

ChEMBL csv file

Transformed into

Ligands

Erlotinib

SMILES:  
c1cc1nc...

MorganFingerprint featurizer

Sequence:  
GSGWWQL...

OneHotSequence featurizer

$y$  = pIC50

$x$  and  $y$  are  
ML-ready values

kinoml.ml

File

Data Acquisition

Featurization

ML

1

2

3

