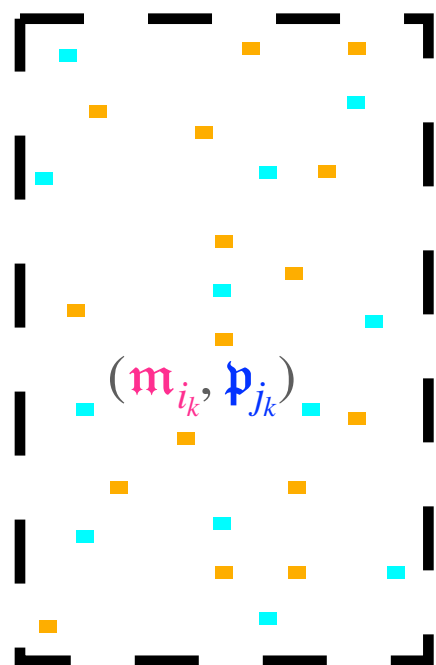


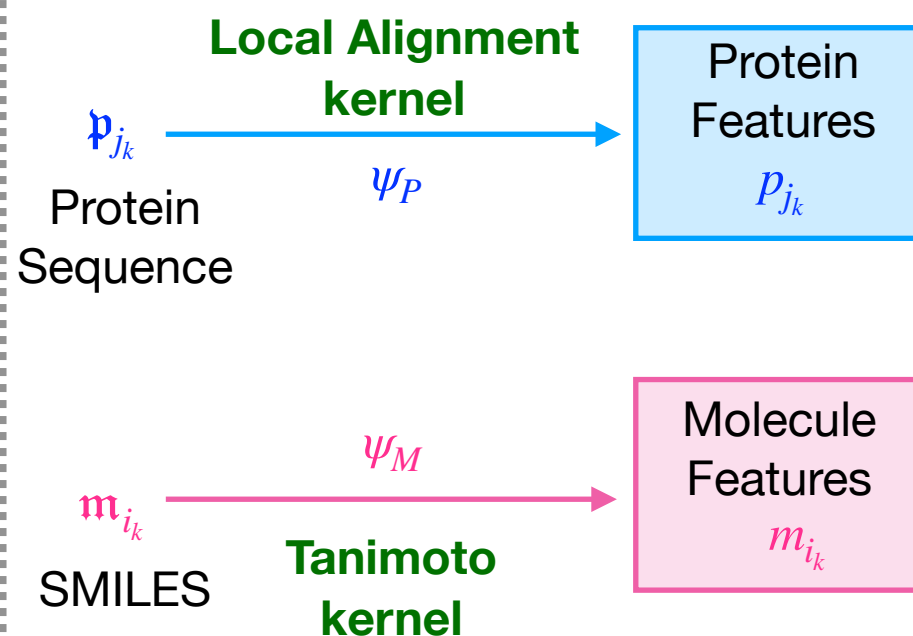


of proteins : 2,060
 # of molecules : 271,180
 # of Positive DTI : 396,798
 # of Negative DTI: 7,965 (+ 388,833 balanced)



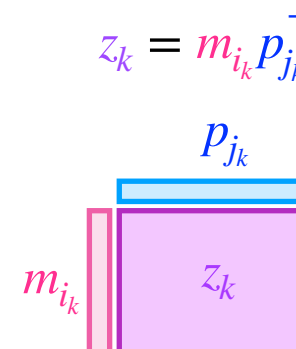
Training dataset

Nystrom approximation Dimension reduction



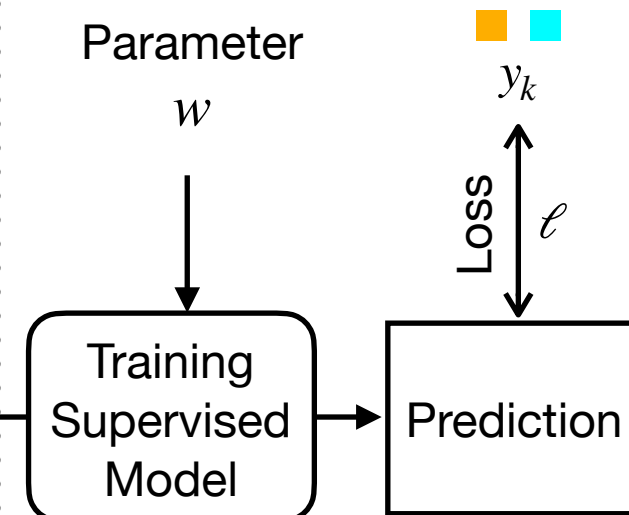
Step 1
Features for
protein and molecule

Kronecker kernel



Step 2
Features for
(molecule, protein) pairs

SVM



Step 3
DTI prediction model