GMUM.R Package Demo

Igor Sieradzki

tfml 2015

theoretical foundations
of machine learning, Będlewo

Presentation plan

- 1. What, why, who?
- 2. Current package modules
  - 2.1. CEC module
  - 2.2. SVM module
  - 2.3. GNG module

What is gmum.R

- R language package containing various machine learning models

What is gmum.R

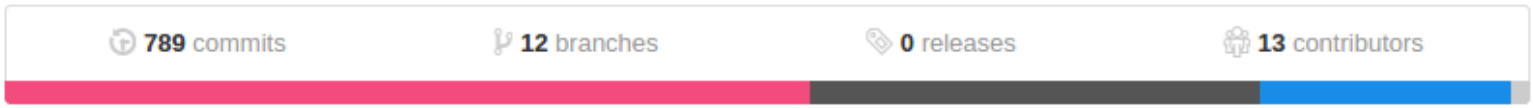
- R language package containing various machine learning models
- Started as a cooperation project between GMUM members and students

What is gmum.R

- R language package containing various machine learning models
- Started as a cooperation project between GMUM members and students
- This is still a **development** version of the package

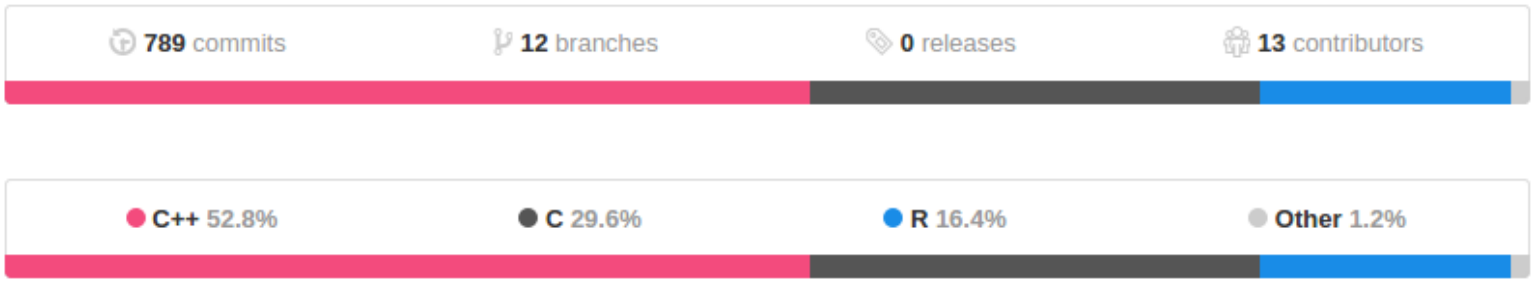
What is gmum.R

- R language package containing various machine learning models
- Started as a cooperation project between GMUM members and students
- This is still a **development** version of the package



What is gmum.R

- R language package containing various machine learning models
- Started as a cooperation project between GMUM members and students
- This is still a **development** version of the package



Why we created gmum.R

- Package for different machine learning tasks

Why we created gmum.R

- Package for different machine learning tasks
- Fast and efficient C++ based engines

Why we created gmum.R

- Package for different machine learning tasks
- Fast and efficient C++ based engines
- Implementing GMUM group algorithms

Why we created gmum.R

- Package for different machine learning tasks
- Fast and efficient C++ based engines
- Implementing GMUM group algorithms
- User friendly package not only for computer scientists

GMUM algorithms

- Tabor, J., & Spurek, P. (2014). Cross-entropy clustering. Pattern Recognition, 47(9), 3046-3059.
- Czarnecki, W. M., & Tabor, J. (2014). Two ellipsoid Support Vector Machines. Expert Systems with Applications, 41(18), 8211-8224.
- Podolak, I. T., & Jastrzębski, S. K. (2013, January). Density Invariant Detection of Osteoporosis Using Growing Neural Gas. In Proceedings of the 8th International Conference on Computer Recognition Systems CORES 2013 (pp. 629-638). Springer International Publishing.

Who is gmum.R

- Currently the main team consists of 9 master/bachelor students, 1 phd student and 3 GMUM members

Who is gmum.R

- Currently the main team consists of 9 master/bachelor students, 1 phd student and 3 GMUM members
- Team:
  - Project leader - Wojciech Czarnecki
  - Maintainer/GNG developer - Stanisław Jastrzębski
  - SVM team - Igor Sieradzki, Piotr Kowenzowski, Konrad Talik
  - CEC team - Marcin Data, Karol Jurek, Michał Pletty
  - Developers - Szymon Nakoneczny, Marcin Hatalski
  - Website and additional code - Matuesz Bruno-Kamiński
  - Former member - Maciej Zgliczyński

CEC

- CEC aims to efficiently implement Cross Entropy Clustering Algorithm as R extension.
- Cross-entropy clustering (shortly CEC) joins advantages of classical k-means with those of EM.

```
CEC(k=3,
    x=dataset,
    control.nstart=10,
    method.type='sphere',
    method.init='random')
```

SVM

- SVM Wrapper is a part of the gmum.R project which provides a popular Support Vector Machine implementations wrapped in the R package.
- Currently we support 2 SVM engines: libSVM and SVMLight

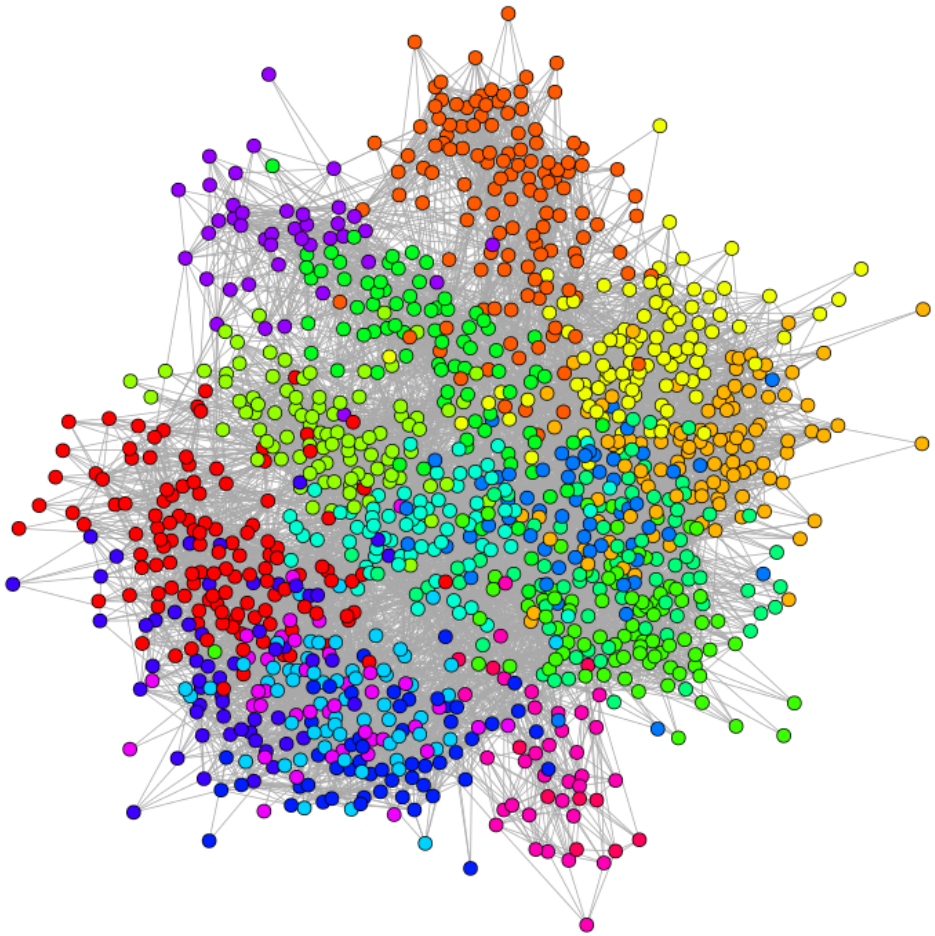
```
SVM(formula = Y~.  ,
    data = ds,
    lib = "libsvm",
    kernel = "linear",
    C = 1)
```

GNG

- Subpackage containing efficient, online GNG algorithm
- Produces graph, that you can easily convert to igraph and save

```
GNG(wine,
    labels=wine$Type,
    max_nodes=20)
```

MNIST with GNG



MNIST with GNG