



Learning Deep Feature Representations for Kinase Polypharmacology

Derek Jones^{1,2}, Nathan Jacobs¹, Sally R. Ellingson¹



University of Kentucky¹, Lawrence Livermore National Laboratory²

Abstract:

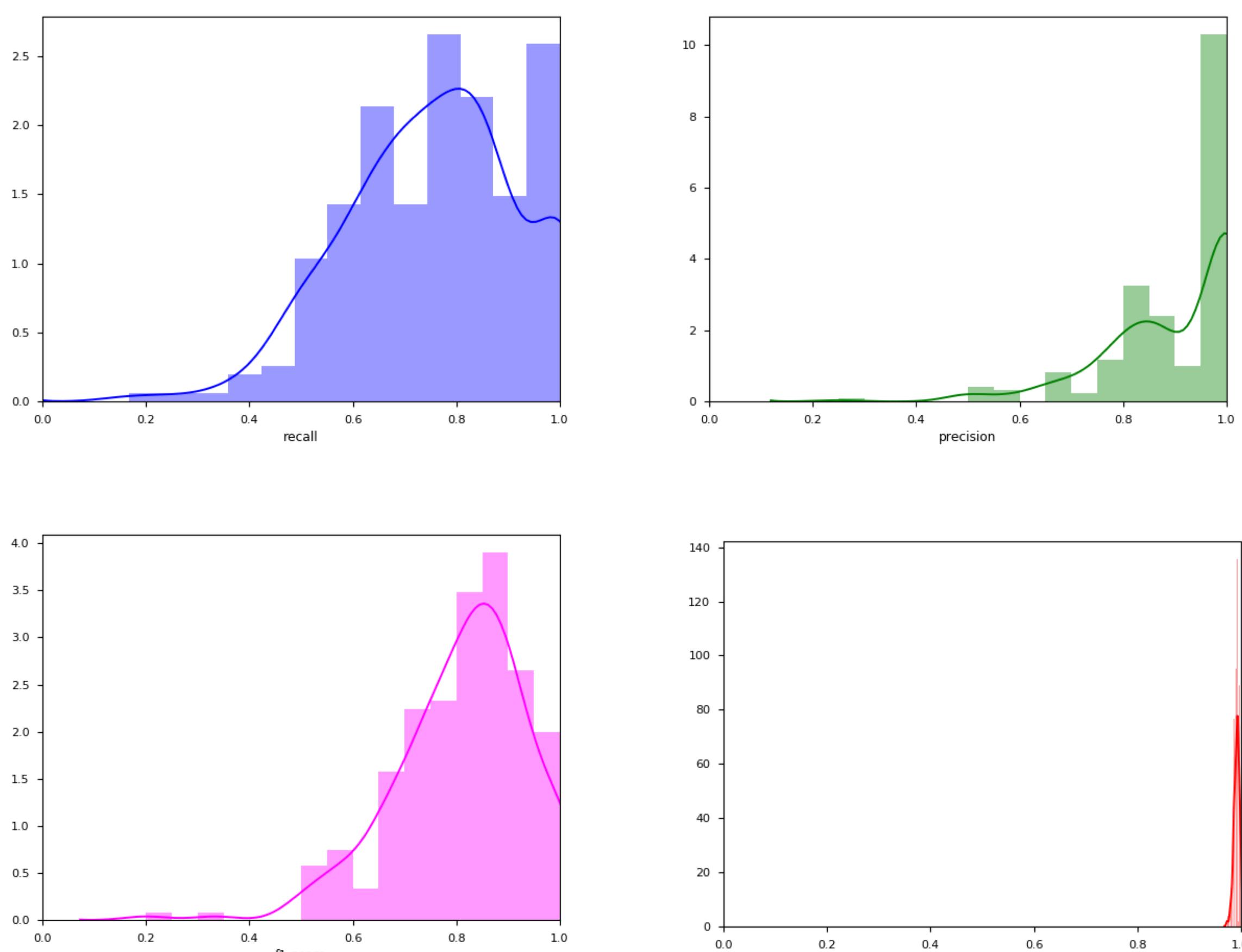
- In traditional ML drug discovery pipeline, use expert knowledge to generate features for drug molecules
- Molecular Descriptors:** typically scalar values computed using rules based upon expert knowledge
- Molecular Fingerprints:** computed using a hash function that encodes structural information as a binary vector. Requires some expert knowledge to choose algorithm parameters
- Goal:** Can learned features be effective in the application of identifying drug molecules that bind to a special class of proteins known as kinases?

Results:

- Experiment 1 (Regression):** Can the MPNN learn to predict molecular properties?
- Experiment 2 (Classification):** Can the MPNN consistently identify which drug molecules bind to a kinase?
- To capture uncertainty (variance) in test predictions, compute metrics over batches and visualize their distribution

Table 1: Molecular property prediction results

Task	Mean R^2	Min R^2	Max. R^2	σ
Hy	0.971	0.953	0.982	0.005
MLogP2	0.937	0.905	0.968	0.01
MLogP	0.961	0.943	0.971	0.004
PDI	0.977	0.969	0.984	0.003
SAacc	0.99	0.983	0.99	0.002
SAdon	0.977	0.96	0.991	0.005
SAtot	0.992	0.99	0.994	0.001
TPSA(No)	0.996	0.991	0.997	0.001
UC	0.993	0.981	0.996	0.002
Ui	0.995	0.991	0.997	0.001
Vvdwmg	0.999	0.998	0.999	0.0002
VVdwzaz	0.998	0.997	0.999	0.0003
Vx	0.999	0.999	0.999	0.0001



Classification metric distributions for testing set

Table 2: Binding affinity prediction results

Metric	Mean	Min.	Max.	σ
Accuracy	0.992	0.973	1.0	0.005
Precision	0.90	0.25	1.0	0.131
Recall	0.757	0.167	1.0	0.167
F1	0.810	0.20	1.0	0.005

Methods:

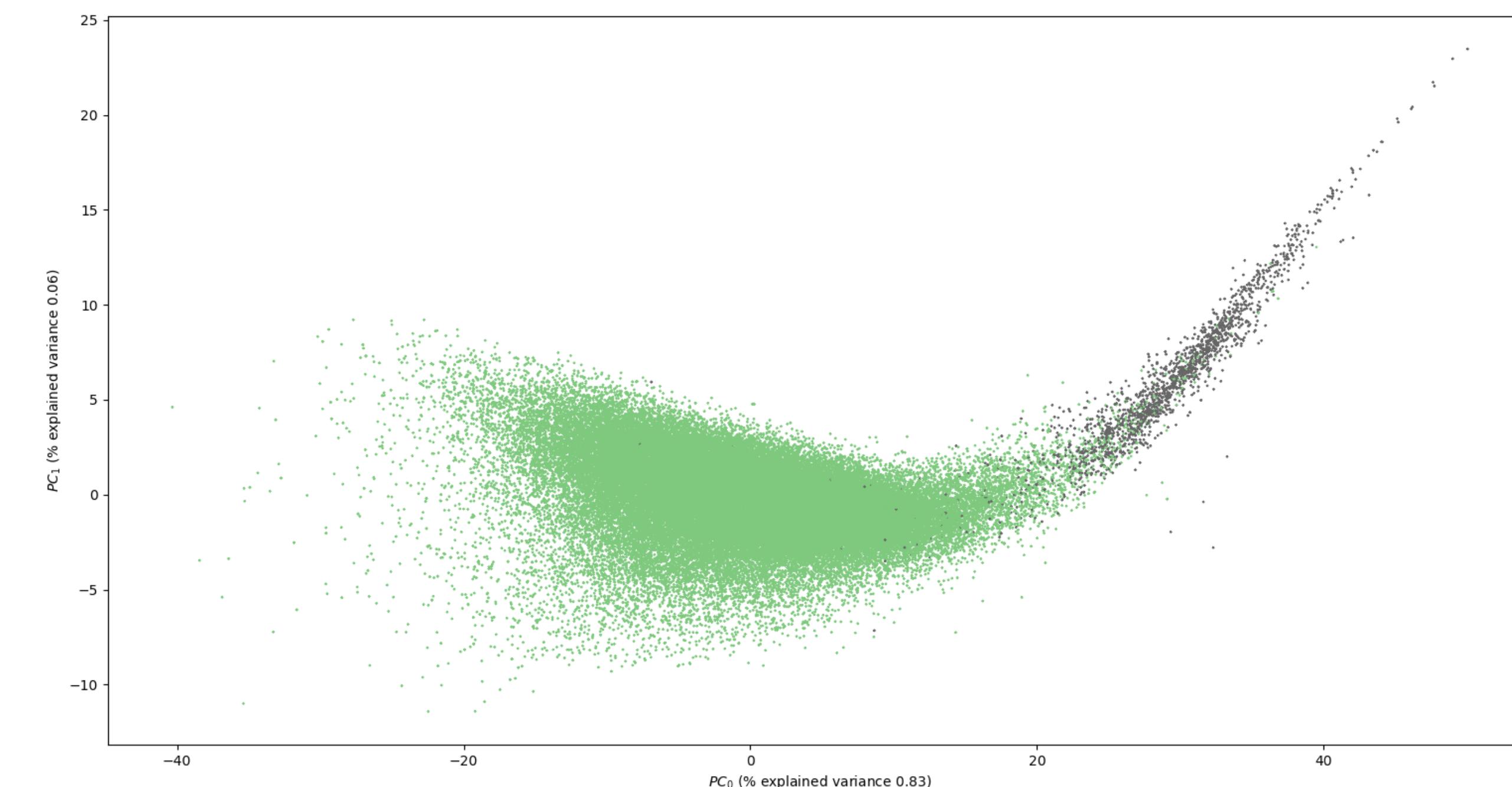
- Data gathered from the kinase subset of the Directory of Useful Decoys Extended (DUD-E)[1]
- Molecular properties extracted using Dragon7 software [2]
- Used message passing neural network (MPNN) as DNN architecture [3],[4]
- Trained networks using asynchronous stochastic gradient descent [5]

$$\text{Message-Passing Phase} \quad m_v^{t+1} = \sum_{w \in N(v)} M_t(h_v^t, h_w^t, e_{vw}) \quad (1)$$

$$\text{Readout Phase} \quad \hat{y} = R(\{h_v^T \mid v \in G\}) \quad (3)$$

$$h_v^{t+1} = U_t(h_v^t, m_v^{t+1}) \quad (2)$$

Discussion:



PCA reduction of hidden layer features on testing set when trained on binding classification task

- Our results show that the MPNN is able to successfully predict the molecular properties of the molecular data
- Further, we show that ability of the MPNN to learn a discriminatory feature for the task of classifying drug molecules that “bind” to a kinase

References:

1. Mysinger, Michael M. et al. “Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking.” *Journal of Medicinal Chemistry* 55.14 (2012): 6582–6594. PMC. Web. 12 Sept. 2018.
2. K. srl. Dragon (software for molecular descriptor calculation) version 7.0.6, 2016. <https://chm.kode-solutions.net>.
3. D. K. Duvenaud, D. Maclaurin, J. Iparragirre, R. Bombarell, T. Hirzel, A. Aspuru-Guzik, and R. P. Adams. Convolutional networks on graphs for learning molecular fingerprints. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems* 28, pages 2224–2232. Curran Associates, Inc., 2015.
4. J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl. Neural message passing for quantum chemistry. Apr. 2017.
5. B. Recht, C. Re, S. Wright, and F. Niu. Hogwild: A Lock-Free approach to parallelizing stochastic gradient descent. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems* 24, pages 693–701. Curran Associates, Inc., 2011.