

ユーザ特性の分析 およびBotの作成

作島大智

背景

- ✦ OSS開発においてボランティアである開発者の離脱がプロジェクトの存続に大きく影響を与える
- ✦ バグ修正パッチの検証を行うコミッターの作業量が非常に多い
- ✦ プロジェクトに対する貢献を維持するためには開発者のモチベーションを維持することが重要

問題点

- ✦ 開発者がプロジェクトから離脱する
→ 開発者間における対応の差に失望

目的

- ✦ 開発者間のコミュニケーションにおいて
開発者のモチベーションを維持できるように支援を行う

アプローチ

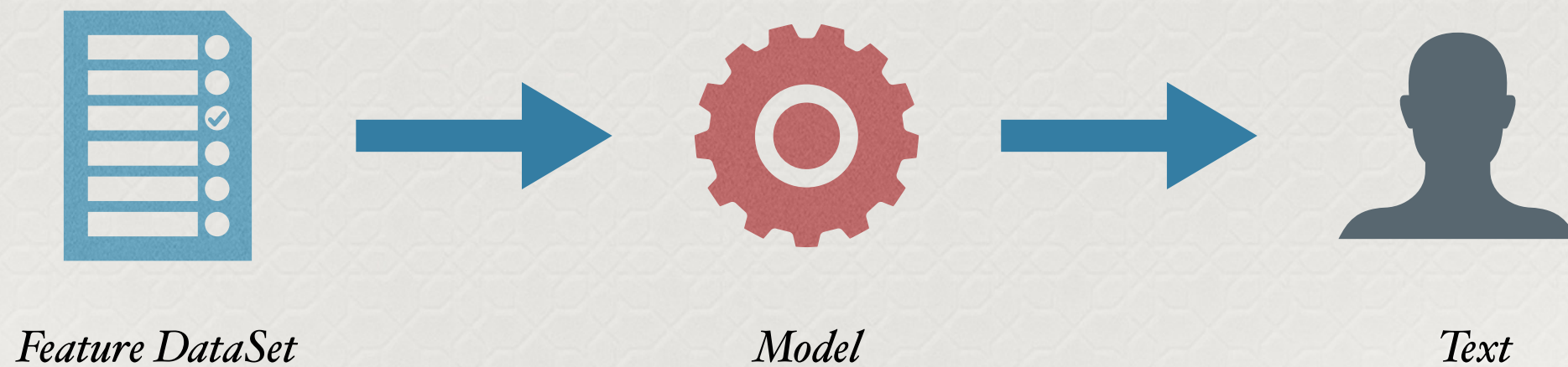
- ✿ *Pull Request*が承認された開発者, 承認されない開発者とコミッターの会話テキストデータを対象
- ✿ *LIWC*を用いて性格データセットの分析と検証
- ✿ 性格データセットを使って*Bot*を作成

大まかな流れ

分析



Bot作成



分析&検証

分析



- ✦ ユーザの発言テキストを収集しLIWCに投げる
- ✦ ユーザの性格特性(*Big5*)を獲得
- ✦ 性格分析のデータを検証

LIWC

- ✿ *Linguistic Inquiry and Words Count*の略称
- ✿ テキスト分析アプリケーション
 - 様々な感情的、認知的および構造的な要素の研究のために効率的で効果的な方法を提供するために開発

Big 5

- ✦ 外向性-内向性(*Extraversion*)
- ✦ 情緒安定性(*Neuroticism*)
- ✦ 開放性(*Openness to experience*)
- ✦ 誠實性(*Conscientiousness*)
- ✦ 協調性(*Agreeableness*)

Big 5

	High	Low
外向性 <i>Extraversion</i>	社交的・エネルギッシュ	孤独志向・控えめ
情緒安定性 <i>Neuroticism</i>	繊細・神経質	情緒安定・自信家
開放性 <i>Openness</i>	独創的・好奇心が強い	着実・警戒心が強い
誠実性 <i>Conscientiousness</i>	手際が良い・マメな人	楽天的・不注意
協調性 <i>Agreeableness</i>	人当たりの良い・温情あり	冷たい・不親切

良識性 (conscientiousness)

- ✦ 短期的な衝動や欲よりも長期的・合理的な利益を優先する節度の度合いを測る性格軸

情緒安定性 (neuroticism)

- ✦ 神経症傾向とも
- ✦ 不安を引きずる度合いを示す

知的好奇心（openness to experience）

- ✿ 「経験への開放性」とも
- ✿ 高い人は、新しい経験に対してオープンであり、知的で、連想が広がりやすい傾向
- ✿ 低い人は、変化を好まず、感覚的であり、大雑把でときに頑固で、権威に従順な面がある
- ✿ オープンであるということは、自分と環境の境界意識が薄い

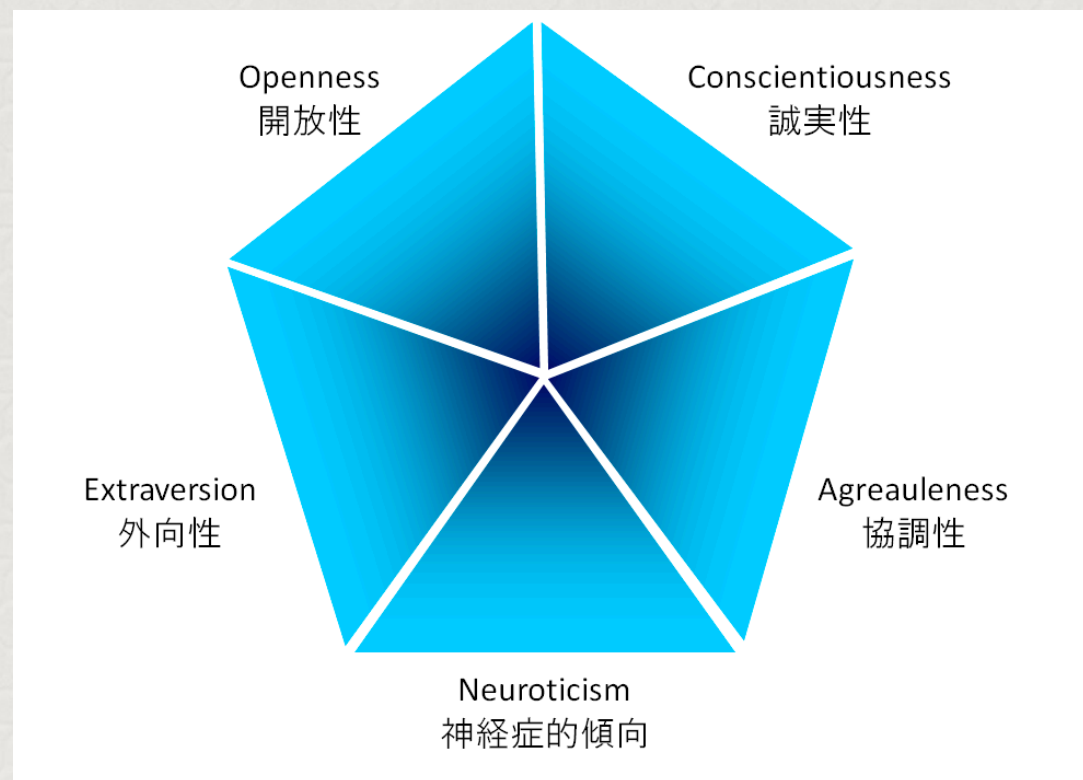
外向性 (Extraversion)

- ✦ 活動性や元気の良さを表す
- ✦ 高い人は社交関係では相手の気をひくことに関心を持ち、より支配的な立場を取りたがる傾向

協調性 (Agreeableness)

- ✿ 集団と一体である度合いを示しており、また同時に自分を見失う度合いでもある
- ✿ 高い人は、親切で気前が良く、だまされやすい一方で、自分なりの考えを持ちにくい
- ✿ 低い人は、利己的で非協力的だが、集団と距離を置いて自分の考えを持つことは得意

Big 5



◀ 衝動的

良 識 性

秩 序 ▶

◀ 神經質

情緒安定性

樂觀的 ▶

◀ 保守的

知的好奇心

空想的 ▶

◀ 消極的

外 向 性

活 潑 ▶

◀ 個 別

協 調 性

愛 着 ▶

Feature DataSet

Type A

外向性	0.2
情緒安定性	0.3
開放性	0.7
誠実性	0.2
協調性	0.1

Type B

外向性	0.6
情緒安定性	0.3
開放性	0.3
誠実性	0.4
協調性	0.2

Type C

外向性	0.7
情緒安定性	0.8
開放性	0.5
誠実性	0.9
協調性	0.8

検証

- ★ うまくいっている人とそうでない人の振る舞いの差をとって比較する

例(

- Pull Request* が通るか否かで開発者を分類する
- ある程度の活動後、役職に就けているか否か

Bot作成

Bot作成



Feature DataSet



Model



Text

- ✿ 検証した分析結果を使用し複数のモデルに投入
- ✿ 求める人物像のテキストを取得する

性格類型論

- ✿ 人の性格に関するいくつかの典型的なパターンを作成し，それに個人を当てはめることでその全体的な特徴を捉える手法
- ✿ 手軽で単純明快なことがメリット
- ✿ 3類型や8類型がある

性格特性論

- ✦ それぞれの特性を数値化してその高低を比較する手法
- ✦ 各特性について明確に議論できるのがメリット
- ✦ *Big5*のほか、*3*特性や*I2*特性がある

参考文献

- ✿ ビッグファイブとは？

<https://achievement-hrs.co.jp/ritori/?p=2925>

- ✿ ビッグ・ファイブ(Big5)とは？

<https://star.careers/company/20180727/1474/>

- ✿ *The Development and Psychometric Properties of LIWC2015*