

Prioritisation of disease-associated genes with cNMTF

Centre for Integrative Systems Biology and Bioinformatics - Imperial College London

cNMTF is a data fusion framework for prioritising reliable associations between single nucleotide variants (SNVs), genes and diseases. **In this webpage you will find the step-by-step process to run the algorithm on your own data using R scripts.**

Installation

Install the ***cnmtf*** package directly from the github repository. This process requires the **devtools** package :

```
install.packages("devtools")  
  
devtools::install_github("lg115/cnmtf")
```

You can check the development version of this R package on Github.

What is inside the package?

The ***cnmtf*** package provides four categories of functions for preprocessing data, clustering, scoring SNVs and comparing results.

- **Preprocessing functions:**
These functions will help you to create the inputs for the algorithm.
 - **Factorisation functions:**
Main functions to integrate the input data, generate the low-dimensional matrices and find consensus clusters.
 - **Scoring functions:**
A set of functions to score SNVs and prioritise significant SNV-trait associations from the low-dimensional matrices.
 - **Comparing functions:**
Auxiliary functions to compare your results across different settings of the algorithm.
-

Citation

Leal, L.G., David, A., Jarvelin, M.-R., Sebert, S., Männikkö, M., Karhunen, V., Seaby, E., Hoggart, C., and Sternberg, M.J.E. (2019). **Identification of disease-associated loci using machine learning for genotype and network data integration.** *Bioinformatics*, btz310 (<https://doi.org/10.1093/bioinformatics/btz310>)

Corresponding author:

Luis G. Leal, lg115@imperial.ac.uk

Centre for Integrative Systems Biology and Bioinformatics (CISBIO)

Department of Life Sciences

Imperial College London, London, UK, SW7 2AZ

Hands on

You must start by preparing Your input data.