

# auto-sklearnを用いたソフトウェアバグ予測による ソフトウェア品質評価作業の効率化

富士通株式会社 古川 智章

## 開発における問題点

- 品質評価作業が高頻度化
  - アジャイル型開発で定期的にリリースが発生。
  - これに伴い品質評価作業も発生
- 機械学習での自動化にもハードルあり
  - 素人には大変
  - 現場には時間的な余裕なし

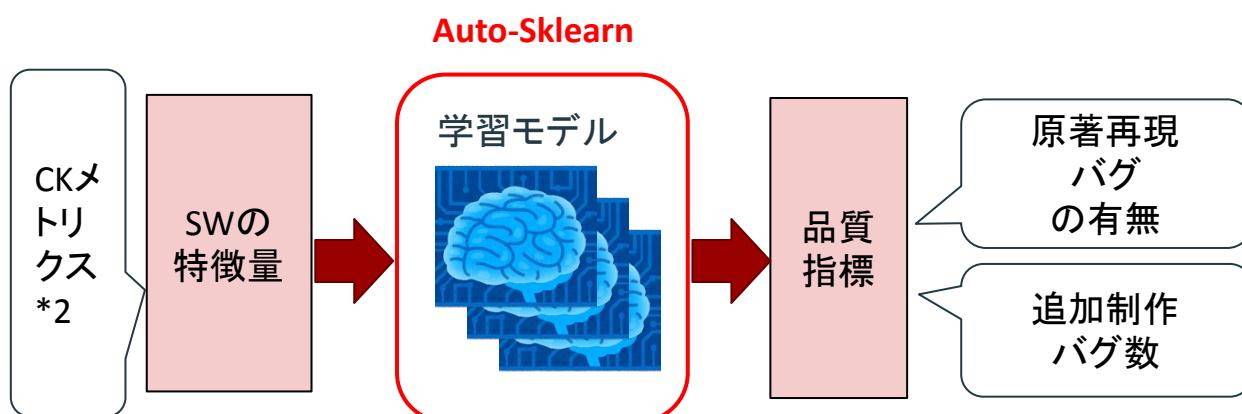
## ツールの適用による解決

- 機械学習による品質評価の自動化
  - 機械学習の自動化
- ↓
- 先行論文\*1からAuto-Sklearnを採用し、機械学習を含めて自動化

## 制作概要

### Auto-Sklearnとは？

- 入力データに適した前処理や予測モデルを自動的に選択
- ベイズ最適化手法にてハイパーパラメータも自動決定
- 特徴
  - メタ学習
  - ベイズ最適化
  - アンサンブル構築

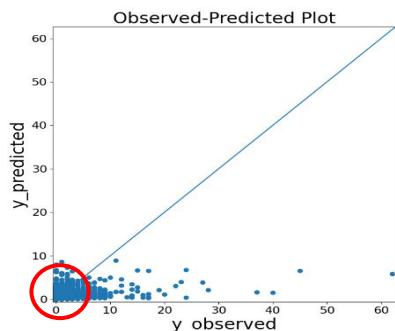


## 評価

- 2値分類モデル (原著論文再現)
  - 原著論文と同等
  - 学習時間の影響評価: 8通り

学習時間	AUC
5分~20分	0.731
25分, 30分	0.789
60分	0.792
90分	0.791

- 回帰モデル (新規の取り組み)
  - 学習時間: 8時間
  - MAE: 0.418(件)
  - MAEが1以下に収まっており実用上有望
  - 正解値と予測値をプロット: より深い分析が必要



## まとめ

- 制作内容
  - 2値分類モデル (原著論文再現)
  - 回帰モデル (新規の取り組み)
- 制作結果
  - 現場で使用する情報として有望
  - バグ数を基に強化試験/強化レビューが実施可能
- 現場適用における課題
  - 特徴量の収集
  - バグ数の調査

\*1: 田中和也, 門田 暁人, Zeynep Yücel, ソフトウェアバグ予測におけるauto-sklearnの有効性評価, コンピュータソフトウェア, 2019, 36巻, 4号, p. 4\_46-4\_52  
 \*2: S.R. Chidamber, C.F. Kemerer, A metrics suite for object oriented design, IEEE Transactions on Software Engineering 20 (6), 476-493, 1994-06