

GitHub を用いた開発フローの判別分析

プロジェクトマネジメントコース・ソフトウェア開発管理グループ 矢吹研究室 1242132 若月 純

1. 研究の背景

ソフトウェア開発では、GitHub を用いることが多い。GitHub は、バージョン管理システムに加え、branch、PullRequest、Issues といった開発を補助する機能を提供するサービスである。

GitHub を使用する手順を開発フローと呼ぶ。現在わかっている開発フローの数は 13 個ある [1]。開発フローの例を 1 つあげる。GitHub Flow は、作業をする branch を作成し、完成したら統合する。というような開発フローである。この開発フローはとてもシンプルなため、開発フローを実施するまでの学習コストは、抑えられるが、開発規模が大きい場合、PullRequest がたまりやすく、コードレビューに時間がかかってしまうことがある [2]。

このように開発フローは、メリット・デメリットがある。そのため、開発フローをプロジェクトの性質から選択する基準が必要である。

2. 目的

GitHub を用いたソフトウェア開発プロジェクトの性質において、適切な開発フローを選択できるようにするための基準を提供する。

3. 研究方法

初めに、GitHub 上のプロジェクトから、プロジェクトの性質と開発フローを調査する。開発フローは、GitHub 上のプロジェクトと、開発フローの特徴を照らしあわせることで求められる。

次に、調査したデータを分析する。分析は、調査したデータをランダムに 2 種類のデータに分け、決定木分析を行う。決定木分析により、プロジェクトがどのような性質を持つときに、どの開発フローが使われているかを明らかにする。

4. 成果物のイメージ

32 件の調査データを 2 種類にわけ、決定木分析を行い、開発フローを選択する基準を求める。

5. 進捗状況

GitHub 上の 32 件のプロジェクトから、プロジェクトの性質と開発フローを調査し、分析を行った。

プロジェクトの性質は、行数、ファイル数、バイト数、Watch 数、Star 数、Fork 数、Commit 数、branch 数、Release 数、人数、Issues 数、PullRequest 数、Label 数、Milestone 数、Wiki 数、言語を調査した。

開発フローは、Git Flow、GitHub Flow、LINE Flow、Stable Flow、WIP Flow の 5 種類だった。

決定木分析結果は、図 1 と図 2 である。図 1 は、Issues 数とファイル数、図 2 は、言語、行数、バイト数により分けられた。2 種類のデータに分け、分析を行うことで、異なった開発フローを選択する基準を求めることが出来た。

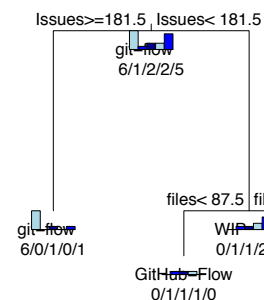


図 1 決定木 1

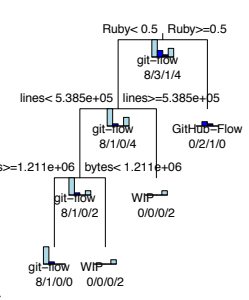


図 2 決定木 2

6. 今後の計画

時系列データを分析に含めた場合、基準にどのような変化があるか調査する。そうすることで、開発スピードを含めた開発フローを選択する基準が求められる。

参考文献

- [1] 小野寺航己. バージョン管理システムを活用するソフトウェア開発の開発フロー. 卒業論文, 千葉工業大学, 2015.
- [2] 大塚弘記. GitHub 実践入門 Pull Request による開発の変革. 技術評論社, 2014.