

AMSC/MATH 420, HW7, SPRING 2015

Due: April 7, 2015

Let  $x, y \in \mathbb{R}^d$ , and let

$$dist_p(x, y) = \|x - y\|_p$$

for any  $1 \leq p \leq \infty$ .

Use the handwritten digits data set to construct different graphs associated with different notions of distance as introduced above. One way in which such distance impacts the data analysis is through the notion of nearest neighbors and the associated classification. Is any of the above distances more suitable for the handwritten digit classification problem?

In your solution you must consider at least the following 3 distances:  $dist_1$ ,  $dist_2$ ,  $dist_\infty$ , and at least 2 different values of  $k$  in  $kNN$  classification.