

Supporting Information

Permutationally Invariant Deep Learning

Approach to Molecular Fingerprinting with

Application to Compound Mixtures

Andrei Buin,^{*} Hung Yi Chiang, S. Andrew Gadsden,^{*} and Faraz A. Alderson

*College of Engineering and Physical Sciences, University of Guelph, Guelph, Ontario,
Canada, N1G 1Y4*

E-mail: phquanta@gmail.com; gadsden@uoguelph.ca

Figures S1 to S4 portray some of the models compared and used in the Classification and Regression tasks described in Section 3.1 Classification Task with Permutational layer and Section 3.2 Regression Task with Permutationally Invariant Layers. Confusion Matrices for the Neural Fingerprints approach is provided in Figure S5, complementing those already included in Section 3.1 Classification Task with Permutational layer.

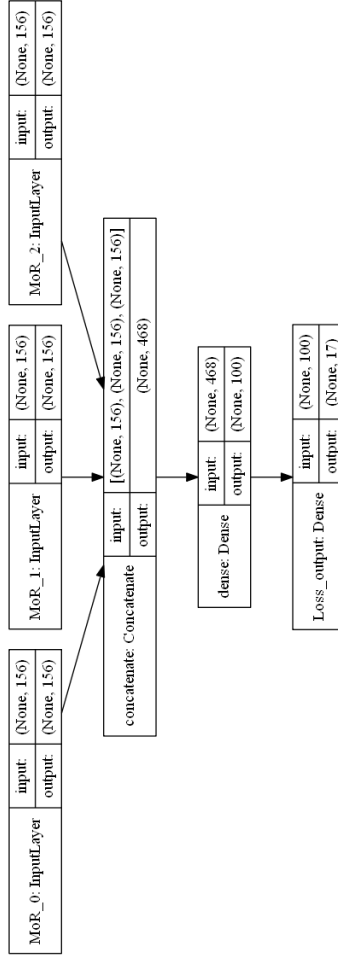


Figure S1: Wei's Original model with Neural Fingerprints. Please note that fingerprint length in this case is 156, however in more general case it does not have to be 156. Length of Seq2Seq FP is 256.

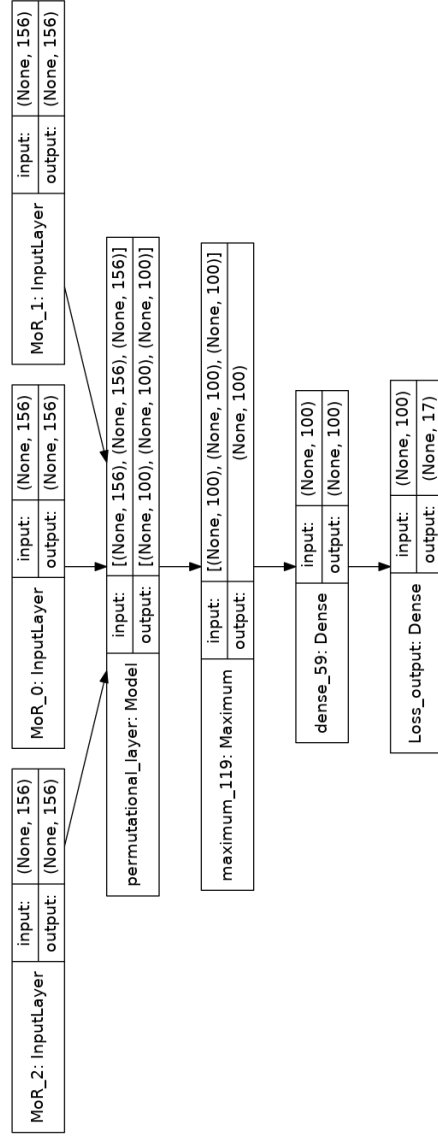


Figure S2: Fingerprints with permutational layers used in modified classification task.

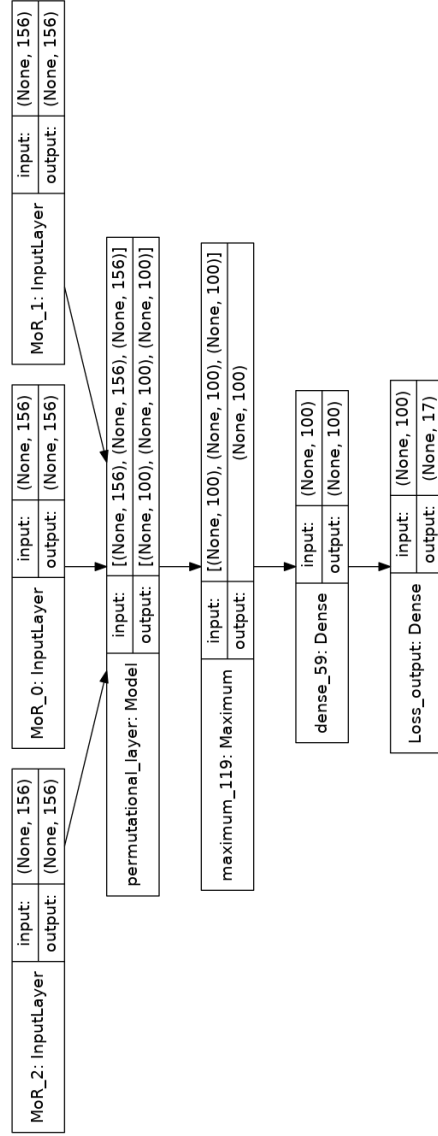


Figure S3: Fingerprints with permutational layers used in modified classification task.

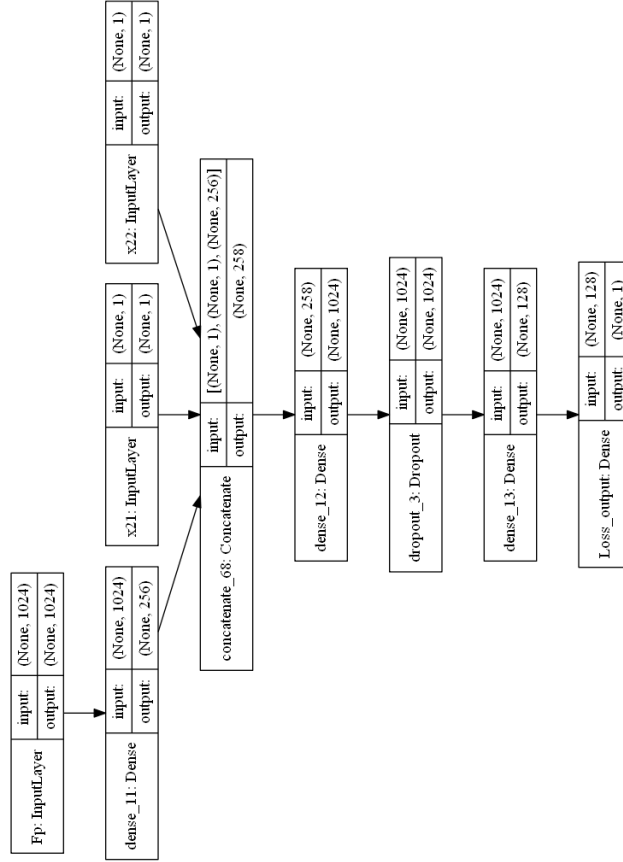
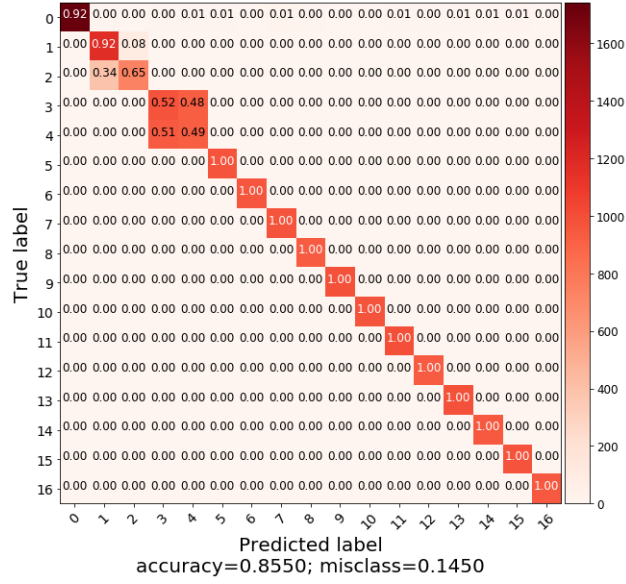
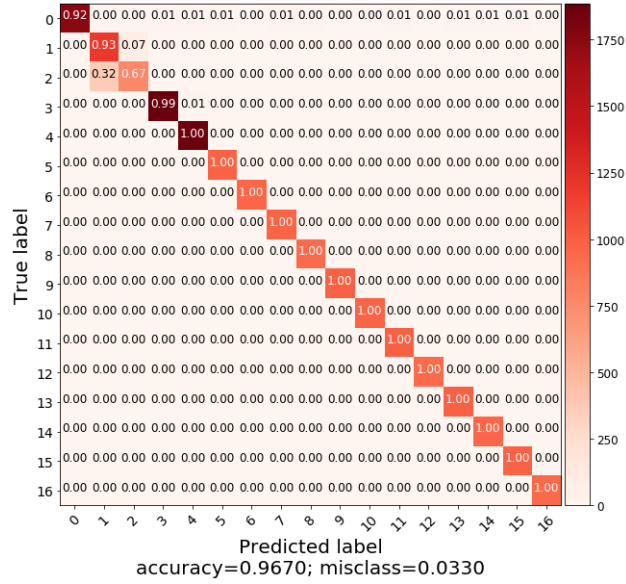


Figure S4: Baseline1 model.



(a)



(b)

Figure S5: Confusion matrices using Neural Fingerprints on test data (a),(b) - Original and $\epsilon(0.1)$ accuracies.