Dimensionality Reduction

Due to the large features dimensions, the training time of the machine learning will be large. Therefore, we performed the dimensionality reduction and reduced the components to 100

Text

Description automatically generated

K means clustering

After the feature extraction, we use K means for the implementation of the headline news clustering task. For K means clustering, the choosing of the value of K is important. We used elbow method to examine the results of different values of k.

Chart, line chart

Description automatically generated

From the above graph, the elbow of curve is at k=7.