

## GDSC Data set

The first data set used in this study [1] was obtained from Genomics of Drug Sensitivity in Cancer (GDSC) project initiated by Wellcome Trust Sanger Institute version release, June 2014 [2, 3]. 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{Garnett2012,  
title={Systematic identification of genomic markers of drug sensitivity  
in cancer cells},  
author={Garnett, Mathew J and Edelman, Elena J and Heidorn, Sonja J  
and Greenman, Chris D and Dastur, Anahita and Lau, King Wai and Greninger,  
Patricia and Thompson, I Richard and Luo, Xi and Soares, Jorge and  
others},  
journal={Nature},
```

```

volume={483},
number={7391},
pages={570--575},
year={2012},
publisher={Nature Publishing Group}
}
@article{yang2013genomics,
title={Genomics of Drug Sensitivity in Cancer ({GDSC}): a resource
for therapeutic biomarker discovery in cancer cells},
author={Yang, Wanjuan and Soares, Jorge and Greninger, Patricia and
Edelman, Elena J and Lightfoot, Howard and Forbes, Simon and Bindal,
Nidhi and Beare, Dave and Smith, James A and Thompson, I Richard and
others},
journal={Nucleic Acids Res.},
volume={41},
number={D1},
pages={D955--D961},
year={2013},
publisher={Oxford Univ Press}
}
@article{liberzon2011molecular,
title={Molecular signatures database {MSigDB} 3.0},
author={Liberzon, Arthur and Subramanian, Aravind and Pinchback, Reid
and Thorvaldsd{\o}ttir, 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] W. Yang, J. Soares, P. Greninger, E. J. Edelman, H. Lightfoot, S. Forbes, N. Bindal, D. Beare, J. A. Smith, I. R. Thompson, *et al.*, “Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells,” *Nucleic Acids Res.*, vol. 41, no. D1, pp. D955–D961, 2013.
- [3] M. J. Garnett, E. J. Edelman, S. J. Heidorn, C. D. Greenman, A. Dastur, K. W. Lau, P. Greninger, I. R. Thompson, X. Luo, J. Soares, *et al.*, “Systematic identification of genomic markers of drug sensitivity in cancer cells,” *Nature*, vol. 483, no. 7391, pp. 570–575, 2012.
- [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.