Who predicts handwritten digits the best?

KNN, SVM or CNN

Fiona Xu, Amberley Su

Motivation

- Curiosity of lab7 - Why does SVM achieve higher test scores than Random Forests?
- Is there any other model that is good at recognizing handwritten digits?
- Gain deeper understanding of these models.

Data - Semeion Handwritten Digit

n - 1593

p - 256

y - 0-9

training - 80%

validation - 20%

Methods

KNN

SVM

CNN

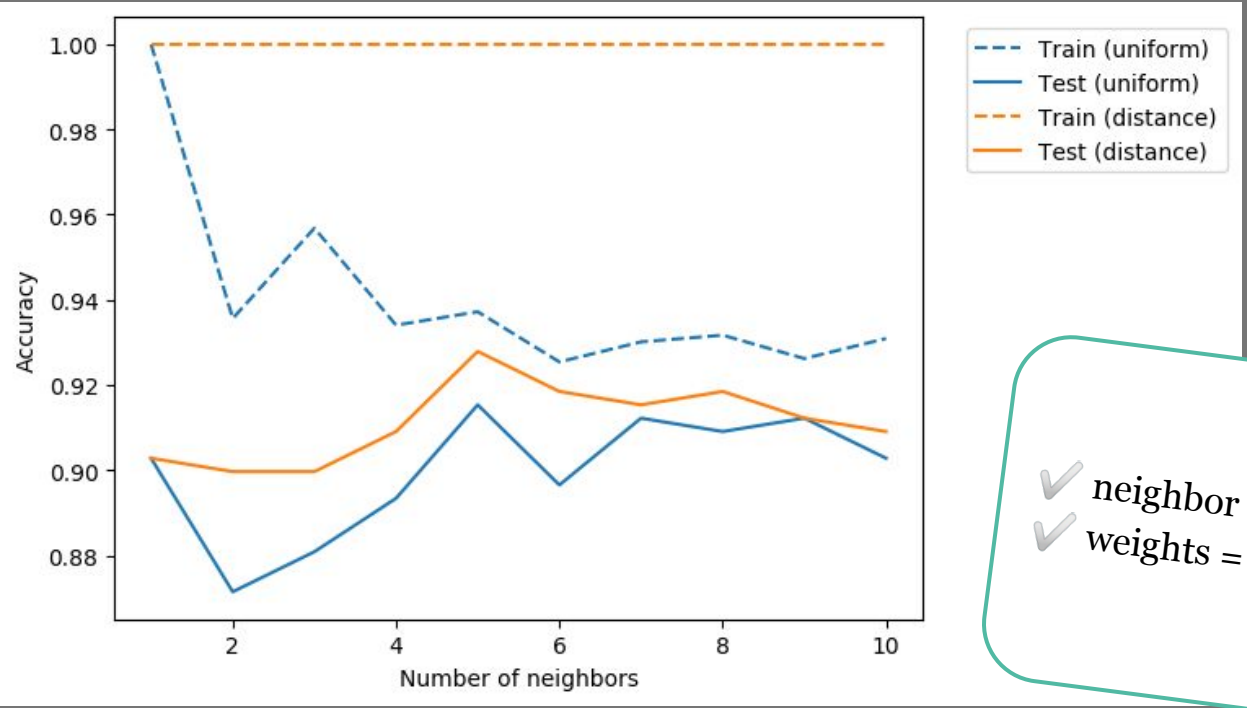
scikits-learn

TensorFlow

– KNN

➔ **Find the best parameters**

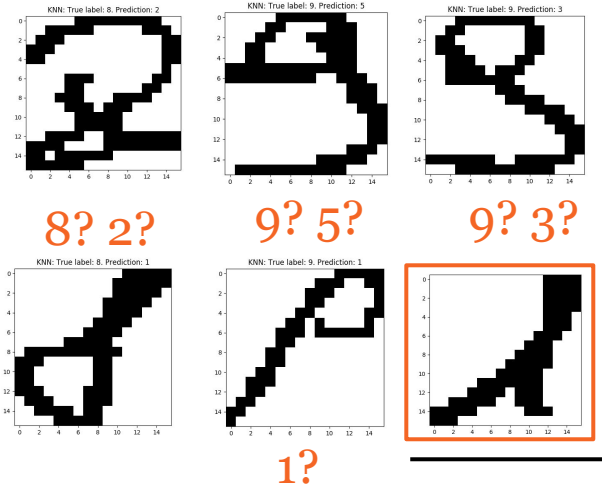
- Neighbors number & weights

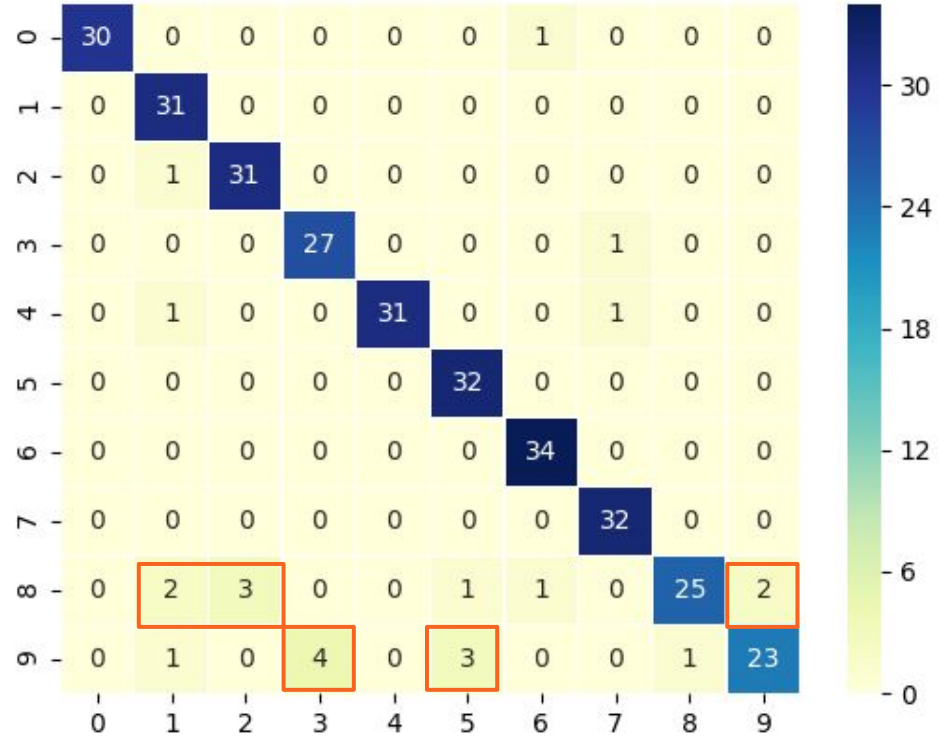


✔ neighbor number = 5
✔ weights = distance

KNN

- Training accuracy: 1.0
- Validation accuracy: 0.9279 (296/319)

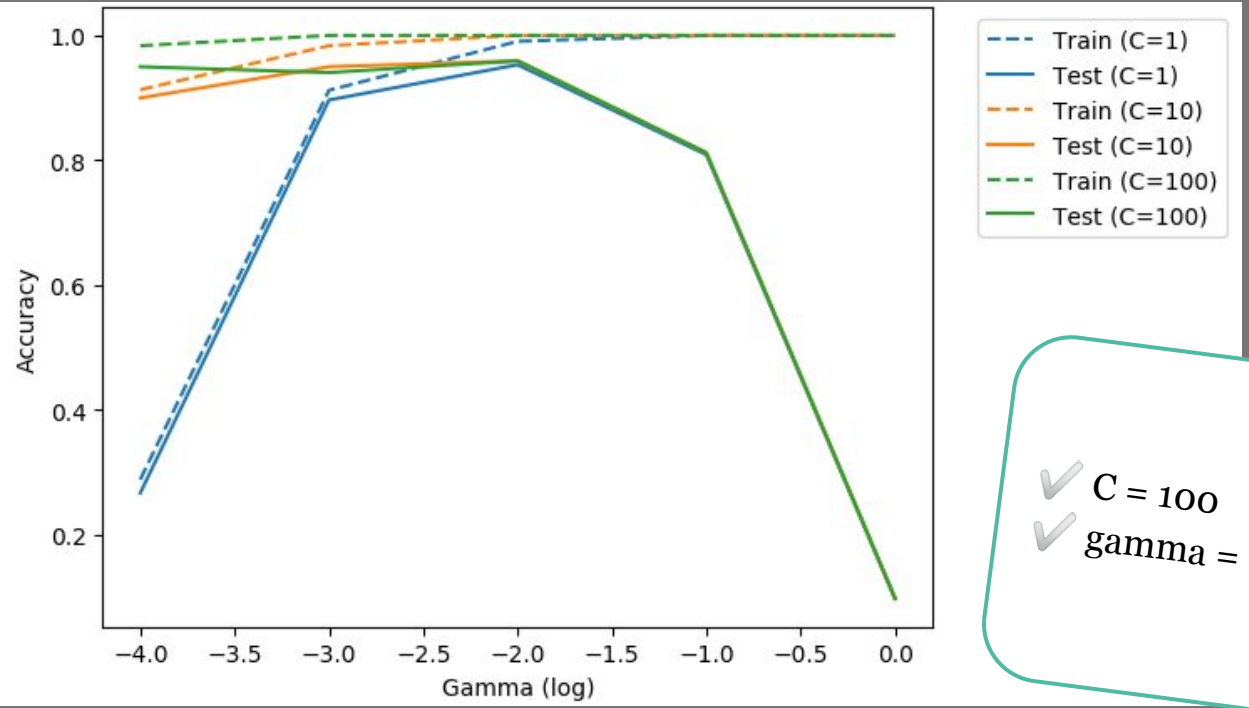




– SVM

➔ **Find the best parameters**

- C & gamma

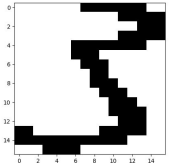


✔ C = 100

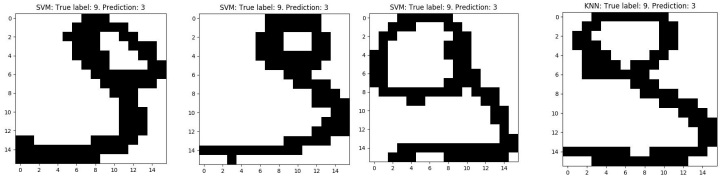
✔ gamma = 0.01

SVM

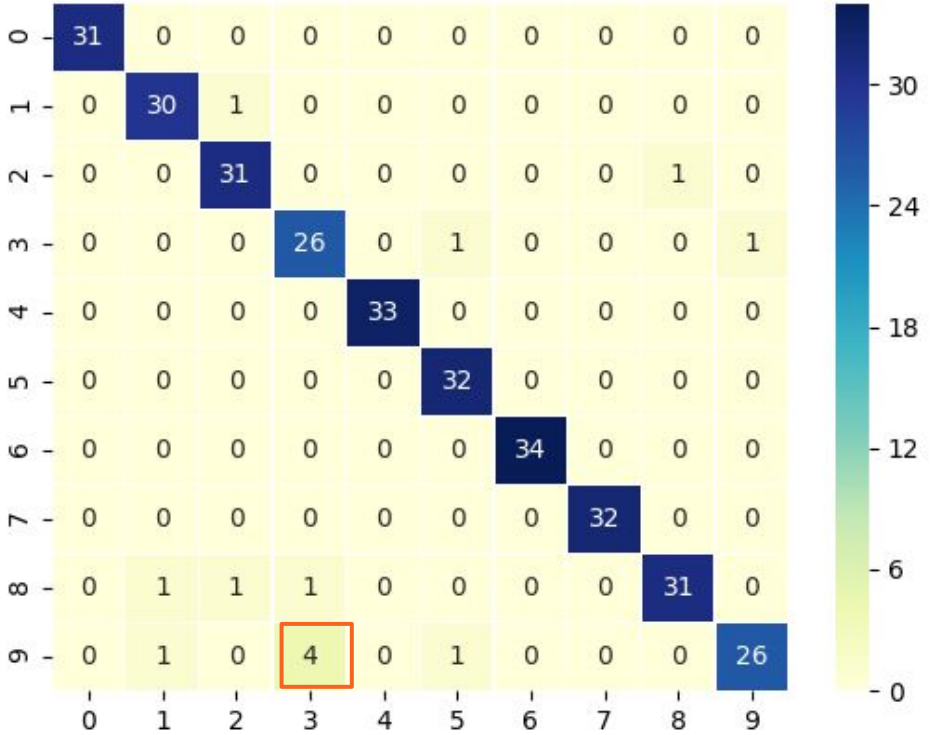
- Training accuracy: 1.0
- Validation accuracy: 0.9592 (306/319)



3 in training dset



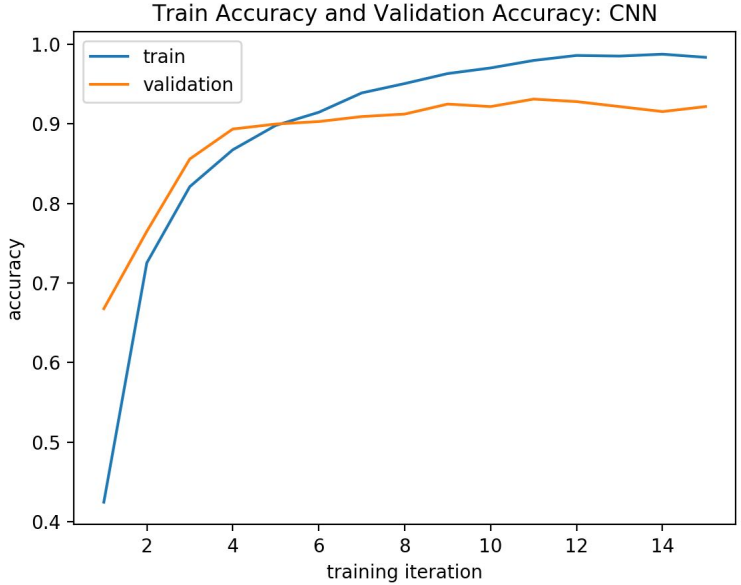
9 in validation dset

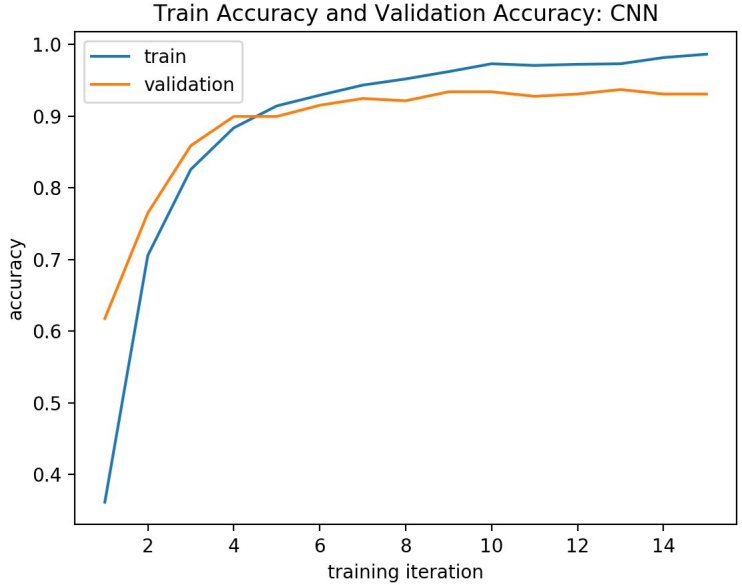


– CNN

➔ **Find the best classification**

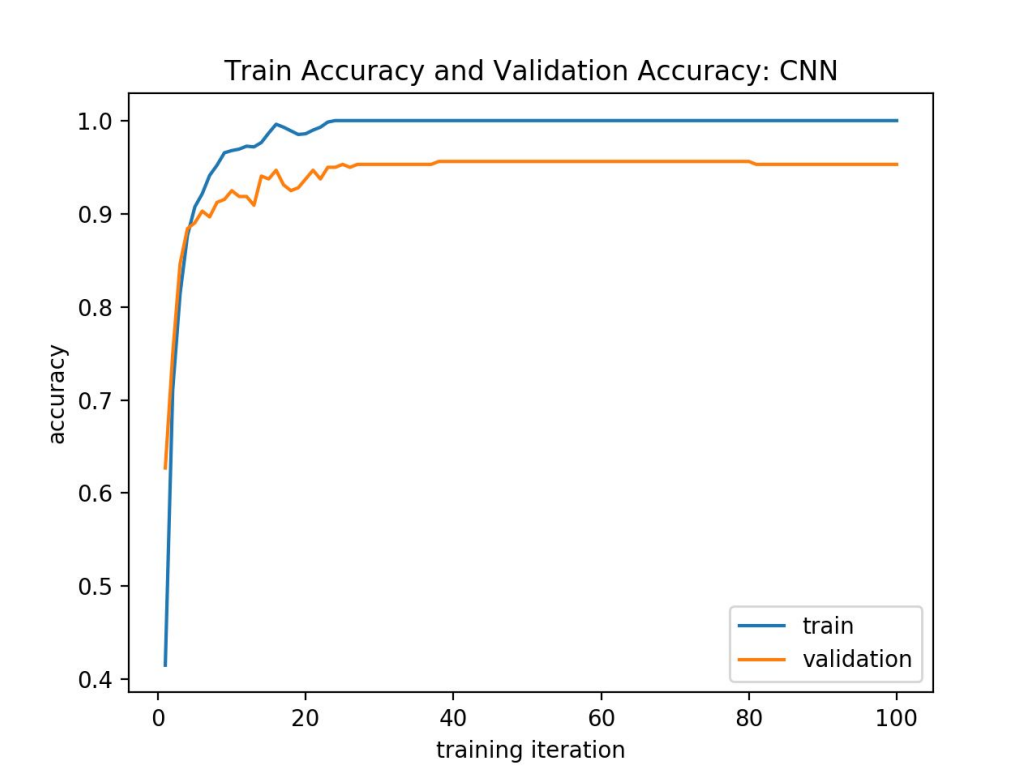
◆ Input ➞ CNN ➞ Output





Best iteration: around 5

– CNN



- Train and Validation Accuracy both plateau in the long term
- Validation Accuracy continues to increase along with Train Accuracy.

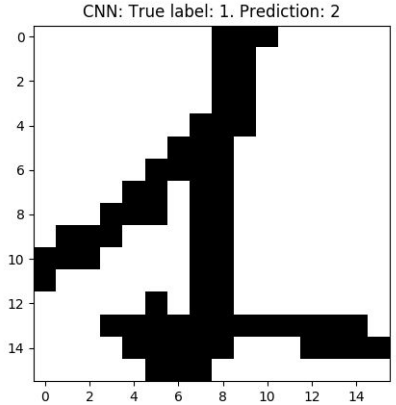
CNN

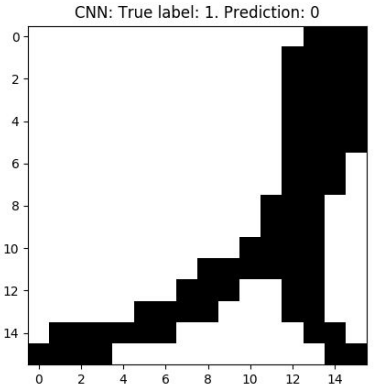
Random results for accuracy

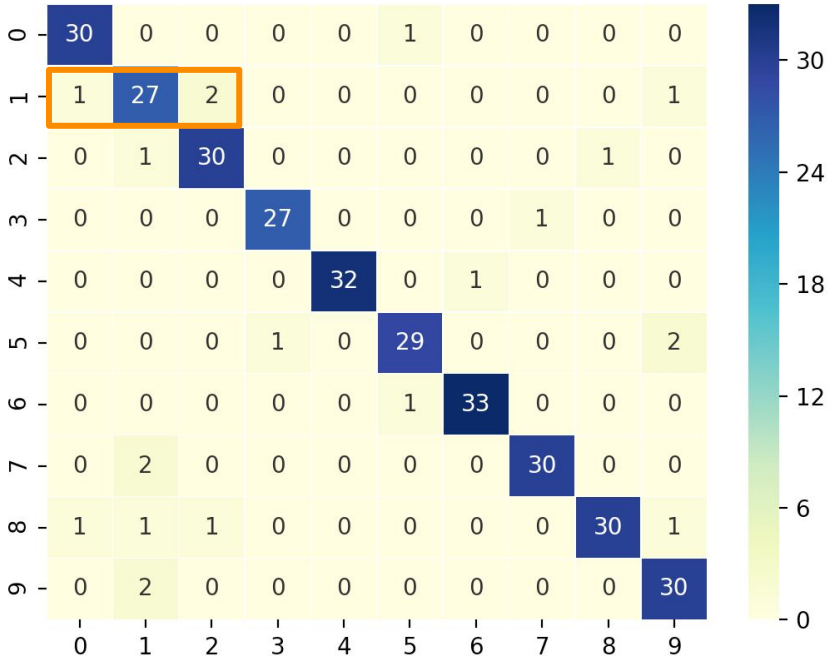
Training data: 97%~100%

Test(Validation) data: 87%~96%

Wrong classifications of CNN:







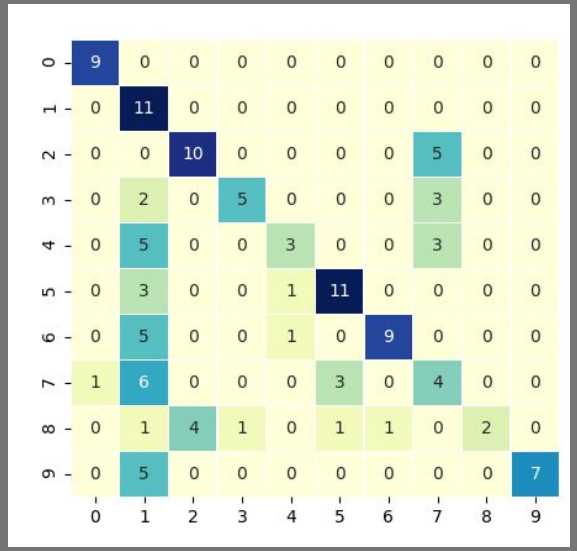
Conclusion

Model	KNN	SVM	CNN
Accuracy	0.9279	0.9592	0.87～0.96

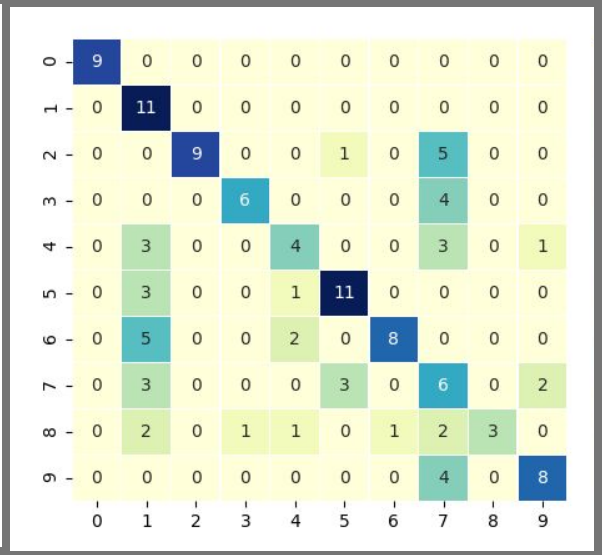
– FUN PART: Reading our own digits!

- ➔ 122 examples!
- ➔ OpenCV

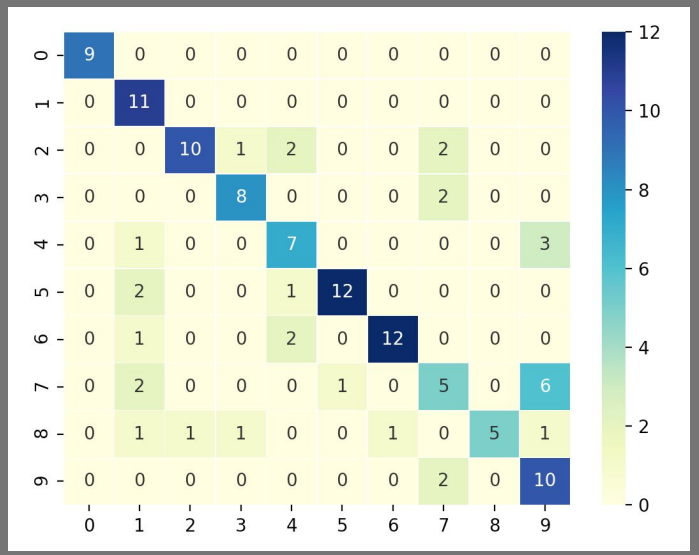
KNN
0.5820



SVM
0.6148

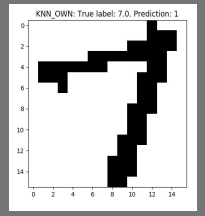


CNN
~0.73

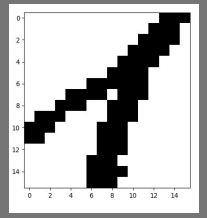


– FUN PART: Reading our own digits!

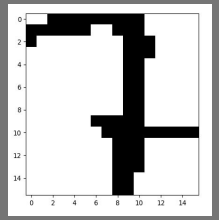
➔ Different handwriting styles



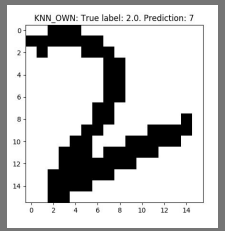
Our 7,
Predicted as 1



Train's 1



Train's 7



Our 2,
Predicted as 7

➔ Future work

- Training with more data!