

Scikit-Learn



Cole Smith, Yuvraj Chugh, Grace Yang, Won
Jong Yang

Project Introduction

- A Python module for machine learning and data science
- Provides tools for data mining and analysis
- Built on top of NumPy, SciPy, and Matplotlib

Motivation

- Interest in the Python community and the wish to improve it, give back
- Project has “Good First Issue” and “Easy” Issues – Easy to get started
- Community is responsive and open to suggestion

Health Assessment

- **Difficulty:** Wide range from easy to complex bugs
- **Responsiveness:** ~ 2-7 comments per issue
- **PR Turnaround:** Easy issues: <1 week, harder issues: 1 week +
- **Total Open Issues:** 1,222 (as of Mar 8th)

Getting Started

- Identify and leverage the best means of communication suitable for all team members
- Explore issues and identify those for beginners/newcomers to start contributing
- Divide issues not only by strengths, but by interests and time to ensure fairness and to maintain a level playing field
- Move on to harder issues once all team members get used to contributing