molecular pharmaceutics

Article

pubs.acs.org/molecularpharmaceutics

Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data

Alexander Aliper,*[,†] Sergey Plis,[‡,§] Artem Artemov,[†] Alvaro Ulloa,[§] Polina Mamoshina,[†] and Alex Zhavoronkov*[,†,∥]

[†]Insilico Medicine, ETC, B301, Johns Hopkins University, Baltimore, Maryland 21218, United States

[‡]Datalytic Solutions, 1101 Yale Boulevard NE, Albuquerque, New Mexico 87106, United States

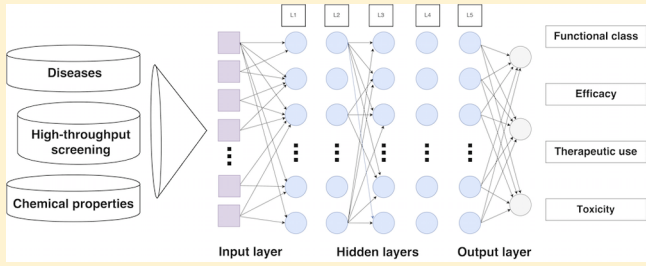
[§]The Mind Research Network, Albuquerque, New Mexico 87106, United States

[∥]The Biogerontology Research Foundation, Oxford, U.K.

深度学习正在迅速推进许多科学和技术领域，在图像、文本、声音和视频识别、机器人技术和自主驾驶方面有着多个成功的事例。本文演示了如何深层神经网络（DNN）训练对大数据集进行分类转录反应不同的药物完全基于他们的转录图谱的治疗类别。我们用678种药物在A549、MCF-7扰动样品，和PC-3细胞从有项目和联系那些12治疗应用类来自网。训练DNN的，我们利用基因的转录组数据和转录组数据使用途径激活的评分算法进行处理，对合并数据集样本的扰动与不同浓度的6和24小时的药物。在通路和基因水平分类，DNN取得了较高的分类精度和令人信服的表现优于支持向量机（SVM）模型对每一个多类分类问题，然而，基于路径的数据进行更好的模型。我们首次展示了深度学习神经网络训练的转录组数据，识别不同的生物系统和环境多种药物的药理特性。我们还提出了利用深层神经网络混淆矩阵进行药物再定位。这项工作证明了将深入学习应用于药物发现和开发的原则。

Supporting Information

ABSTRACT: Deep learning is rapidly advancing many areas of science and technology with multiple success stories in image, text, voice and video recognition, robotics, and autonomous driving. In this paper we demonstrate how deep neural networks (DNN) trained on large transcriptional response data sets can classify various drugs to therapeutic categories solely based on their transcriptional profiles. We used the perturbation samples of 678 drugs across A549, MCF-7, and PC-3 cell lines from the LINCS Project and linked those to 12 therapeutic use categories derived from MeSH. To train the DNN, we utilized both gene level transcriptomic data and transcriptomic data processed using a pathway activation scoring algorithm, for a pooled data set of samples perturbed with different concentrations of the drug for 6 and 24 hours. In both pathway and gene level classification, DNN achieved high classification accuracy and convincingly outperformed the support vector machine (SVM) model on every multiclass classification problem, however, models based on pathway level data performed significantly better. For the first time we demonstrate a deep learning neural net trained on transcriptomic data to recognize pharmacological properties of multiple drugs across different biological systems and conditions. We also propose using deep neural net confusion matrices for drug repositioning. This work is a proof of principle for applying deep learning to drug discovery and development.



KEYWORDS: *deep learning, DNN, predictor, drug repurposing, drug discovery, confusion matrix, deep neural networks*

INTRODUCTION

Drug discovery and development is a complicated and time and resource consuming process, and various computational approaches are regularly being developed to improve it. *In silico* drug discovery[1,2] has evolved over the past decade and offers a targeted, efficient approach compared to those of the past, which often relied on either identifying active ingredients in traditional remedies or, in many cases, serendipitous discovery. Modern methods include data mining, structure modeling (homology modeling), traditional machine learning[3] (ML), and its biologically inspired branch technique, deep learning (DL).[4]

DL[4] methods model high-level representations of data using deep neural networks (DNNs). DNNs are flexible multilayer systems of connected and interacting artificial neurons that perform various data transformations. They have several hidden layers of neurons, which number variation allows adjusting the level of data abstraction. DL now plays a dominant role in the areas of physics[5] and speech, signal, image, video, and text mining and recognition,[6] improving state of the art performances by more than 30%, where the prior decade struggled to obtain 1−2% improvements. Traditional machine learning approaches have achieved significant levels of classification accuracy, but at the price of manually selected and tuned features. Arguably, feature engineering is the dominating research component in practical applications of ML. In contrast, the power of NNs is in automatic feature learning from massive data sets. Not only does it simplify manual and laborious feature engineering but also it allows learning task-optimal features.

Modern biology has entered the era of Big Data, wherein data sets are too large, high-dimensional, and complex for classical computational biology methods. The ability to learn at

Received: March 18, 2016
Revised: May 13, 2016
Accepted: May 20, 2016
Published: May 20, 2016

DOI: 10.1021/acs.molpharmaceut.6b00248
Mol. Pharmaceutics 2016, 13, 2524−2530

药物发现和开发是一个复杂的时间和资源消耗过程。各种计算方法正在不断地发展以改进它。在硅片药物discovery1,2发展在过去的十年中，提供有针对性的，有效的方法，相比过去，这往往依赖于确定活性成分的中草药，在很多情况下，意外的发现。现代方法包括数据挖掘、结构建模（同源建模）、传统机器学习。（ML），其生物启发分支技术，深入学习（DL）4。4方法模型使用深度神经网络（DNNs）高层表示数据。DNNs的连接和相互作用的人工神经元，进行各种数据变换灵活的多层系统。它们有几个隐藏的神经元层，数字变化允许调整数据抽象级别。DL在物理和语音、信号、图像、视频领域起着主导作用，而文本挖掘和识别，6改善国家的艺术表演超过30%，在之前的十年努力获得1−2%改进。传统的机器学习方法已经实现了显著的分类精度，但以手工选择和调整的特征为代价。可以说，特征工程是主要研究构件在ml对比实际应用中，神经网络的电力是从大量的数据中自动学习功能。它不仅简化了手动和繁琐的工程特征也可以学习taskoptimal特征。现代生物学已进入大数据时代，其中数据集太大，高维，复杂的经典计算生物学方法。在更高的抽象层次上学习的能力使DL成为一个有前途的和有效的工具，用于生物和化学数据。7。使用DL体系结构能够处理稀疏和复杂信息，这是高维基因表达数据分析中特别需要的。“维数灾难”是一种基因表达数据，可以通过特征选择实施标准数据的投影方法PCA或更多的生物相关的途径分析解决的主要问题。8 DNNs展示国家的最先进的性能特征的提取数据稀疏转录（mRNA和miRNA数据），9的癌症分类使用基因表达data10预测剪接代码patterns.11 DL已被有效地应用于biomodeling和结构基因组学预测蛋白质序列的蛋白质的三维结构（有序或无序蛋白（与固定的三维结构缺乏））12、13和可能成为成功实施预测药物−靶相互作用的新drugs.14 DL方法发展的一个重要工具，15分子模型反应性能，16和计算毒性drugs.17深网结合生物学的更多特性，18的应用广度和准确性可能会增加。药物再利用或目标扩展允许新的潜在应用药物或甚至新的治疗类药物使用基因表达数据的预测治疗前和治疗后（例如，在

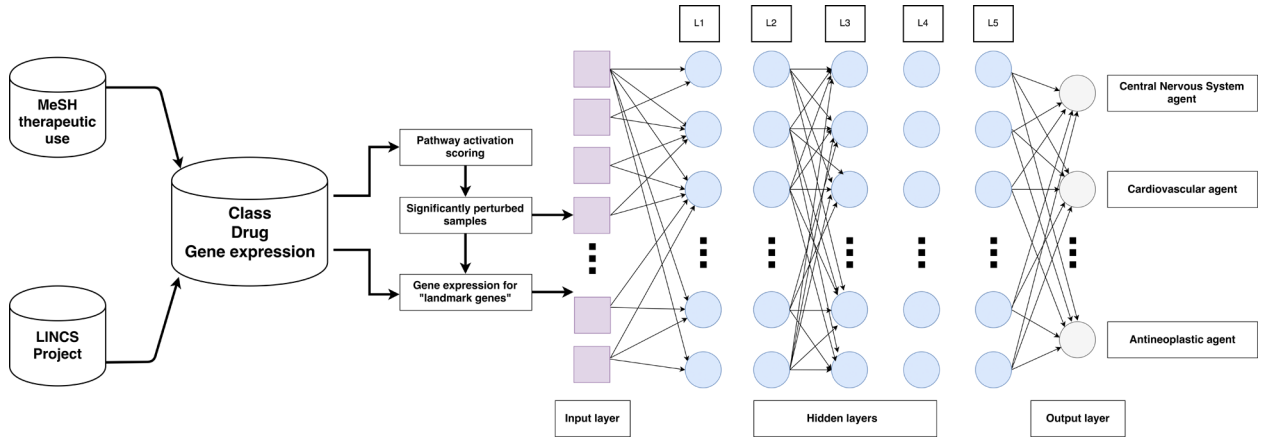


Figure 1. Study design. Gene expression data from LINCS Project was linked to 12 MeSH therapeutic use categories. DNN was trained separately on gene expression level data for “landmark genes” and pathway activation scores for significantly perturbed samples, forming input layers of 977 and 271 neural nodes, respectively.

the higher levels of abstraction made DL a promising and effective tool for working with biological and chemical data.[7] Methods using DL architecture are capable of dealing with sparse and complex information, which is especially demanded in the analysis of high-dimensional gene expression data. “Curse of dimensionality” is one of the major problems of gene expression data that can be solved by feature selection implementing standard data projection methods as PCA or more biologically relevant as pathway analysis.[8] DNNs demonstrate the state-of-the-art performance extracting features from sparse transcriptomics data (both mRNA and miRNA data),[9] in classifying cancer using gene expression data[10] and predicting splicing code patterns.[11] DL has been effectively applied in biomodeling and structural genomics to predict protein 3-D structure using protein sequence (ordered or disordered protein (with lack of fixed 3-D structure))[12,13] and may become an essential tool for development of new drugs.[14] DL approaches were successfully implemented to predict drug−target interactions,[15] model reaction properties of molecules,[16] and calculate toxicity of drugs.[17] As deep networks incorporate more features from biology,[18] application breadth and accuracy will likely increase.

Drug repurposing or target extension allows prediction of new potential applications of medications or even new therapeutic classes of drugs using gene expression data before and after treatment (e.g., before and after incubation of a cell line with multiple drugs). There are multiple *in silico* approaches to drug discovery and classification,[19−21] and many attempts were made to predict transcriptional response with functional properties of drugs.[22−24] In this study we addressed this problem by classifying various drugs to therapeutic categories with DNN solely based on their transcriptional profiles. We used the perturbation samples of X drugs across A549, MCF-7, and PC-3 cell lines from the LINCS Project and linked those to 12 therapeutic use categories derived from MeSH therapeutic use section (Figure 1). After that we independently used both gene expression level data for “landmark genes” and pathway activation scores to train DNN classifier.

RESULTS

The main aim of this study was to apply and estimate the accuracy of DL methods to classify various drugs to therapeutic categories solely based on their transcriptional profiles. In total, we analyzed 26,420 drug perturbation samples for three cell lines from the Broad LINCS database. All samples were assigned to 12 specific therapeutic use categories according to MeSH classification of the particular drug (Supplementary Table 1). Since a number of drugs were present in multiple categories, we considered only those drugs that belong only to one category. To increase the number of samples in each of the categories and to make the classification more robust for each given drug, we aggregated all samples corresponding to all possible perturbation time, perturbation concentration, and cell line parameters (Supplementary Table 2).

When dealing with transcriptional data at the gene level, a common problem is the so-called “curse of dimensionality”. Indeed, when we applied DNN on gene level data for whole data set of 12,797 genes, it did not perform very well, achieving only 0.24 mean F1 score on 12 classes. So our first step was proper feature selection. Here we investigated two approaches: pathway activation scoring and using “landmark genes” as new features.

Pathway Level. For pathway level analysis we used a previously established pathway analysis method called OncoFinder.[25−30] It preserves biological function and allows for dimensionality reduction at the same time. In contrast to other pathway analysis tools, which mostly implement pathway enrichment analysis, OncoFinder performs quantitative estimation of signaling pathway activation strength, and the sign of the resulting value indicates how significantly the pathway is up- or downregulated. All perturbation samples were analyzed with this tool and for each sample we calculated pathway activation profiles for 271 signaling pathways. Samples with zero pathway activation score for all of the pathways were considered as insignificantly perturbed and were excluded from further analysis. That resulted in a final data set containing 308, 454, and 433 drugs for A549, MCF7, and PC3 cell lines, respectively, and totalling 9352 samples (Supplementary Table 2).

Using this data set we built a deep learning classifier based only on pathway activation scores for drug perturbation profiles of 3 cell lines: A549, MCF-7, and PC-3. Making a classifier based on a pooled data set with different cell lines, drug concentration, and perturbation time, we are able to estimate the classification performance in recognizing complex drug

DOI: 10.1021/acs.molpharmaceut.6b00248
Mol. Pharmaceutics 2016, 13, 2524−2530

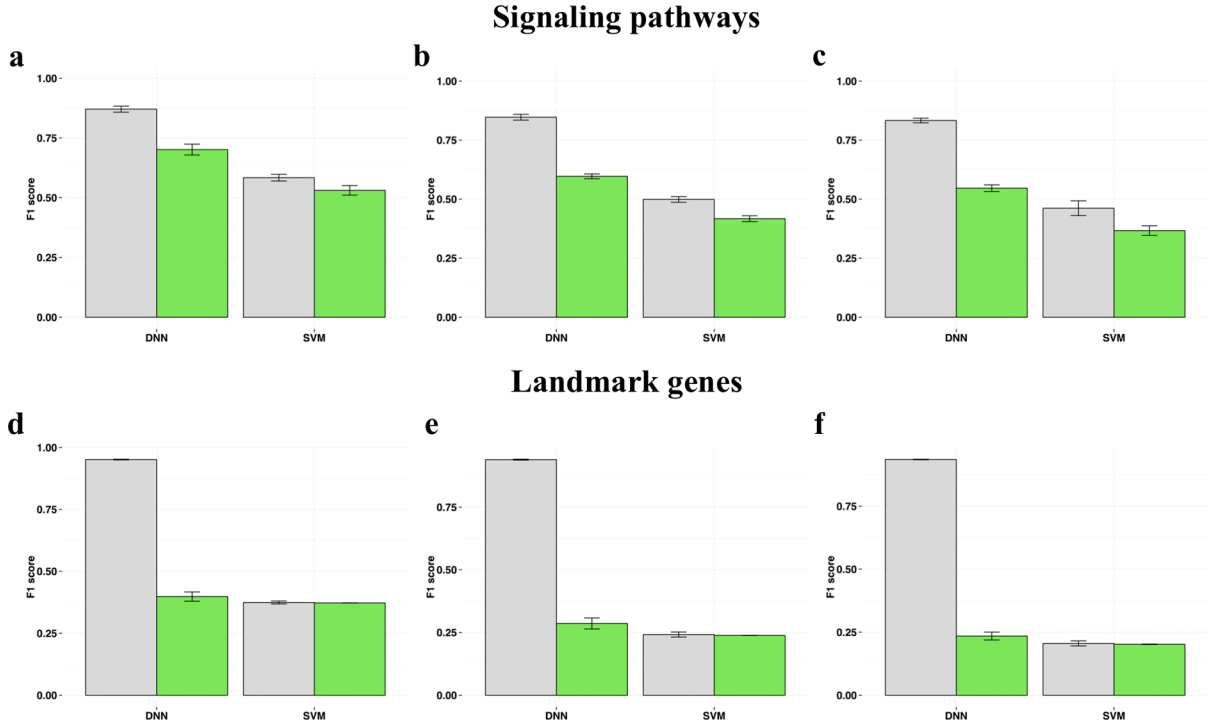


Figure 2. Classification results. Classification performance of DNN and SVM trained on signaling pathways (a, b, c) and landmark genes (d, e, f) for 3, 5, and 12 drug classes, respectively, after 10-fold cross validation. Training and validation set results are shown in gray and green colors, respectively.

action patterns across different biological conditions. For the 3-class classification problem we chose the most abundant categories: antineoplastic, cardiovascular, and central nervous system agents. DNN achieved 10-fold cross-validation mean F1 score of 0.701. We compared the results of DNN to another popular classification algorithm called support vector machine (SVM) trained via nested 3-fold cross validation for several hyperparameters (see Materials and Methods). On 3-class classification problem SVM performed with mean F1 score of 0.530.

Addition of gastrointestinal and anti-infective classes decreased the mean F1 score of DNN to 0.596. Mean F1 score for SVM dropped as well, down to 0.417.

When all 12 classes were considered, the classification neural performance decreased in a minor way, with a cross-validation mean F1 score of 0.546. SVM performed with cross-validation mean F1 score of 0.366 on the same 12-class classification problem. The performance comparison of DNN and SVM on investigated classification problems is depicted in Figure 2a–c. These results indicate that our model performance far exceeds random chance,[31] and we can conclude that DNN outperformed SVM on every multiclass classification problem.

Landmark Gene Level. In our second feature selection approach we used a data set containing normalized gene expression data for 977 "landmark genes". According to the authors of LINCS Project they can capture approximately 80% of the information and possess great inferential value. For fair comparison we trained DNN exactly the same way we did on the pathway level. We used the same data set of 9352 significantly perturbed samples and tested the performance of DNN on the same classification problems. DNN trained on "landmark gene" data performed with 10-fold cross-validation mean F1 scores of 0.397, 0.285, and 0.234 for 3-, 5-, and 12-class classification tasks, respectively. The SVM model showed mean F1 scores of 0.372, 0.238, and 0.202 for respective tasks (Figure 2d–f).

DNN as Drug Repurposing Tool. Here we tried to dig a bit deeper into classification results on pathway level, since the DNN model worked best with pathways as features. To determine which of the 12 therapeutic use categories are the most detectable by DNN, we calculated 10-fold cross-validation classification accuracy of each category. Antineoplastic agents turned out to be the most "recognizable" category by a large margin, with 0.686 accuracy on 12 classes. This was followed by anti-infective, central nervous system, and dermatologic categories, with 0.513, 0.506, and 0.505 accuracy, respectively. The least "recognizable" on the same number of classes was hematologic agents, with accuracy of 0.23. On 3- and 5-class classification problems, the category antineoplastic drugs was on top as well, with accuracy of 0.82 and 0.742. Separability of therapeutic categories by DNN can be illustrated with confusion matrices (Figure 3). Here we observed that the cardiovascular category drugs was relatively often misclassified as central nervous system and antineoplastic agents. In contrast, the level of false positives for the antineoplastic category was relatively small. If we look even closer into the results, sometimes these misclassified false positive drugs may in fact represent a possibility for drug repurposing. For instance, well-known muscarinic receptor antagonist otenzepad was misclassified as central nervous system agent, but despite its obvious role in brain function,[32,33] according to the MeSH therapeutic use section, it is only used against cardiac arrhythmia. Another example includes vasodilator pinacidil, a cyanoguanidine drug that opens ATP-sensitive potassium channels, which was misclassified as central nervous system agent in several cross-validation iterations, although it is used only in cardiovascular conditions. It is known that potassium channels play important roles in different brain regions,[34] and

DOI: 10.1021/acs.molpharmaceut.6b00248
Mol. Pharmaceutics 2016, 13, 2524–2530

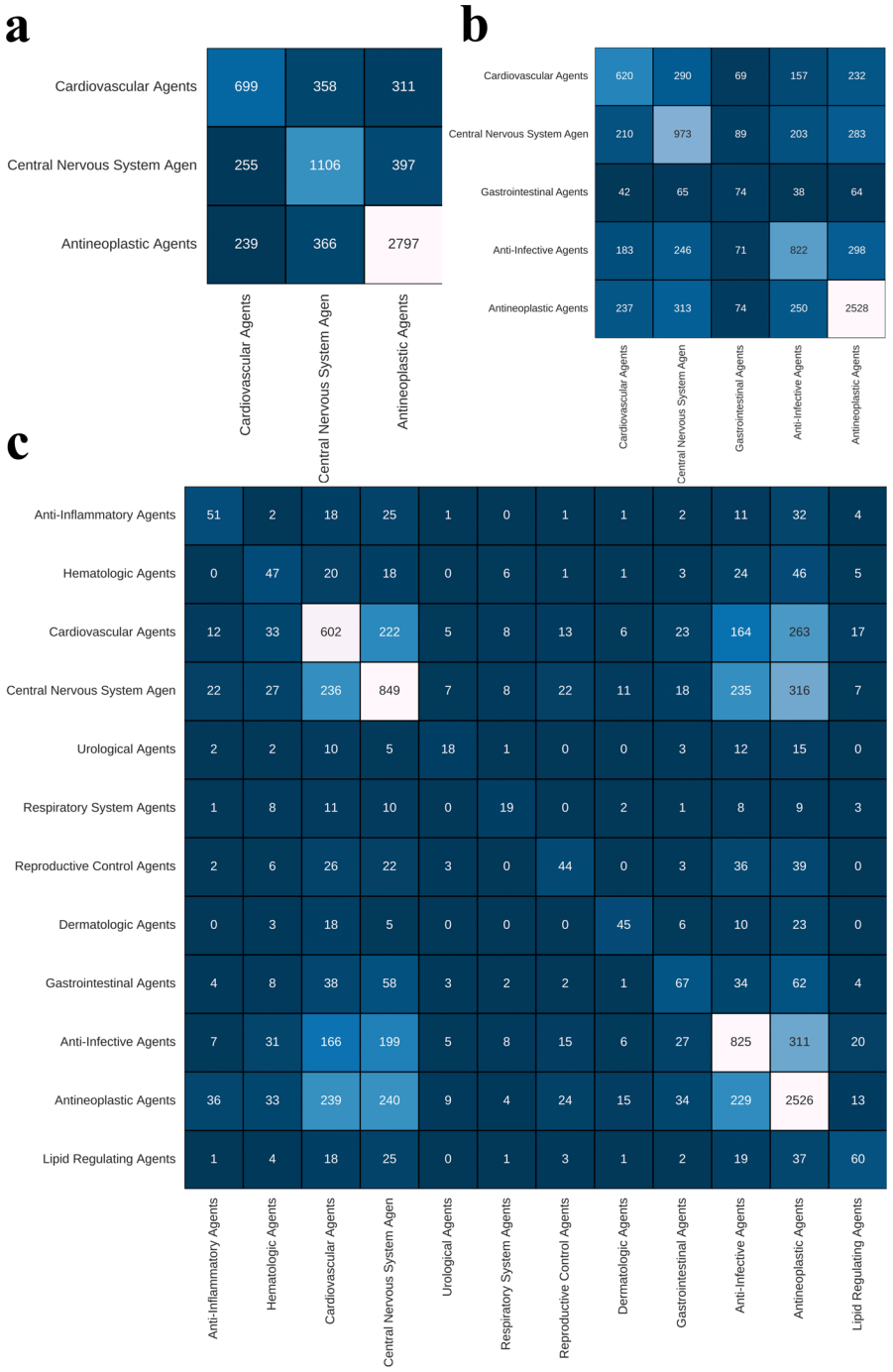


Figure 3. Validation confusion matrix representing deep neural network classification performance over a set of drugs profiled for A549, MCF7, and PC3 cell lines, belonging to 3 (a), 5 (b), and 12 (c) therapeutic classes. C(i,j) element is a sample count of how many times i was the truth and j was predicted.

pinacidil might influence some of them. Aforementioned cases hint to the fact that imperfect accuracy here might not be a bad thing and the DNN model could serve as powerful drug repositioning tool.

DISCUSSION

With increasing availability of big data and GPU computing, the entire field of deep learning is experiencing very rapid development, and the breadth of DNN applications goes far beyond text, voice, and image recognition problems. In this paper we explored the possibility of using DL to classify various drugs into therapeutic categories solely based on their transcriptomic data. To our knowledge, this is the first DL model to map transcriptomic data onto therapeutical category.

DNN trained on gene level data did not perform very well, achieving only 0.24 F1 score on 12 classes. Thus, as a way to reduce dimensionality and keep biological relevance, we decided to apply pathway activation scoring.[27] Translation of perturbation profiles onto the pathway level turned out to be very beneficial. Pathways served as excellent features, and we were able to exclude insignificantly perturbed samples and demonstrate the ability of deep neural network to recognize

DOI: 10.1021/acs.molpharmaceut.6b00248
Mol. Pharmaceutics 2016, 13, 2524−2530

sophisticated drug action mechanisms on the pathway level. The power of this approach is further highlighted by the fact that it performs with high accuracy on a pooled data set with samples from cells treated by different drug concentrations and perturbation times. Furthermore, the DNN achieves significant classification accuracy even across different cell lines. Its performance turned out to be better than SVM on every multiclass classification problem. When we used the same set of significantly perturbed samples that we selected with the pathway activation approach, and trained DNN on a data set with gene expression data for "landmark genes", the results turned out to be significantly worse. Hence, we can conclude that pathway level data is more complementary for DNN and more suitable for classifying drugs into therapeutic use categories. Proper comparison to reference group of samples plays an essential role, and merely normalized gene expression data from perturbed samples is not sufficient for complex classification tasks.

It is possible to interpret the classification results from different angles. For instance, as it was shown, in confusion matrices (Figure 3), that the "misclassified" samples for a certain drug might in fact be an indication of its potential for novel use, or repurposing, in these exact "incorrectly" assigned conditions. Misclassification, therefore, may lead to unexpected new discoveries. This approach opens a great avenue for application of DL in the drug repurposing field.

MATERIALS AND METHODS

Data Collection. In this study, we performed the analysis of gene expression data produced by the LINCS Project participants (http://www.lincsproject.org/). We utilized the level 3 (Q2NORM) gene expression data for three cell lines: MCF7, A549, and PC3. Q2NORM data contains the expression levels of both directly measured landmark transcripts plus inferred genes, which were normalized using invariant set scaling followed by quantile normalization. Mapping probes onto official HGNC[35] gene symbols resulted in a gene expression data set comprising 12,797 genes total.

Drug Selection. To link the drugs profiled by LINCS Project to medical conditions, we utilized MeSH (medical subject headings) classification (https://www.nlm.nih.gov/mesh/). We picked only those drugs that had a link to a disease in the "therapeutic use" section of the classification tree. Drugs belonging to several categories simultaneously were excluded from the analysis.

Gene Expression Analysis. All of the selected drugs' samples collected for three cell lines were grouped by the combination of the following parameters: cell line, drug, perturbation concentration, and perturbation time. This resulted in a total number of 26,420 samples. For gene level analysis we trained our models on both whole data set of 12,797 genes and a subset of 977 so-called "landmark genes" defined in the LINCS Project.

In pathway level analysis, for each given case sample group we generated a reference group consisting of samples perturbed with DMSO, that came from the same RNA plate as samples from the case group. After that, each case sample group was independently analyzed using an algorithm called OncoFinder.[27] Taking the preprocessed gene expression data as an input, it allows for cross-platform data set comparison with low error rate and has the ability to obtain functional features of intracellular regulation using mathematical estimations. For each investigated sample group it performs a case-reference comparison using Student's t test, generates the list of significantly differentially expressed genes, and calculates the pathway activation strength (PAS), a value which serves as a qualitative measure of pathway activation. Positive and negative PAS values indicate pathway up- and downregulation, respectively. In this study the genes with FDR-adjusted p-value <0.05 were considered significantly differentially expressed. Samples with zero pathway activation score for all of the pathways were considered as insignificantly perturbed and were excluded from further analysis. The filtered data set contained 308, 454, and 433 drugs for A549, MCF7, and PC3 cell lines, respectively, and comprised 9352 samples total (Supplementary Tables 1 and 2).

Classification Methods. Among the multitude of available classification methods we have employed two that are highly robust and widely successful in fields outside drug prediction: SVM[36,37] and deep neural network.[38] SVMs are a celebrated classification method for their flexibility and ease of use, while deep learning approaches are continuing to break records in many pattern recognition tasks.

Flexibility of SVM, as a maximum margin classifier, is in part reflected in provided ability to select a kernel that fits the data best and choose a soft-margin parameter that allows for best generalization. However, these parameters are not evident for any given data set. It has been previously shown that radial basis function (RBF) kernel SVMs perform well on largely different data. In order to be more objective in selection of kernel we have allowed our nested cross validation choice of three. We used grid search for hyperparameter optimization. We have trained the SVM via nested 3-fold cross validation for hyperparameters that include kernel type (linear, RBF, or polynomial) and soft-margin parameter C embedded in the outer 10-fold cross-validation loop. For each fold, the algorithm could have selected different kernels and different soft-margin parameters. Nevertheless, RBF was indeed the most preferred.

The deep learning method used in our work was the standard fully connected multilayer perceptron with 977 input nodes for gene expression level data and 271 for pathway activation scores. Similarly to SVM, we used grid search for hyperparameter optimization. We used cross entropy as cost function and AdaDelta[39] as cost function optimizer. We searched for the optimal number of layers, number of hidden units, and dropout rejection rate in a nested cross-validation framework. For the hyperparameter search, the number of layers varied from 3 to 8, where the number of hidden units per layer was reduced with the depth of the network to half the previous layer. We set the search for the starting hidden layer from 300 to 900 in steps of 150. Each layer was initialized with the Glorot uniform approach.[40] The experiment shows that the best combination of parameters was 3 hidden layers with 200 in each with rectified linear activation function. The dropout rejection ratio was tested for 20% and 80% at each layer. We chose an antisymmetric activation function, hyperbolic tangent, as the nonlinear function of hidden neurons because the data is normalized to zero mean and unit variance, thus we rely on deviations from the mean to train the network. Also, it has been reported that the hyperbolic tangent speeds up convergence compared to sigmoid functions.[39] The number of output nodes was equal to the number of classes to explore in each particular experiment with a softmax activation function. The code that implements the feed-forward neural network used in our experiments is publicly available at https://github.com/

DOI: 10.1021/acs.molpharmaceut.6b00248
Mol. Pharmaceutics 2016, 13, 2524−2530

alvarouc/mlp (commit: 8b07a1a18b17ca530fdcb482fce-c24e26e36b27a).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.molpharmaceut.6b00248.

MeSH category stratification binary matrix (XLSX)

Number of drugs selected for A549, MCF7, and PC3 cell lines (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: aliper@insilicomedicine.com.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We would like to thank Dr. Leslie C. Jellen of Insilico Medicine for editing this manuscript for language and style. Authors thank Mark Berger and NVIDIA Corporation for providing valuable advice and high performance GPU equipment for deep learning applications.

■ REFERENCES

(1) Loging, W.; Harland, L.; Williams-Jones, B. High-Throughput Electronic Biology: Mining Information for Drug Discovery. *Nat. Rev. Drug Discovery* **2007**, DOI: 10.1038/nrd2345.

(2) Kirchmair, J.; Göller, A. H.; Lang, D.; Kunze, J.; Testa, B.; Wilson, I. D.; Glen, R. C.; Schneider, G. Predicting Drug Metabolism: Experiment and/or Computation? *Nat. Rev. Drug Discovery* **2015**, *14*, 387.

(3) Schirle, M.; Jenkins, J. L. Identifying Compound Efficacy Targets in Phenotypic Drug Discovery. *Drug Discovery Today* **2016**, *21*, 82.

(4) LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521* (7553), 436−444.

(5) Baldi, P.; Sadowski, P.; Whiteson, D. Searching for Exotic Particles in High-Energy Physics with Deep Learning. *Nat. Commun.* **2014**, *5*, 4308.

(6) Schmidhuber, J. Deep Learning in Neural Networks: An Overview. *Neural Networks* **2015**, *61*, 85−117.

(7) Mamoshina, P.; Vieira, A.; Putin, E.; Zhavoronkov, A. Applications of Deep Learning in Biomedicine. *Mol. Pharmaceutics* **2016**, *13* (5), 1445−1454.

(8) Hira, Z. M.; Gillies, D. F. A Review of Feature Selection and Feature Extraction Methods Applied on Microarray Data. *Adv. Bioinf.* **2015**, *2015*, 198363.

(9) Ibrahim, R.; Rania, I.; Yousri, N. A.; Ismail, M. A.; El-Makky, N. M. Multi-Level gene/MiRNA Feature Selection Using Deep Belief Nets and Active Learning. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*; 2014.

(10) Fakoor, R.; Ladhak, F.; Nazi, A.; Huber, M. *Using Deep Learning to Enhance Cancer Diagnosis and Classification*. In *Proceedings of the International Conference on Machine Learning*; 2013.

(11) Leung, M. K. K.; Xiong, H. Y.; Lee, L. J.; Frey, B. J. Deep Learning of the Tissue-Regulated Splicing Code. *Bioinformatics* **2014**, *30* (12), i121−i129.

(12) Lyons, J.; Dehzangi, A.; Heffernan, R.; Sharma, A.; Paliwal, K.; Sattar, A.; Zhou, Y.; Yang, Y. Predicting Backbone Cα Angles and Dihedrals from Protein Sequences by Stacked Sparse Auto-Encoder Deep Neural Network. *J. Comput. Chem.* **2014**, *35* (28), 2040−2046.

(13) Wang, S.; Weng, S.; Ma, J.; Tang, Q. DeepCNF-D: Predicting Protein Order/Disorder Regions by Weighted Deep Convolutional Neural Fields. *Int. J. Mol. Sci.* **2015**, *16* (8), 17315−17330.

(14) Lusci, A.; Pollastri, G.; Baldi, P. Deep Architectures and Deep Learning in Chemoinformatics: The Prediction of Aqueous Solubility for Drug-like Molecules. *J. Chem. Inf. Model.* **2013**, *53* (7), 1563−1575.

(15) Wang, C.; Caihua, W.; Juan, L.; Fei, L.; Yafang, T.; Zixin, D.; Qian-Nan, H. *Pairwise Input Neural Network for Target-Ligand Interaction Prediction*. In *2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*; 2014.

(16) Hughes, T. B.; Miller, G. P.; Swamidass, S. J. Modeling Epoxidation of Drug-like Molecules with a Deep Machine Learning Network. *ACS Cent. Sci.* **2015**, *1* (4), 168−180.

(17) Xu, Y.; Dai, Z.; Chen, F.; Gao, S.; Pei, J.; Lai, L. Deep Learning for Drug-Induced Liver Injury. *J. Chem. Inf. Model.* **2015**, *55* (10), 2085−2093.

(18) Solovyeva, K. P.; Karandashev, I. M.; Zhavoronkov, A.; Dunin-Barkowski, W. L. Models of Innate Neural Attractors and Their Applications for Neural Information Processing. *Front. Syst. Neurosci.* **2015**, DOI: 10.3389/fnsys.2015.00178.

(19) Newby, D.; Freitas, A. A.; Ghafourian, T. Comparing Multilabel Classification Methods for Provisional Biopharmaceutics Class Prediction. *Mol. Pharmaceutics* **2015**, *12* (1), 87−102.

(20) Wenlock, M. C.; Barton, P. In Silico Physicochemical Parameter Predictions. *Mol. Pharmaceutics* **2013**, *10* (4), 1224−1235.

(21) Broccatelli, F.; Cruciani, G.; Benet, L. Z.; Oprea, T. I. BDDCS Class Prediction for New Molecular Entities. *Mol. Pharmaceutics* **2012**, *9* (3), 570−580.

(22) Herrera-Ruiz, D.; Faria, T. N.; Bhardwaj, R. K.; Timoszyk, J.; Gudmundsson, O. S.; Moench, P.; Wall, D. A.; Smith, R. L.; Knipp, G. T. A Novel hPepT1 Stably Transfected Cell Line: Establishing a Correlation between Expression and Function. *Mol. Pharmaceutics* **2004**, *1* (2), 136−144.

(23) Iskar, M.; Zeller, G.; Blattmann, P.; Campillos, M.; Kuhn, M.; Kaminska, K. H.; Runz, H.; Gavin, A.-C.; Pepperkok, R.; van Noort, V.; Bork, P. Characterization of Drug-Induced Transcriptional Modules: Towards Drug Repositioning and Functional Understanding. *Mol. Syst. Biol.* **2013**, *9*, 662.

(24) Kutalik, Z.; Beckmann, J. S.; Bergmann, S. A Modular Approach for Integrative Analysis of Large-Scale Gene-Expression and Drug-Response Data. *Nat. Biotechnol.* **2008**, *26* (5), 531−539.

(25) Spirin, P. V.; Lebedev, T. D.; Orlova, N. N.; Gornostaeva, A. S.; Prokofjeva, M. M.; Nikitenko, N. A.; Dmitriev, S. E.; Buzdin, A. A.; Borisov, N. M.; Aliper, A. M.; Garazha, A. V.; Rubtsov, P. M.; Stocking, C.; Prassolov, V. S. Silencing AML1-ETO Gene Expression Leads to Simultaneous Activation of Both pro-Apoptotic and Proliferation Signaling. *Leukemia* **2014**, *28* (11), 2222−2228.

(26) Zhu, Q.; Izumchenko, E.; Aliper, A. M.; Makarev, E.; Paz, K.; Buzdin, A. A.; Zhavoronkov, A. A.; Sidransky, D. Pathway Activation Strength Is a Novel Independent Prognostic Biomarker for Cetuximab Sensitivity in Colorectal Cancer Patients. *Hum. Genome Var.* **2015**, *2*, 15009.

(27) Buzdin, A. A.; Zhavoronkov, A. A.; Korzinkin, M. B.; Venkova, L. S.; Zenin, A. A.; Smirnov, P. Y.; Borisov, N. M. Oncofinder, a new method for the analysis of intracellular signaling pathway activation using transcriptomic data. *Front. Genet.* **2014**, *5*, 55.

(28) Artemov, A.; Aliper, A.; Korzinkin, M.; Lezhnina, K.; Jellen, L.; Zhukov, N.; Roumiantsev, S.; Gaifullin, N.; Zhavoronkov, A.; Borisov, N.; Buzdin, A. A Method for Predicting Target Drug Efficiency in Cancer Based on the Analysis of Signaling Pathway Activation. *Oncotarget* **2015**, *6* (30), 29347−29356.

(29) Venkova, L.; Aliper, A.; Suntsova, M.; Kholodenko, R.; Shepelin, D.; Borisov, N.; Malakhova, G.; Vasilov, R.; Roumiantsev, S.; Zhavoronkov, A.; Buzdin, A. Combinatorial High-Throughput Experimental and Bioinformatic Approach Identifies Molecular Pathways Linked with the Sensitivity to Anticancer Target Drugs. *Oncotarget* **2015**, *6* (29), 27227−27238.

(30) Makarev, E.; Cantor, C.; Zhavoronkov, A.; Buzdin, A.; Aliper, A.; Csoka, A. B. Pathway Activation Profiling Reveals New Insights

into Age-Related Macular Degeneration and Provides Avenues for Therapeutic Interventions. *Aging* **2014**, *6* (12), 1064−1075.
(31) Combrisson, E.; Jerbi, K. Exceeding Chance Level by Chance: The Caveat of Theoretical Chance Levels in Brain Signal Classification and Statistical Assessment of Decoding Accuracy. *J. Neurosci. Methods* **2015**, *250*, 126−136.
(32) Langmead, C. J.; Watson, J.; Reavill, C. Muscarinic Acetylcholine Receptors as CNS Drug Targets. *Pharmacol. Ther.* **2008**, *117* (2), 232−243.
(33) Volpicelli, L. A.; Levey, A. I. Muscarinic Acetylcholine Receptor Subtypes in Cerebral Cortex and Hippocampus. *Prog. Brain Res.* **2004**, *145*, 59−66.
(34) Trimmer, J. S. Subcellular Localization of K+ Channels in Mammalian Brain Neurons: Remarkable Precision in the Midst of Extraordinary Complexity. *Neuron* **2015**, *85* (2), 238−256.
(35) Gray, K. A.; Yates, B.; Seal, R. L.; Wright, M. W.; Bruford, E. A. Genenames.org: The HGNC Resources in 2015. *Nucleic Acids Res.* **2015**, *43* (D1), D1079−D1085.
(36) Boser, B. E.; Guyon, I. M.; Vapnik, V. N. A Training Algorithm for Optimal Margin Classifiers. *Proc. 5th Annu. Workshop Comput. Learn. Theory - COLT '92* **1992**, 144.
(37) Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20* (3), 273−297.
(38) Hinton, G. E.; Salakhutdinov, R. R. Reducing the Dimensionality of Data with Neural Networks. *Science* **2006**, *313* (5786), 504−507.
(39) Zeiler, M. D. ADADELTA: An Adaptive Learning Rate Method *arXiv* **2012**, *6*.
(40) Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. *Int. Conf. Artif. Intell. Stat.* **2010**, 249−256.

DOI: 10.1021/acs.molpharmaceut.6b00248
Mol. Pharmaceutics 2016, 13, 2524−2530