Home Work -> 5

Chinmay Bhagwat

CWRD -> 20015512

Problem 1)

1) K-nearest neighbors is a ~~far~~ classification of new data points that include ~~few~~ similarity among them. ~~*~~ Such that creates pregnant data points which helps in learning & to draw close & accurate results from its knowledge base.

If we don't reduce the dimensions & remove the irrelevent data features then it will cause ~~vato~~ various issues below.

a) Size ~~the~~ of the data will increse along with the dimensions, which is more difficult to handle.

b) The increse in data size will grow exponential along with dimensions which create the issues in KNN & which reads to much more gap between data poi

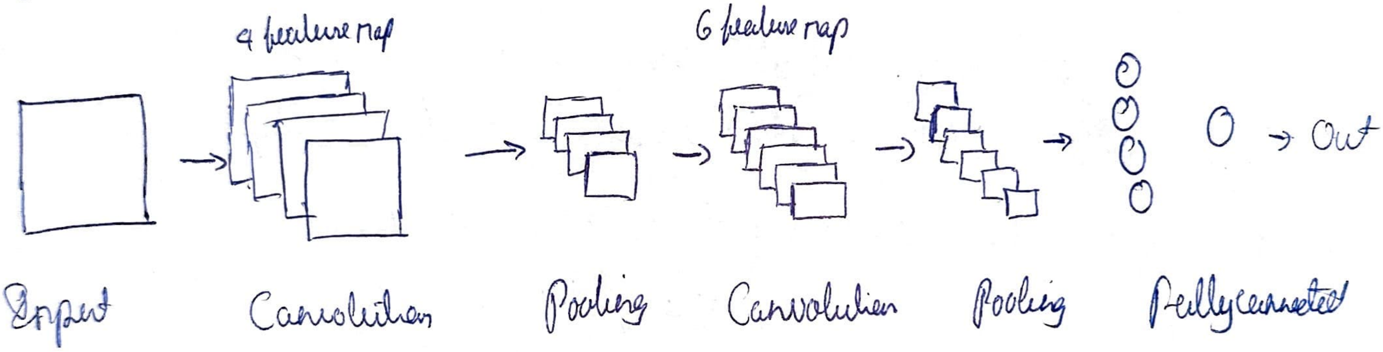
c) Of data points with much gap reads to decrese much more gap reads to decrese in efficiency as KNN used data point to make accurate decisioning.

d) The algorithm will no longer help to take precise decisioning because of the above reasons why this is requirement to remove irrelevent data features & to reduce data dimensions

2) The K-means algorithm works best under certain conditions. These conditio nclude the clusters having equal variances. Without these conditions the algorith ith fail to converge to an optimal solution. Depending on the locations of the d initial centroid, the algorithm may find non ideal clusters. A way to ddress this issue is to adjust the location of the initial centroids. e can run the algorithm several times with different centroids to see which will gives the best solution.

3) A GMM is a probabilistic model that assumes that the instance were generated from the mixture of several Gaussian distributions whose parameters were unknown. We assume that data is grouped into a finite number of cluster, each with an ellipsoidal shape. First random Gaussians are initialized with equal variences. Then the posterior probabilities are evaluated from the current gaussian (E. step). Next new Gaussians are re-estimated using the current responsibilities. The EM steps are repeated until convergence. It is essentially a generalization of K-means. GMM can be estimated using current responsibilities. With GMM, anomalies would be any instance located in a low-density region.

7)



Convolutions: Used to extract features from the input. Convolutions are performed between the input or previous layer. Each portion of the input is calculated. This results in a feature map, which tells us where the features are in the input.

Pooling: Layers that usually follows convolution. The pooling layer looks at each feature map independently, & they scale down the length & the width, keeping the depth intact. This is how this keeps training time & cost as low as possible.

Fully connected :- The last layers of CNN which is for wrap up the architecture of the preceding convolution & pooling layers. Once the input data convolved & scaled down, that result is used as the input to the FC layers which act like NN's.

CNN Algorithms :- AlexNet (2012)

ZFNet (2013)

GoogleNet (2014)

VGGNet (2014)

ResNet (2015).

5) The vanishing gradient problem is when the gradient get smaller & smaller towards lower layer, leaving these layer's weight unchanged, resulting no convergence to good solution.

The exploding gradient problem is similar, except here the gradient grow bigger & bigger. The weight updates become very large, resulting in the algorithm diverging.

- methods to prevent this problem.
  - Improving weight initialization
  - Using non-saturating activation functions
  - ~~Batch~~ Batch ~~A~~ normalization
  - Gradient clipping.

Prablem 2 :-

95 % CI

$= 1-\alpha = 0.95$

$\Rightarrow \alpha/2 = 0.025.$

$Z_{\alpha/2} = Z_{0.025} = 1.96$

Standarderror $E = \sqrt{\frac{E(h) * (1-E(h))}{n}}$

$= \sqrt{\frac{0.8 * (1-0.8)}{100}} = 0.04$

Lower bound $= 0.2 - ((1.96)(004)) = \cancel{0.0}\ 0.1216$

Upper bound $= 0.2 + ((1.96)(0.04)) = 0.2784$

So

$Error_D(h) =$ ~~Error~~ Lower bound $<$ $Error_D(h)$ $<$ Upper bound

$= 0.1216 < Error_D(h) < 0.2784.$