Handwritten English Alphabet Recognition Using Bigram Cost

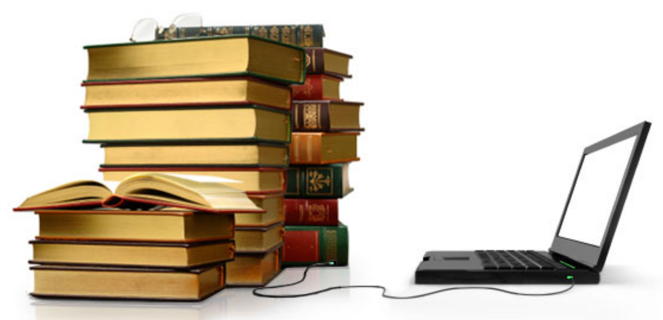
Chengshu (Eric) Li

Introduction & Motivation

Handwritten character recognition has been one of the most challenging and fascinating areas in the field of image processing. It has a wide variety of applications:

- receipt/invoice recognition
- business card information extraction
- books canning
- assistive technology for blind

My approach is to use both image recognition and bigram cost between English alphabets to achieve high performance.



Stanford

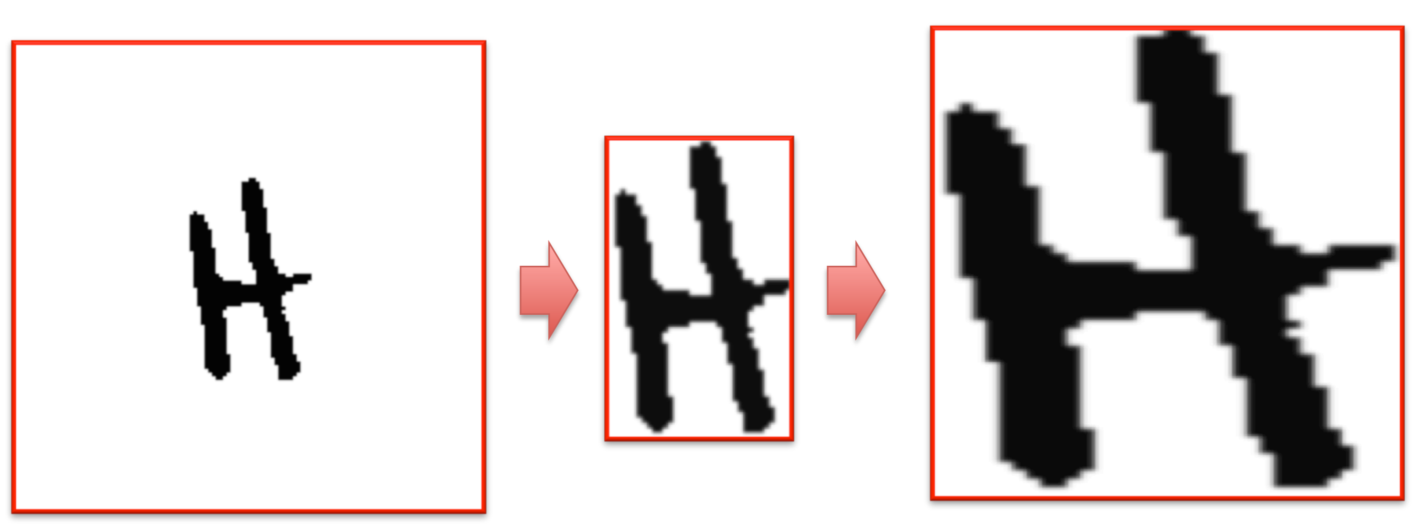
Data & Preprocessing

NIST Database 19 (800,000+ hand-printed samples)

I used 19240 samples (370 samples for each of the 52 upper and lower case English alphabet)

Preprocessing:

- Crop out the central part where the character lies
- Resize it to a standardized size (e.g. 128 × 128 pixels)

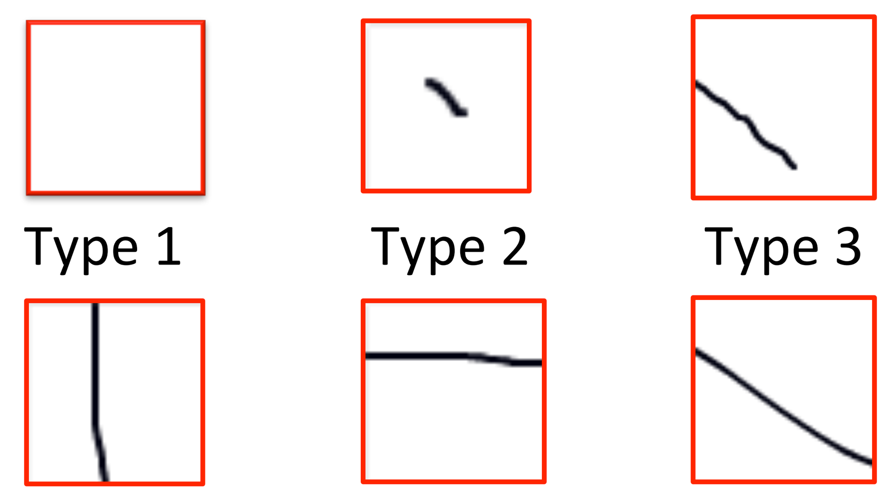


Feature Extraction

- Raw pixels: used as my baseline
- Blackness threshold: an approximation of the original matrix

0	0	1	0	0
1	0	1	1	0
1	1	1	1	1
1	1	0	1	0
0	1	0	1	1

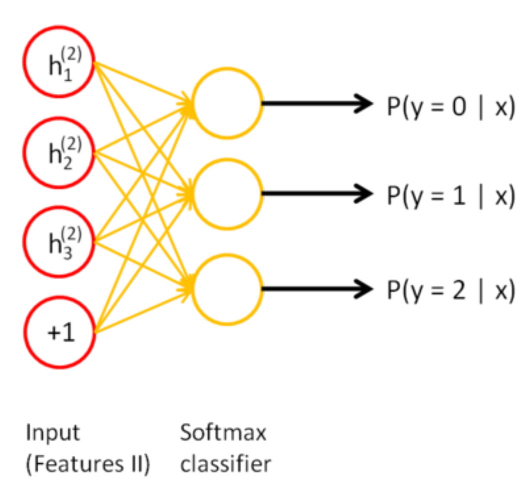
- Blackness percentage: another approximation
- Zoning: put a 3 by 3 grid on top of the original image. Use aspect ratio to classify each grid to six different types



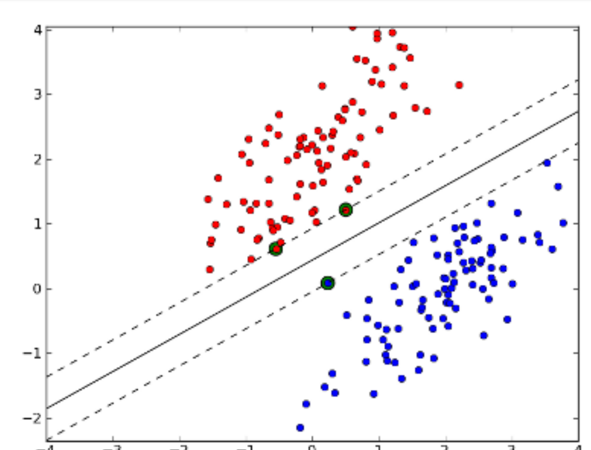
Type 1 Type 2 Type 3

Type 4 Type 5 Type 6

Modeling & Algorithm



Softmax Classification



Support Vector Machine (scikit-learn)

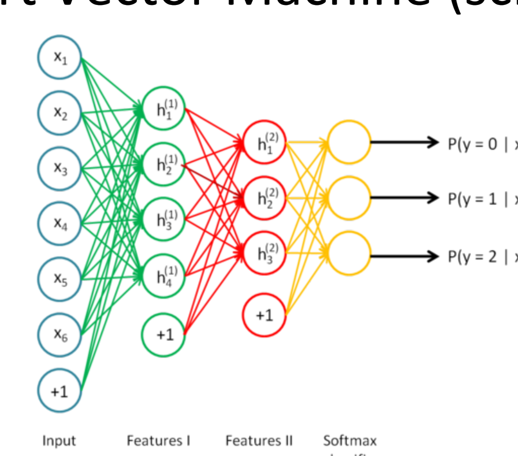
Likelihood Class Prior Probability

$$P(c\mid x)=\frac{P(x\mid c)P(c)}{P(x)}$$

Posterior Probability Predictor Prior Probability

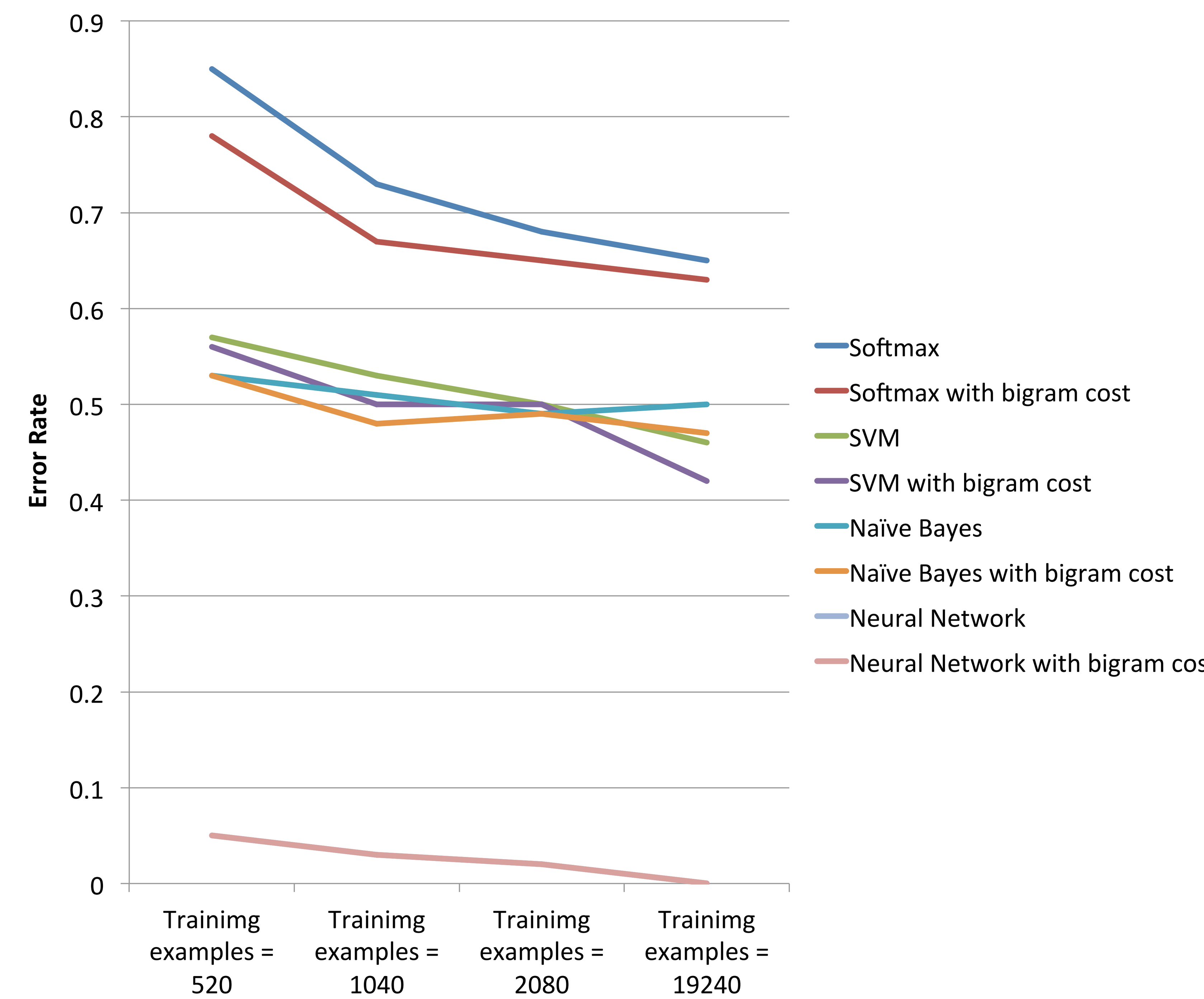
$$P(c\mid X) = P(x_1\mid c)\times P(x_2\mid c)\times\cdots\times P(x_n\mid c)\times P(c)$$

Naïve Bayes (scikit-learn)



Convolutionary Neural Network (PyBrain)

Test Result



Conclusion & Future Work

Conclusion:

- Test error decreases when training data increases
- Convolutionary neural network performs significantly better than other models
- Bigram cost helps improve accuracy

Future Work:

- Find better features to feed in SVM and Naïve Bayes
- Improve performance on 'bottleneck'
- Extend bigram cost method to words, not just characters