# Compute the distance between the training set and the test set. How to vectorize?

Supposing we have the test set and the training set , we want to get a distance matrix , where is the distance between the test sample and the training sample , then we have

where  and are row vectors.

Therefore, to vectorize the above computation in numpy, three steps are performed

1. Square each element in and , and then compute the sum for each row. This corresponds and . We get and as the results.
2. Compute for . Obviously, .
3. The last step is to combine them together according to the above equation. We can utilize the broadcast in numpy here. Reshape and . Then, gives the squared distance matrix of size .

A close up of a logo

Description generated with high confidence

(In fact, according to numpy broadcasting rules, there is no need to reshape *B*.)