Package 'ailm'

February 22, 2025

Type Package

Title Foundations of linear modeling at the age of AI

Version 0.0.1

Author Xu Liu [aut,cre]

Maintainer Xu Liu <liu.xu@sufe.edu.cn>

Description Data sets from the book "Foundations of linear modeling at the age of AI".

License GPL (>=2)

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

Depends R (>= 3.2.0), Matrix, glmnet

NeedsCompilation yes

Repository github

URL https://github.com/xliusufe/ailm

Contents

---

breastcancer *Breast cancer data*

Description

The dataset contains gene expression and gene copy number information from 89 subjects.

Usage

```
data(breastcancer)
```

Arguments

dna	Copy number variation (CNV) data representing genomic DNA amplification or deletion events in tumor samples.
rna	Gene expression profiles measured via RNA transcript levels (e.g., microarray or RNA-seq data).
chrom	Chromosome numbers (1-22, X, Y) corresponding to the genomic location of the measured genes.
nuc	Nucleotide positions (start/end coordinates) of genes or probes on the chromosome (e.g., hg18/hg19 reference).
gene	Unique gene identifiers (e.g., Entrez Gene IDs or probe IDs) linked to genomic features.
genenames	Official gene symbols or names (e.g., BRCA1, ERBB2) standardized by HUGO Gene Nomenclature Committee (HGNC).
genechr	Chromosomal mapping information for each gene (e.g., "chr17" for TP53).
genedesc	Brief functional descriptions of genes (e.g., "tumor protein p53" or "estrogen receptor 1").
genepos	Genomic coordinates of genes (e.g., cytoband or base-pair positions like "17q21.31").

Details

The dataset is derived from molecular bioinformatics data obtained from breast cancer tissue samples treated according to the standard of care between 1989 and 1997. It primarily consists of gene expression profiles and copy number variation data across 22 chromosomal pairs in tumor tissue samples from 89 breast cancer patients. For a detailed explanation of this dataset, please refer to Chin et al. (2006).

Source

http://icbp.lbl.gov/breastcancer/ or R package PMA

References

Chin, K., DeVries, S., et al. (2006). Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer cell*, **10**(6), 529–541. doi:10.1016/j.ccr.2006.10.009

Examples

```
library(glmnet)
library(ailm)
data(breastcancer)
dna = breastcancer$dna[breastcancer$chrom==21,]
rna = breastcancer$rna[which(breastcancer$genechr==21),]
y = dna[1,]
x = t(rna)
set.seed(100)
fit_ridge = cv.glmnet(x,y,alpha = 0)
coef(fit_ridge, s = "lambda.min")
fit_lasso = cv.glmnet(x,y,alpha = 1)
coef(fit_lasso, s = "lambda.min")
```

Index