1. (1) T-SNE uses a symmetrized version of the SNE cost function with simpler gradients;

(2) It use T-test to compute the similarity between two points in the low-dimensional space, which employs a fatter-tailed distribution to alleviate both the crowding problem and optimization problem.

1. Line 3: random choose 6000 numbers between 1 to 60000, as sample number;

Line 4: use the 6000 sample numbers obtain above as index, to filter out the sample records from the total data.

Line 5: use the 6000 sample numbers to filter out the labels of all the samples.