GMM cannot be model the dataset, the reason is we cannot find the of this dataset. Since the dimension of the independent variables is (60000,28,28). Which means we cannot find the covariance.

One way to apply GMM is to use PCA to reduce the dimension of the dataset, thus the dataset will reduce to 2d, then we can apply the dataset into GMM as the same process to BMM.

Another way is to reduce one dimension by removing one features, although it is not recommended as more information will be lost, resulting a high chance of predicting a bad outcome.

Skimage provides a function called rgb2gray, which can convert a coloured picture into gray, thus one dimension of data will be removed.