

# Scikit from the article

David Arroyo Menéndez

April 26, 2019

## Scikit-learn: Machine Learning in Python

- Fabian Pedregosa
- Gaël Varoquaux
- Alexandre Gramfort
- Vincent Michel
- Bertrand Thirion

URL: <http://www.jmlr.org/papers/v12/pedregosa11a.html>

Citations: 15926

Scikit-learn is a Python module integrating a wide range of state-of-the-art machine learning algorithms for medium-scale supervised and unsupervised problems. This package focuses on bringing machine learning to non-specialists using a general-purpose high-level language. Emphasis is put on ease of use, performance, documentation, and API consistency. It has minimal dependencies and is distributed under the simplified BSD license, encouraging its use in both academic and commercial settings. Source code, binaries, and documentation can be downloaded from <http://scikit-learn.sourceforge.net>.

- Lenguaje de programación Python
- Qué hace. Necesidad de hacerlo

# Project Vision

- Code Quality
- BSD Licensing
- Bare-bone design and API
- Community-driven development
- Documentation

# Underlying Technologies

- Numpy
- Scipy
- Cython

- Objects specified by interface, not by inheritance
- Benchmarking between another machine learning tools
- Model Selection: cross-validation (GridSearch and Lasso).

Pipeline can combine several transformers

# High-level yet Efficient: Some Trade Offs

- SVM
- LARS
- Elastic Net
- kNN
- PCA



# Conclusion

- Resumen de secciones
- Qué hace el software
- Future work