Big data can only approximate the whole, but fail to represent the whole data. When there is a huge volume of data, cluster analysis can be tough due to unwanted noise or interference issues (Ref-f913926). The possibility of gathering a large quantity of unwanted data makes data cleaning and padding of missing data necessary.