

---

# **dislib Documentation**

***Release 0.3.0***

## **Workflows and Distributed Computing**

**Nov 04, 2019**



---

## Contents:

---

<b>1</b>	<b>Performance</b>	<b>3</b>
<b>2</b>	<b>Source code</b>	<b>5</b>
<b>3</b>	<b>Support</b>	<b>7</b>
3.1	Indices and tables . . . . .	7





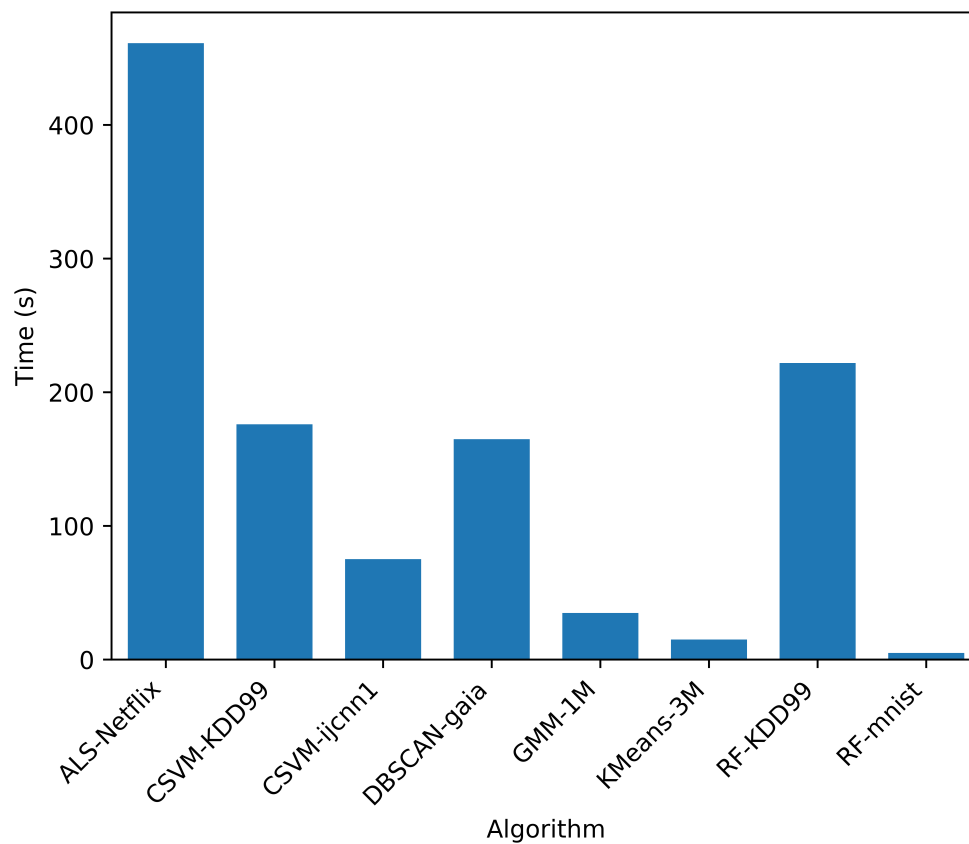
The Distributed Computing Library (dislib) provides distributed algorithms ready to use as a library. So far, dislib is highly focused on machine learning algorithms, and is greatly inspired by [scikit-learn](#). However, other types of numerical algorithms might be added in the future. The main objective of dislib is to facilitate the execution of big data analytics algorithms in distributed platforms, such as clusters, clouds, and supercomputers.

Dislib has been implemented on top of [PyCOMPSs](#) programming model, and it is being developed by the [Workflows and Distributed Computing](#) group of the [Barcelona Supercomputing Center](#).

- [Quickstart](#)
- [API Reference](#)
- [Development](#)
- [FAQ](#)



The following plot shows fit time of some dislib models on the [MareNostrum 4](#) supercomputer (using 8 worker nodes):



Labels on the horizontal axis represent algorithm-dataset, where:

- ALS = AlternatingLeastSquares

- CSVM = CascadeSVM
- GMM = GaussianMixture
- RF = RandomForestClassifier

and:

- Netflix = The Netflix Prize [dataset](#).
- ijcnn1 = The [ijcnn1](#) dataset.
- KDD99 = The [KDDCUP 1999](#) dataset.
- gaia = The Tycho-Gaia Astrometric Solution dataset<sup>1</sup>.
- 1M and 3M = 1 and 3 million random samples.
- mnist = The [mnist](#) dataset.

---

<sup>1</sup> Michalik, Daniel, Lindegren, Lennart, and Hobbs, David, “The Tycho-Gaia astrometric solution - How to get 2.5 million parallaxes with less than one year of Gaia data,” A&A, vol. 574, p. A115, 2015.



## CHAPTER 2

---

### Source code

---

The source code of dislib is available online at [Github](#).



If you have questions or issues about the dislib you can join us in [Slack](#).

Alternatively, you can send us an e-mail to [support-compss@bsc.es](mailto:support-compss@bsc.es).

### 3.1 Indices and tables

- [genindex](#)
- [modindex](#)
- [search](#)