

自動分類における データセット作成時の ミ斯拉ベルの調査

(非機能要件分類を例として)

発表者 木村 祐太

要求分析・要件定義

- 利用者のニーズを明確化し，文書に記述
- 記述されるシステム要件は2種類存在
 - 機能要件 ： システムに要求される機能
 - 非機能要件 ： システムの性能（製品品質）
- システム要件は，設計・実装・テストで利用

要件の抽出とソフトウェア開発

利用者のニーズを明確にし，文書化



要件書からシステム要件を抽出



抽出した要件をもとに
設計・実装・テストを実施



納品・リリース



要件の抽出とソフトウェア開発

利用者のニーズを明確にし，文書化



要件書からシステム要件を抽出

見落とし
解釈の誤り



抽出した要件をもとに
設計・実装・テストを実施



納品・リリース



要件の抽出とソフトウェア開発

利用者のニーズを明確にし，文書化



要件書からシステム要件を抽出

見落とし
解釈の誤り



手戻り
システムの故障
につながる

非機能要件の見落とし

- 機能要件に比べ、非機能要件は見落としやすい
 - 非機能要件が機能要件に含まれている^[2]
 - 記述の表現が曖昧^[3]
 - システムの規模に伴い、要件定義書の文量が増加
- システムの品質を損なう^[1]

非機能要件の管理効率化を目的とした
自動分類手法が提案されている^{[?][?]}

[1] Brooks Jr., F.P., "No Silver Bullet: Essences and Accidents of Software Engineering,"
IEEE Computer, 4, (Apr. 1987), 10-19.

従来の自動分類手法

- 教師ありアルゴリズムを使用
 - Naïve Bayes, SVM (Support Vector Machine)など
- データセット
 - 一文で判断し, ラベル付け
 - 無作為に非機能要件のみを抽出

リサーチクエスチョン

従来研究に対する疑問

– 非機能要件のラベル付け

- ▶ 文脈を考慮していない
 - 人間は文脈を通して内容を理解

リサーチクエスチョン

文脈を考慮すると要件のラベル付けに
どのくらい影響があるか？

本研究の目的・アプローチ

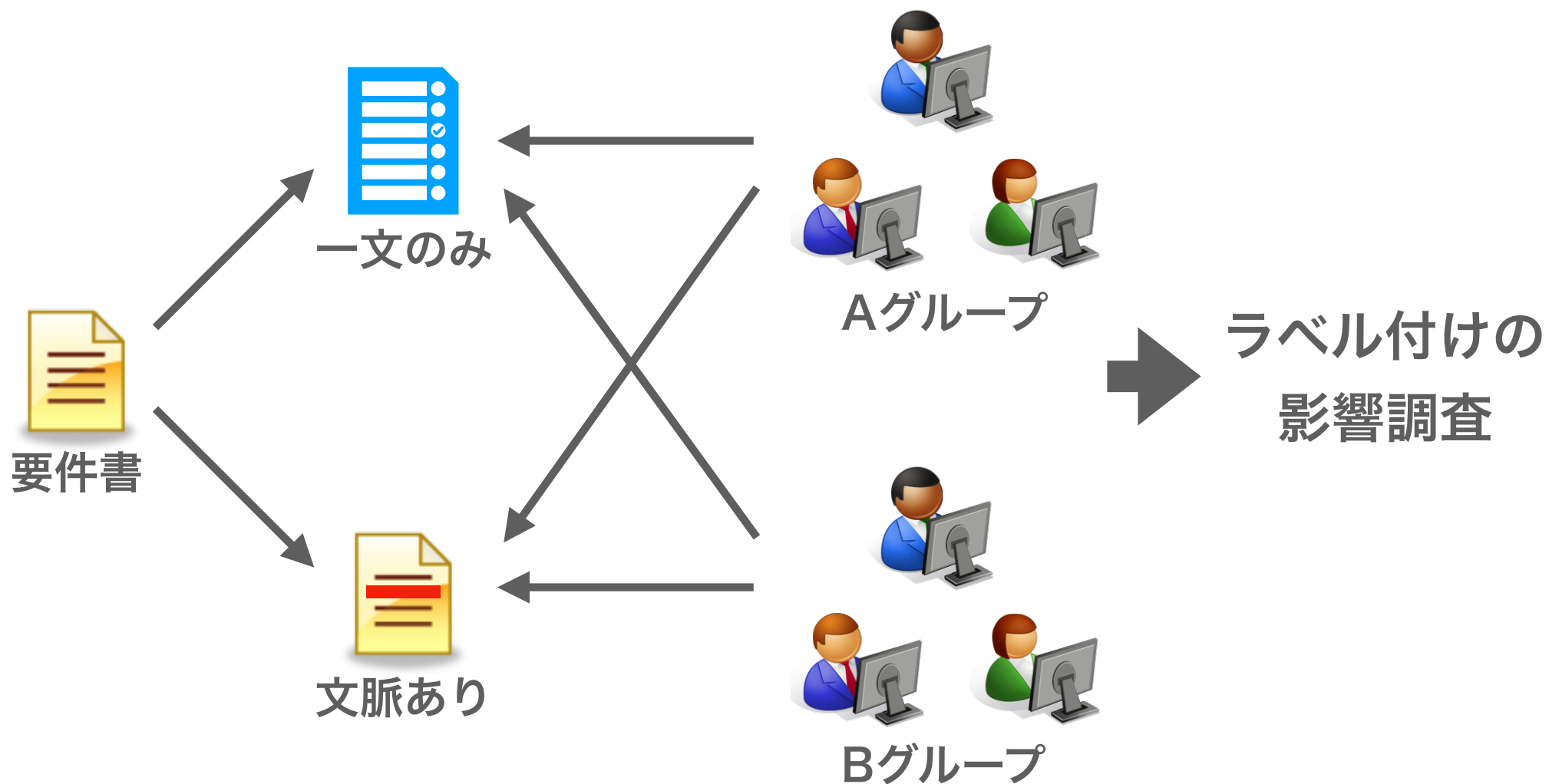
本研究の目的

- 従来研究におけるデータラベリングの検証

アプローチ

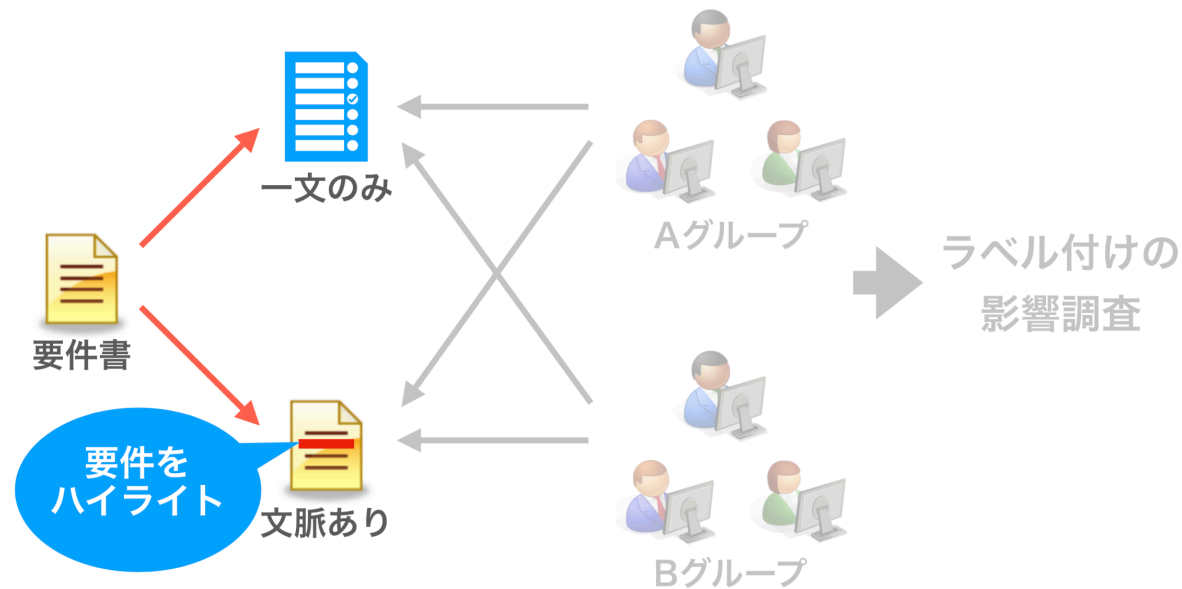
- 一文のみ，文脈ありでラベリングを実施
- 各方法のラベリング結果の違いを調査

実験の概要



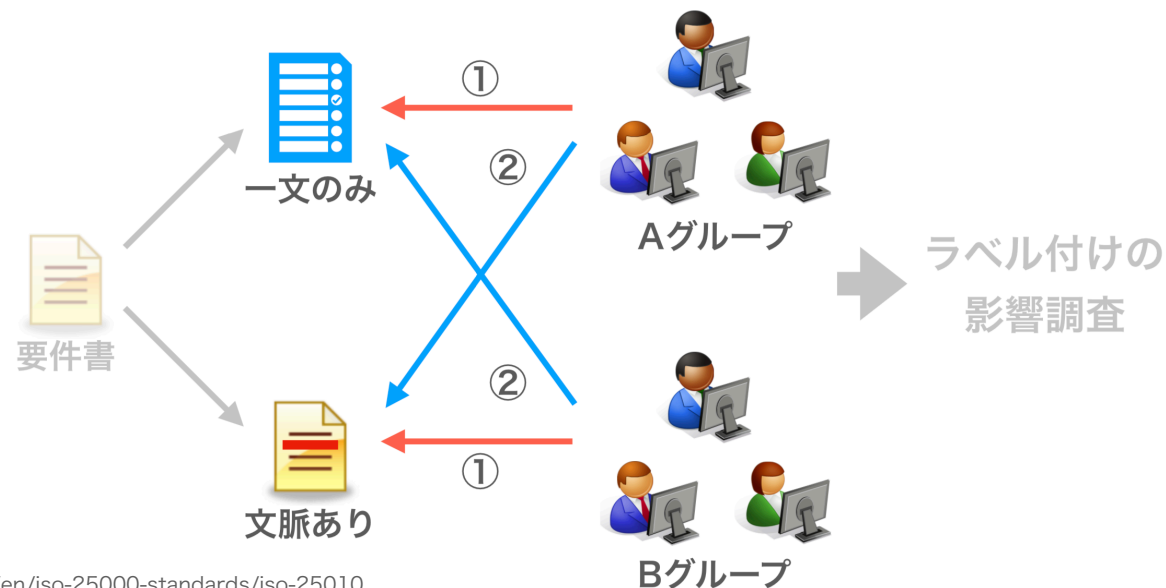
1. データの準備

- 一文のみ
 - 要件書からランダムに一文ずつ
- 文脈あり
 - 要件を抽出し，周辺の変項はそのまま



2. 被験者実験

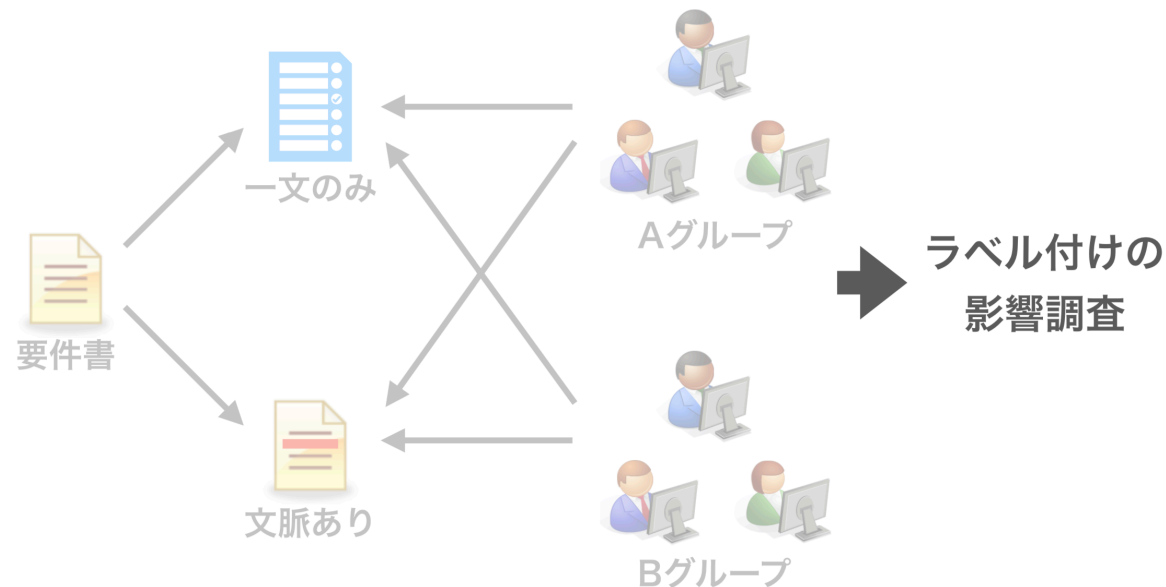
- 2つのグループを用意
- 2つのパターンを実施
 - 一文のみ → 文脈あり, 文脈あり → 一文のみ
 - 8つの品質特性^[?]で分類



[?] <https://iso25000.com/index.php/en/iso-25000-standards/iso-25010>

3. 影響の調査

- 一文のみ，文脈ありのラベルを比較
- 違いが見られた場合，影響があると判断
 - ラベル付けのプロセス見直し？



今後の予定

- 被験者実験の実施（11/13 ～ 11/20）
- 分類結果の分析（11/20 ～ 11/24）