

**Student Name: Guangyu Lin**

**Collaboration Statement:**

Total hours spent: a week

I discussed ideas with these individuals:

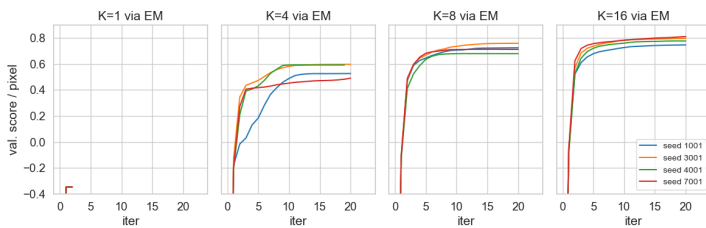
- Mingyang Wu
- TODO
- ...

I consulted the following resources:

- office hour with Professor and TA
- textbook
- ...

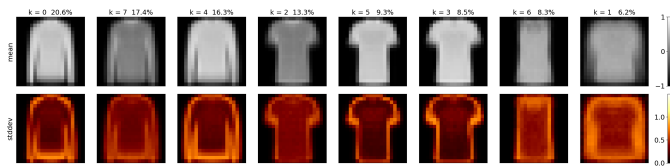
By submitting this assignment, I affirm this is my own original work that abides by the course collaboration policy.

# 1a



Based on the figure, it shows K = 16 did the best on this data.

# 1b



According to figure in 1a, we can know the best score with  $K = 8$  is seed 3001. Based on the figure, we can see when  $k = 0, 7, 4$ , these are long-sleeve, which occupied 54.3% of the whole data. When  $k = 2, 5, 3$ , these are short-sleeve which occupied 31.1% of the entire data. When  $k = 6$  and 1, these are no-sleeve shirts which occupied 14.5% of the entire data.

# 1c

Validation set best score:

k=1	k=4	k=8	k=16
-0.31966905249476085	0.599686761572644	0.7560236877712968	0.8030454076712182

Test set best score:

k=1	k=4	k=8	k=16
-0.34710253107971084	0.5967427374584569	0.7589718286845941	0.8103553227350052