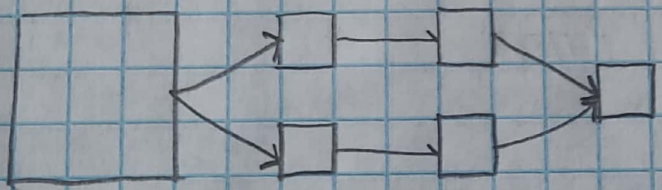
1.

a. With KNN, it is important to clean the data so it doesn't become mislead from irrelevant attributes. PCA can help identify relevant dimensions to reduce this error w/ KNN

b. Kmeans can vary it's grouping when ran multiple times. This can be fixed by selecting several clusterings and picking the one w/ the smallest inertia

c. Gaussian Mixture Modeling assumes instances generated by a mix of gaussian distributions, and works well with identifying non-circular clusters. GMM is good with anomaly detection because it can detect abnormal clusters, and determine which points are not clustered

d.



input layer → CNN input (picture)

Convolution layer → filter that compresses image based on weights

Subsampling layer → Pools and compresses data

neural net → fed outputs from CNN and works as simple NN

Alexnet, VGG16, Resnet are examples

e. Vanishing is when gradients shrink as they move down the layers, creating error. Exploding is the opposite, growing gradients until they diverge. These can be solved by changing activation functions, or by improving weight initialization.

2.

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$.8 \pm 1.96 \sqrt{\frac{.8(.2)}{100}}$$

$$.8 \pm .0784$$

$\alpha = .05$ $\quad (.7216, .8784)$