

CTRP & CCLE Data set

The second data set used in this study [1] was obtained from two sources i.e., Cancer Cell line Encyclopedia CCLE [2] and Cancer Therapeutic Response Portal (CTRP) version release v1 2013 [3] by BROAD Institute. The pathways (and gene sets) information was extracted from Molecular Signature database MSigDB [4].

NB: If you use this data set in your study, remember to cite the original sources of the data [2, 3, 4] in addition to this study [1]. You find the citation information below.

Data Description

From the study [1], the data set is provided in two formats

- R data format in RData folder
- Matlab format in matData folder

where in each folder, the following data objects can be found

1. DataViews is a list and contains two objects
 - (a) DataViews : the input data matrices (or kernels). In paper these are denoted by $\{\mathbf{K}_{x,m}\}$
 - (b) DataViewsNames : the names of the views (i.e., pathways)
2. DrugResponse is also a list and contains three objects
 - (a) DrugResponse: the output matrix , in paper denoted by \mathbf{Y} .
 - (b) CelllineNames: names of the cell lines (denoting to the rows of \mathbf{Y})
 - (c) DrugNames: names of the drugs (denoting to the columns of \mathbf{Y})
3. DrugTargets: names of the primary targets.
4. 5Folds_CVIndexes_* : Cross-Validation(CV) indexes for 5 folds. There are 10 sets of these CV indexes

CITATION information

```
@article{barretina2012cancer,  
title={The Cancer Cell Line Encyclopedia enables predictive modelling  
of anticancer drug sensitivity},  
author={Barretina, Jordi and Caponigro, Giordano and Stransky, Nicolas  
and Venkatesan, Kavitha and Margolin, Adam A and Kim, Sungjoon and  
Wilson, Christopher J and Lehman, Joseph and Kryukov, Gregory V  
and Sonkin, Dmitriy and others},
```

```

journal={Nature},
volume={483},
number={7391},
pages={603--607},
year={2012},
publisher={Nature Publishing Group}
}
@article{basu2013interactive,
title={An interactive resource to identify cancer genetic and lineage
dependencies targeted by small molecules},
author={Basu, Amrita and Bodycombe, Nicole E and Cheah, Jaime H and
Price, Edmund V and Liu, Ke and Schaefer, Giannina I and Ebright, Richard
Y and Stewart, Michelle L and Ito, Daisuke and Wang, Stephanie and
others},
journal={Cell},
volume={154},
number={5},
pages={1151--1161},
year={2013},
publisher={Elsevier}
}
@article{liberzon2011molecular,
title={Molecular signatures database {MSigDB} 3.0},
author={Liberzon, Arthur and Subramanian, Aravind and Pinchback, Reid
and Thorvaldsdottir, Helga and Tamayo, Pablo and Mesirov, Jill
P},
journal={Bioinformatics},
volume={27},
number={12},
pages={1739--1740},
year={2011},
publisher={Oxford Univ Press}
}
@article{Ammad-ud-din01092016,
author = {Ammad-ud-din, Muhammad and Khan, Suleiman A. and Malani,
Disha and Murumgi, Astrid and Kallioniemi, Olli and Aittokallio, Tero
and Kaski, Samuel},
title = {Drug response prediction by inferring pathway-response associations
with kernelized Bayesian matrix factorization},
volume = {32},
number = {17},
pages = {i455-i463},
year = {2016},
doi = {10.1093/bioinformatics/btw433},
URL = {http://bioinformatics.oxfordjournals.org/content/32/17/i455.abstract},
eprint = {http://bioinformatics.oxfordjournals.org/content/32/17/i455.full.pdf+html},

```

```
journal = {Bioinformatics}  
}
```

References

- [1] M. Ammad-ud din, S. A. Khan, D. Malani, A. Murumgi, O. Kallioniemi, T. Aittokallio, and S. Kaski, “Drug response prediction by inferring pathway-response associations with kernelized bayesian matrix factorization,” *Bioinformatics*, vol. 32, no. 17, pp. i455–i463, 2016.
- [2] J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan, A. A. Margolin, S. Kim, C. J. Wilson, J. Lehár, G. V. Kryukov, D. Sonkin, *et al.*, “The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity,” *Nature*, vol. 483, no. 7391, pp. 603–607, 2012.
- [3] A. Basu, N. E. Bodycombe, J. H. Cheah, E. V. Price, K. Liu, G. I. Schaefer, R. Y. Ebright, M. L. Stewart, D. Ito, S. Wang, *et al.*, “An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules,” *Cell*, vol. 154, no. 5, pp. 1151–1161, 2013.
- [4] A. Liberzon, A. Subramanian, R. Pinchback, H. Thorvaldsdóttir, P. Tamayo, and J. P. Mesirov, “Molecular signatures database MSigDB 3.0,” *Bioinformatics*, vol. 27, no. 12, pp. 1739–1740, 2011.