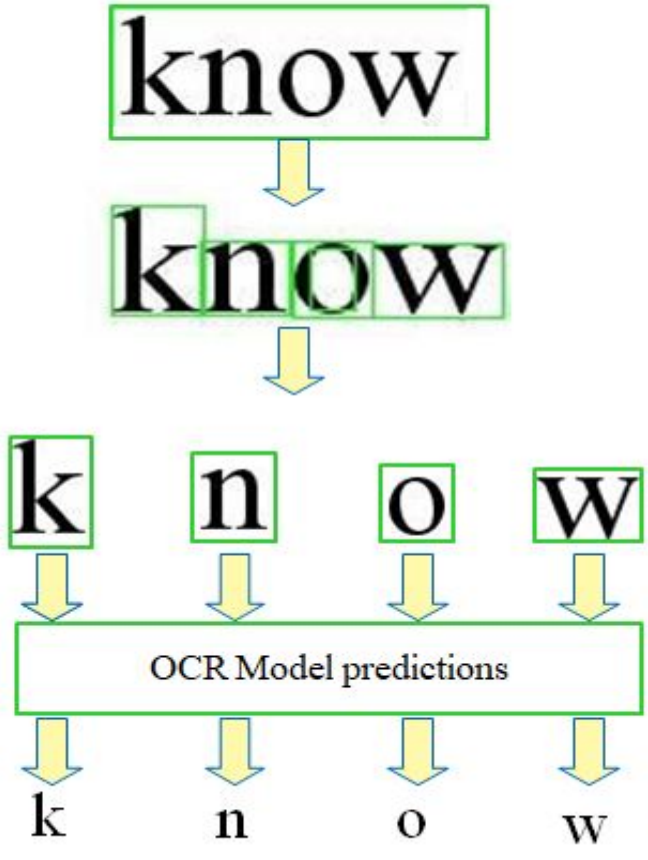
What is Character Recognition?

Process which allows computers to recognize written or printed characters such as numbers or letters and to change them into a form that the computer can use.

Converts a scanned document to an electronic file.



Optical Character Recognition flow diagram



I. Differentiate **word Contours** associated with **Image**.

OpenCV contours, Image cropping

II. Differentiate **letter Contours** associated with **word Contour Image**.

OpenCV contour dilation, Image cropping

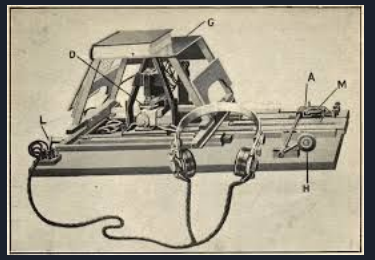
III. Preprocess letter images according to **trained OCR** input.

keras Framework in Detecting, PIL library in Image processing

IV. Consolidate predictions associated **OCR** model to text :-) .

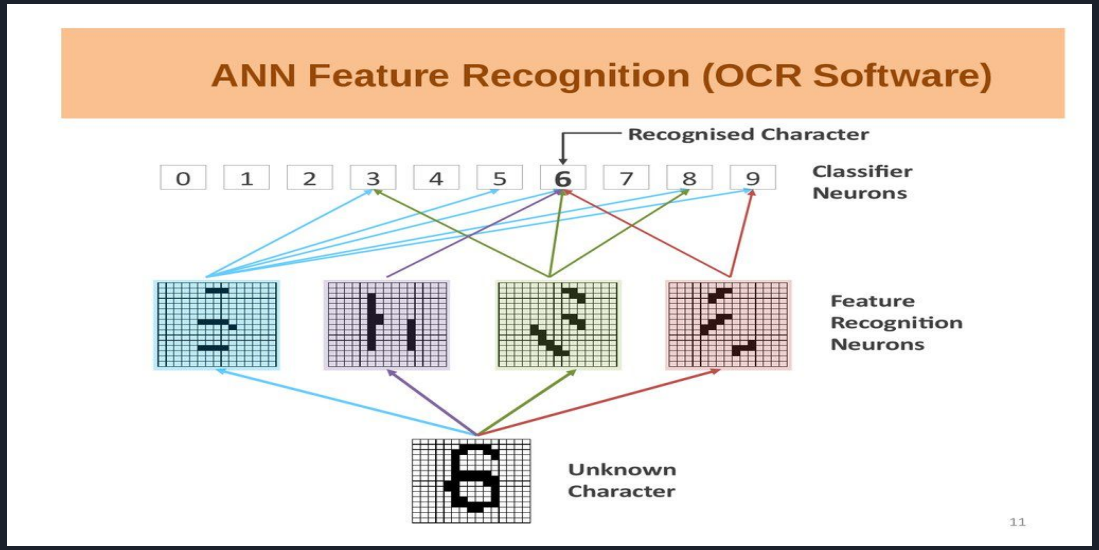
PIL library in Image processing, Python in consolidation

Approaches for character recognition



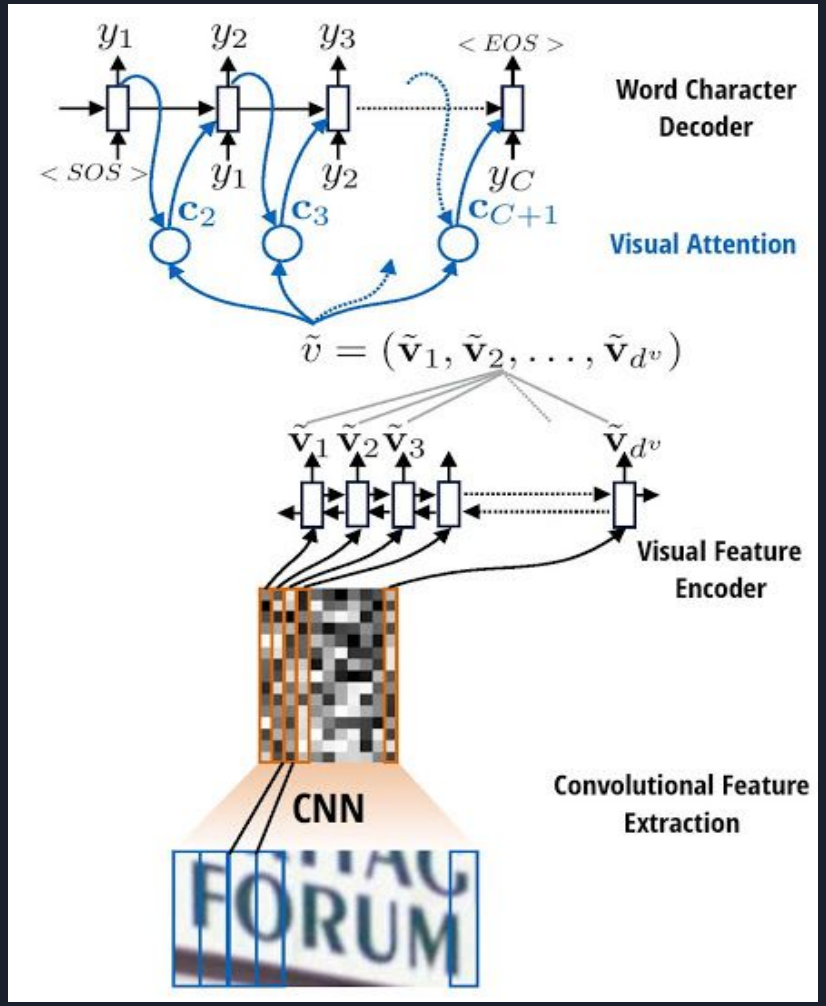
Traditional Optophone

- ANN *Analog Neural Networks*



- CNN

Convolutional Neural Networks



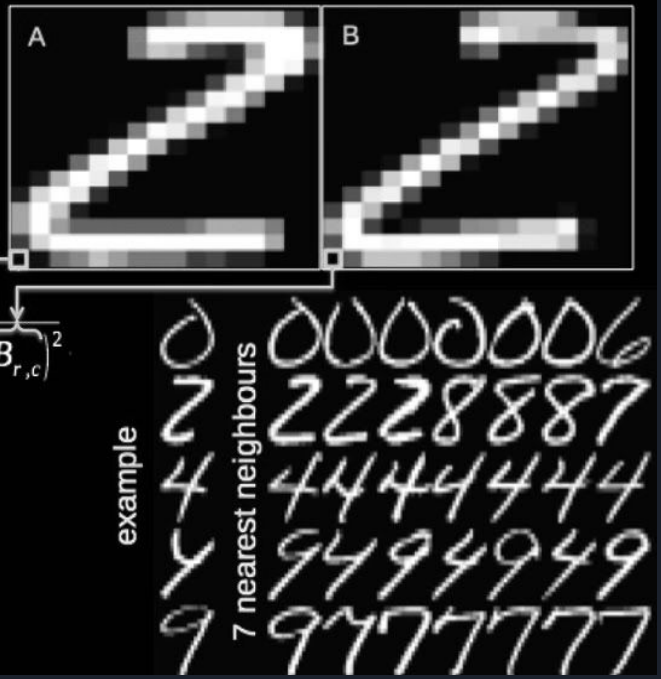
--Approaches for character recognition

- k-NN *K-nearest neighbours*

- 16x16 bitmaps
- 8-bit grayscale
- Euclidian distance
  - over raw pixels

$$D(A,B)=\sqrt{\sum_r \sum_c \left(A_{r,c}-B_{r,c}\right)^2}$$

- Accuracy:
  - 7-NN ~ 95.2%
  - SVM ~ 95.8%
  - humans ~ 97.5%

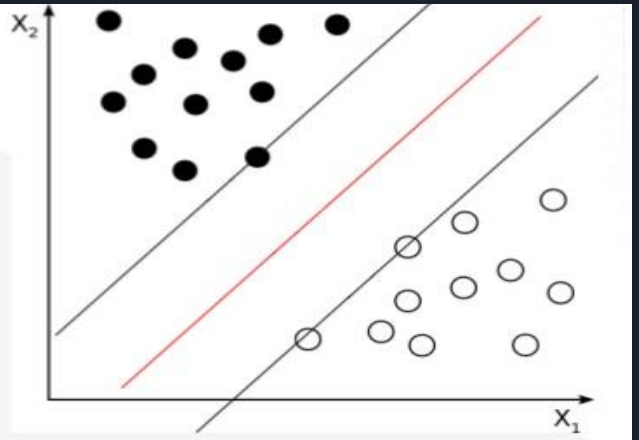


--Approaches for character recognition

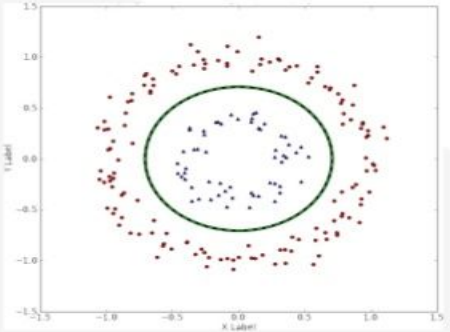
- SVM

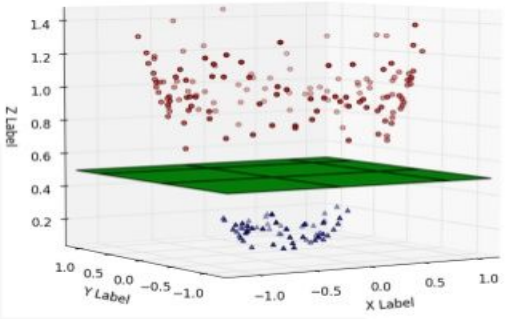
Support Vector Machines

- Optimal hyperplane
- Linear classifier
- Maximum margin
- Classification



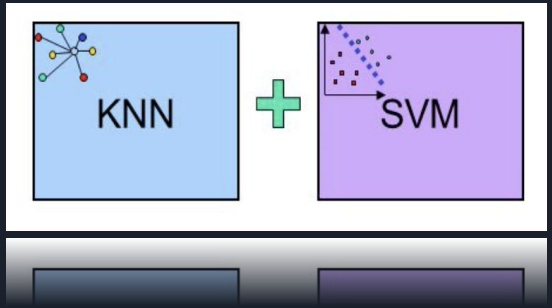
The kernel trick

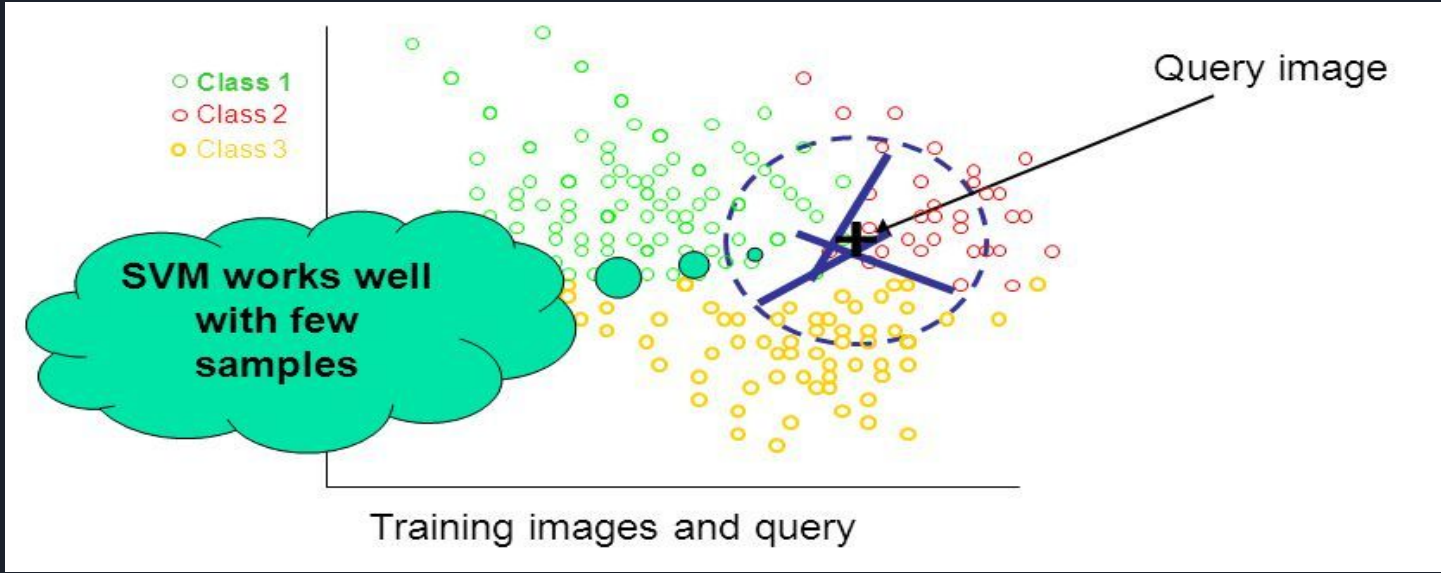




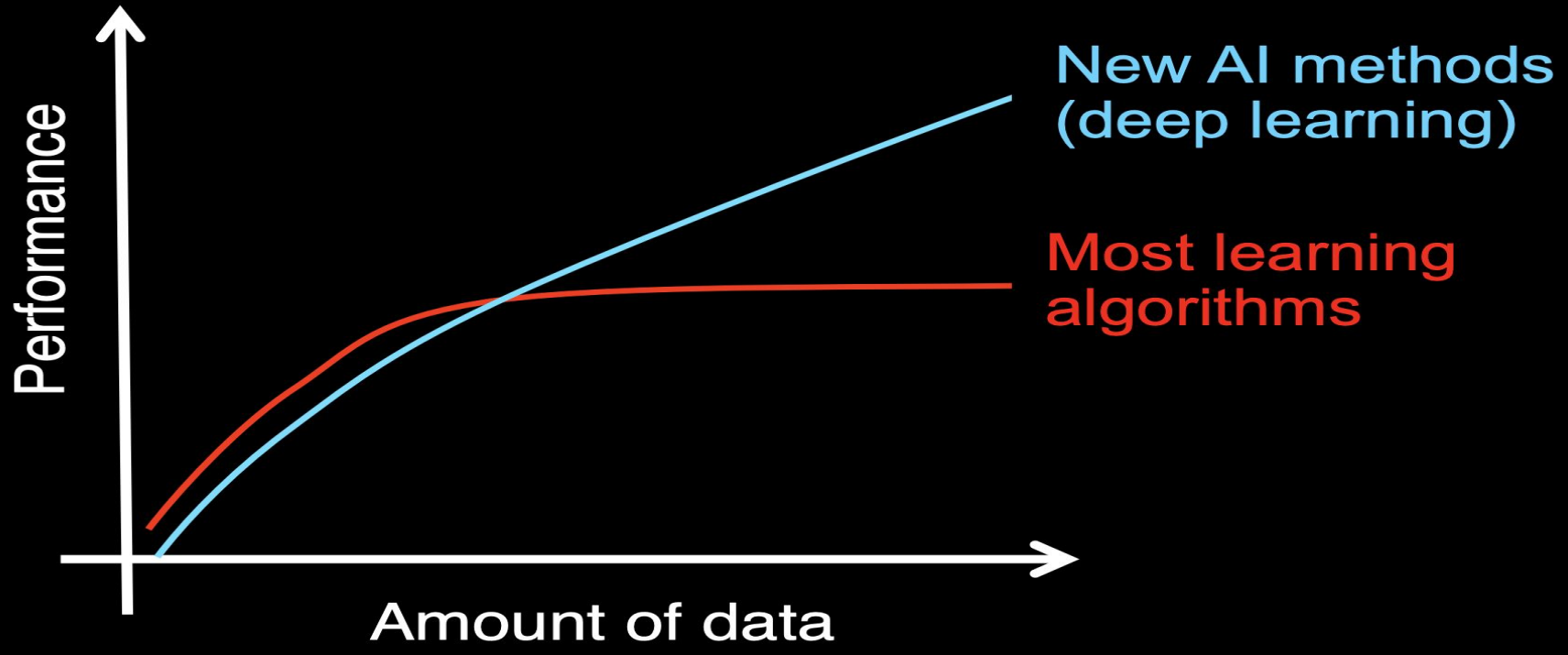
--Approaches for character recognition

- What have we used?

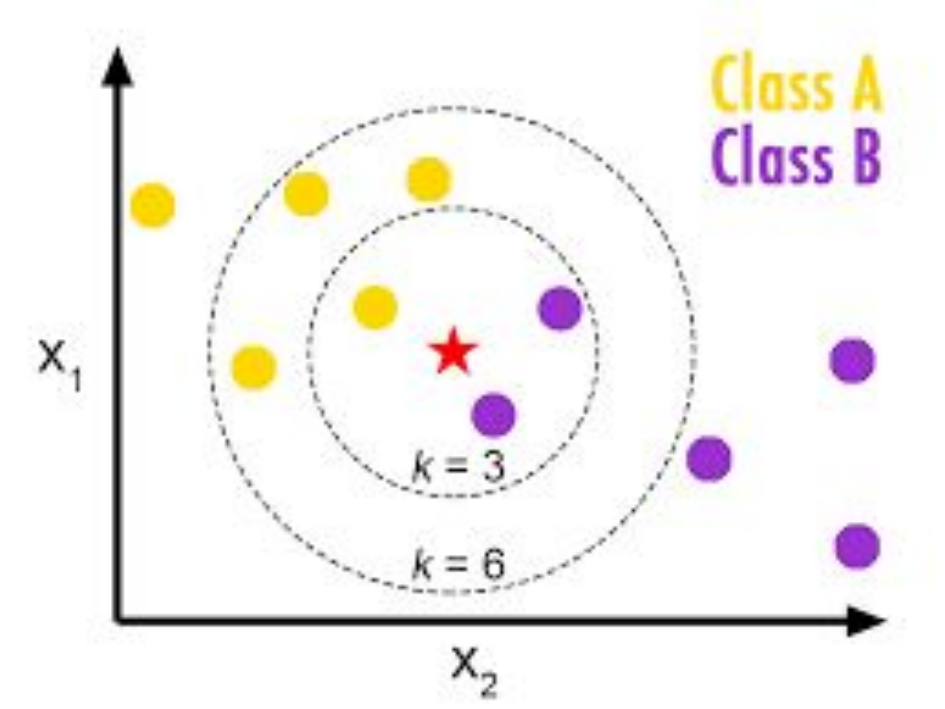




Data and machine learning



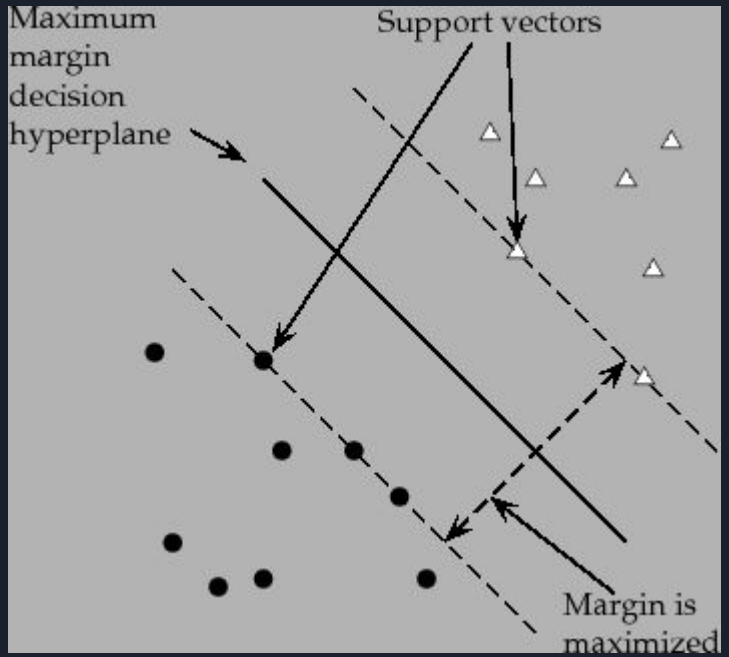
What is KNN?



- K-nearest neighbours Algorithms.
- Instance-based learning
- A Lazy-learning Algorithm
- Uses k-nearest neighbours
- Used for high dimensional data with huge dataset
- The decision boundary in KNN is a convex polygon

--Algorithms Used

What is SVM?



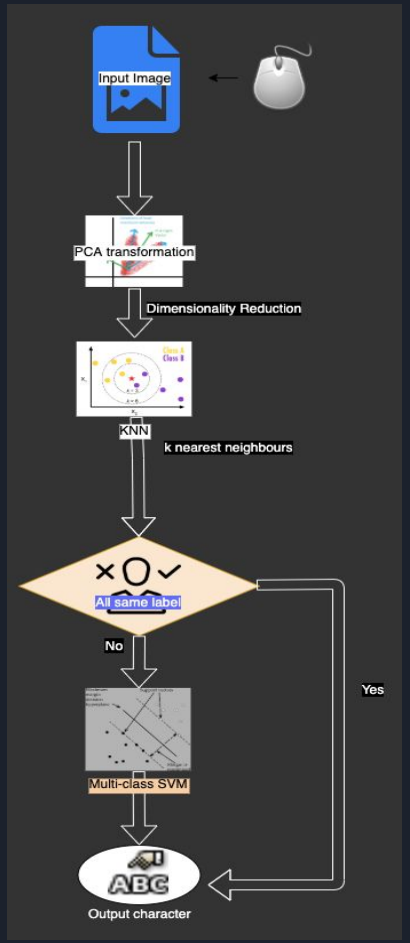
- Supervised learning algorithm.
- SVMs can efficiently perform linear and non-linear classification.
- Objective is to find a hyperplane in a N-dimensional space that distinctly classifies the data points.
- Used for small dataset and low dimensional feature.

--Algorithms Used

How they work Together?

Algorithms Working:-

1. Draw the character for recognition.
2. Apply K-nearest neighbours (k=3)
3. If all neighbours are of same class,
   each neighbour indicate same class label.
   Else
   Apply Multiclass-SVM with
   - ➢ K-neighbours as training data
   - ➢ RBF as kernel function
4. Acquire the Class Label and map the output character.



About our Dataset

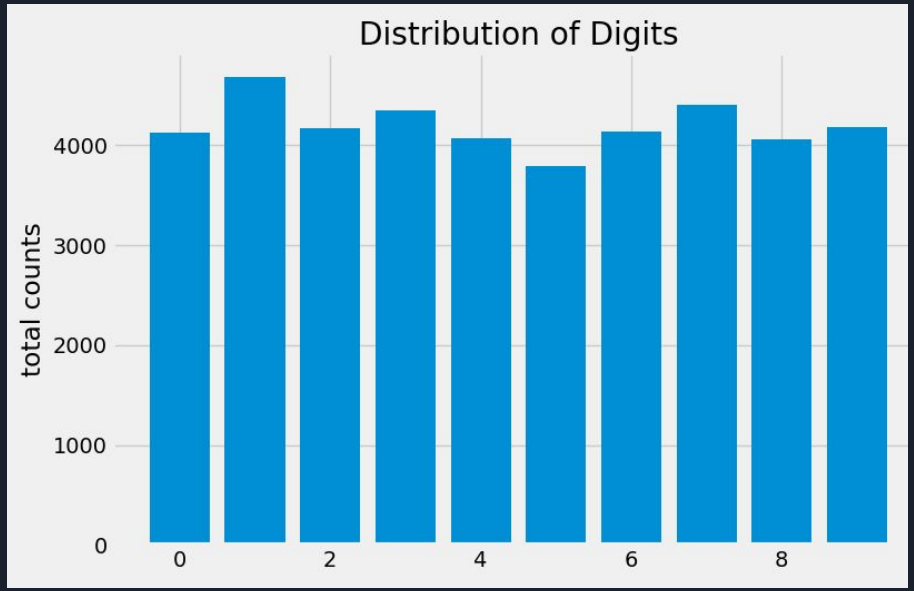
Dataset Used: Emnist -dataset(Character Recognition) & Mnist dataset(Digit Recognition) .

Digit Class:

Image-Size:- 28 X 28

No of bands: 1(Gray Scale)

No of Classes: 10(0.....9)

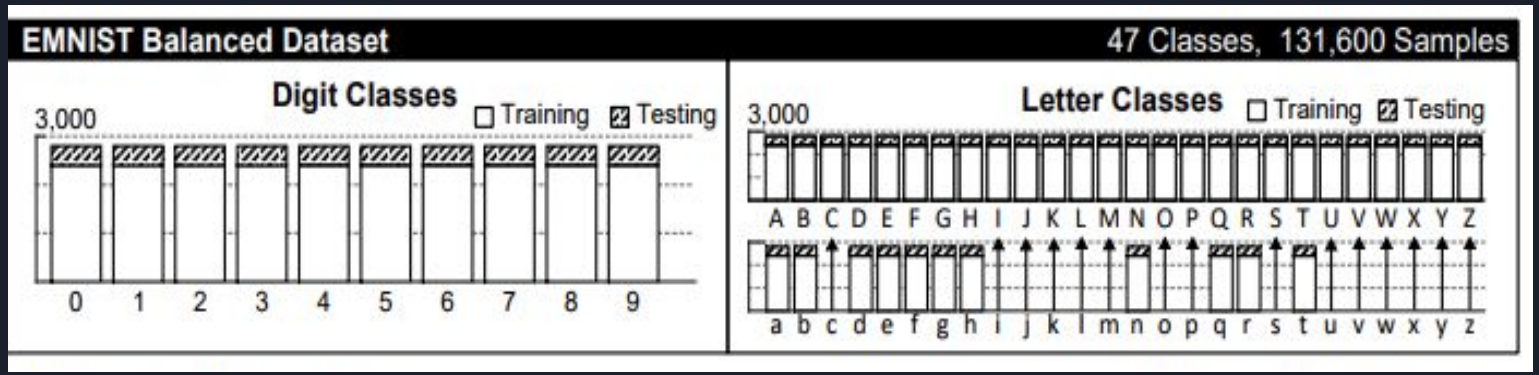


Character Class:

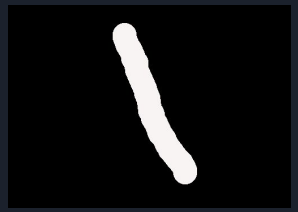
Image-Size: 28 X 28

No of bands: 1(Gray Scale)

No of Classes: 47 (0..9,A..Z,a...Z)



Training



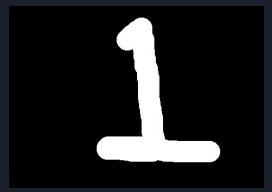
Input Image

Applying K-nearest neighbours

Each Nearest neighbour of same class

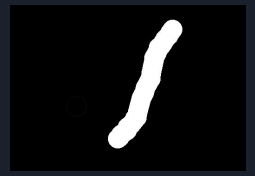
Predicted Output : 1(one)

Training



Input Image

Applying K-nearest neighbours



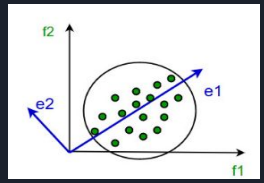
Each Nearest neighbour of different class

Training data for SVM

SVM CLASSIFIER

Predicted Output : 1(one)

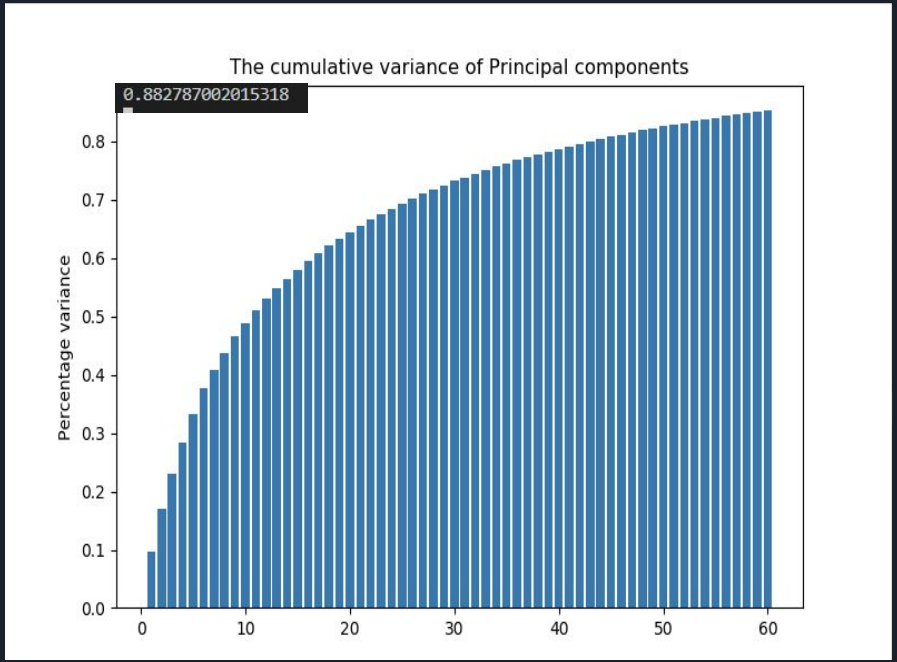
Data Preprocessing: Principal Component Analysis (PCA)



- Reduces data from n-dimensions to k-dimensions.
- Selects the top dimensions with maximum variance.
- Requires every attribute to be equally scaled.

Application in our project:

- Reduces 28 X 28 = 784 pixels of an image to 60 principal components with variance of ~88%.



Some things to be done:
Image Resizing

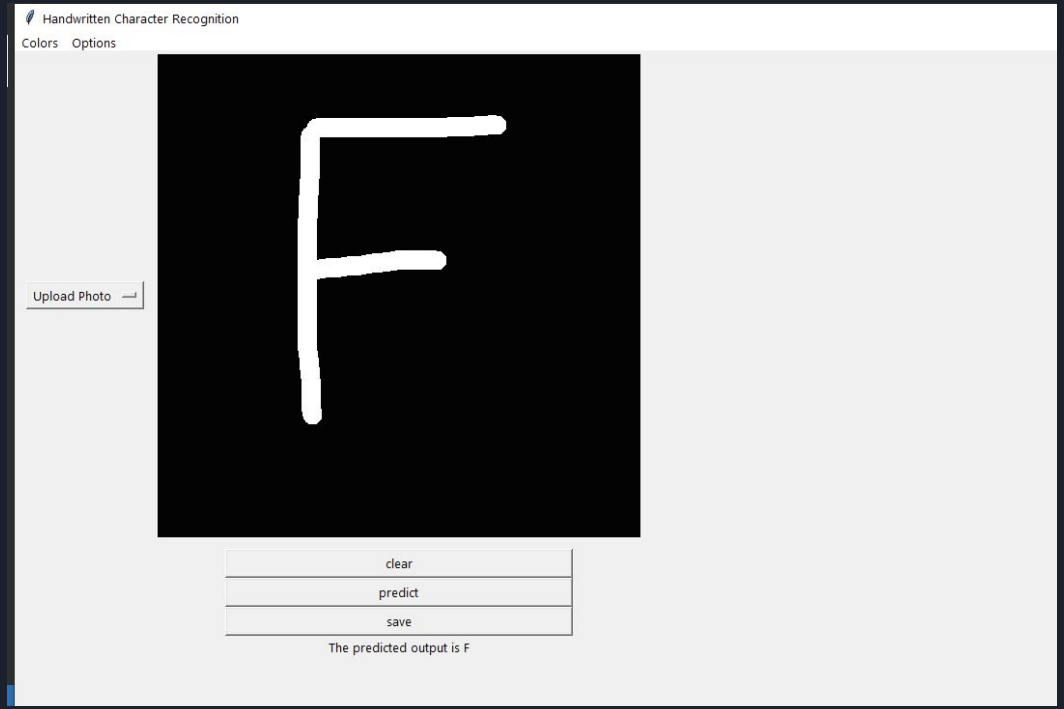
500 X 500 px

Resized by
PIL.image.new

Resized To

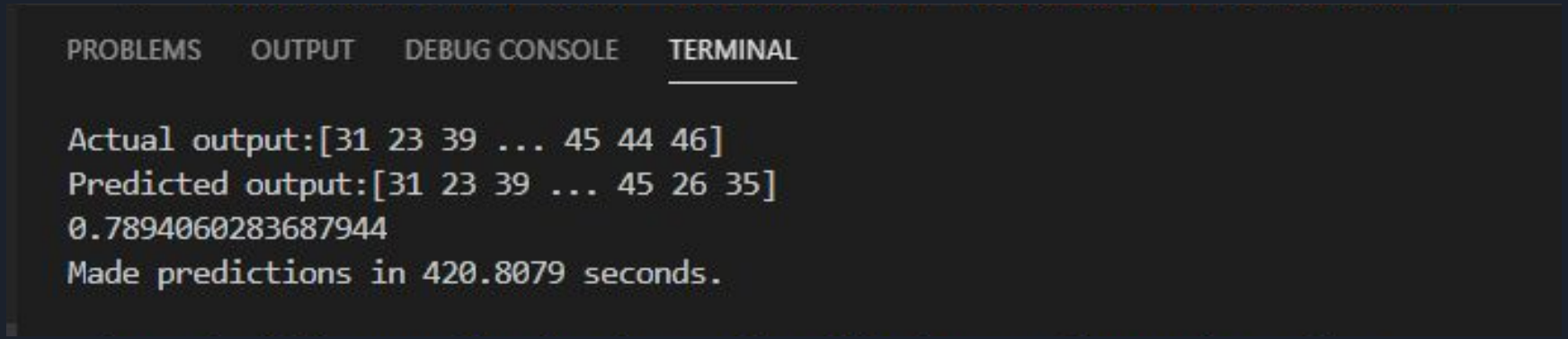
28 X 28 px

App-in Reality



Results

- 98% accuracy on Digit Recognition
- 78% accuracy on Character Recognition.
- ~0.02 second for classification



Limitations

- Similar classes

Class ‘U’

Class ‘J’

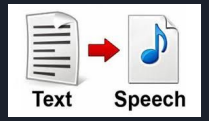
Class ‘n’

- Large Dataset requirements
- Works mediocre with noisy data.

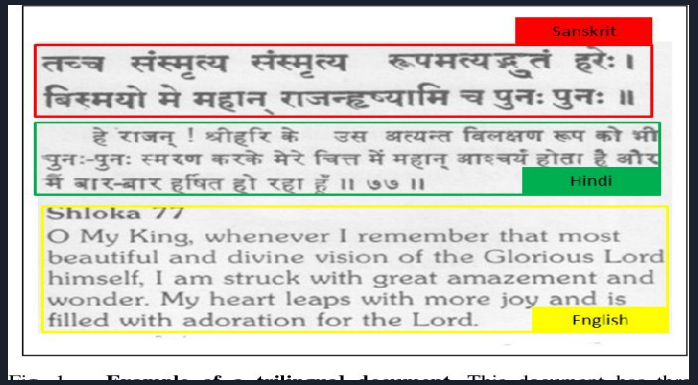
Future Scope

Signboard Translation

A Bilingual or multilingual script OCR



Text-to-Speech Conversion.



Conclusion

- In this report, we proposed a hybrid of SVM and k-NN, which deals naturally with multiclass problems.

- The algorithm works best when the input image is **not**:
  - ➢ A mirrored image,
  - ➢ A water image,
  - ➢ An image rotated at angle >45° .

Prepared By

C11-04 Pavan Agarwal

C11-15 Ankita Bhutra

C11-19 Milan Chandiramani