

S09 不満投稿の内容に基づいた顧客属性間における格差の検出

石黒真生, 三好佑樹, 本多敏明, 齊藤史哲 (青山学院大学)

研究背景

様々な格差が社会問題となっている
格差に直結する不満には
どのような偏りがあるのか？

- ・ 地域格差
⇒ 地域によって得られるサービスの差
- ・ 世代格差
⇒ 年齢層の違いによる差異
- ・ 情報格差
⇒ デジタルリテラシー



使用データ

株式会社 Insight Tech 様より
ご提供いただいた
不満調査データセット

不満の投稿
(自由記述、ジャンル)

デモグラフィクス
(年代、性別、職業等)

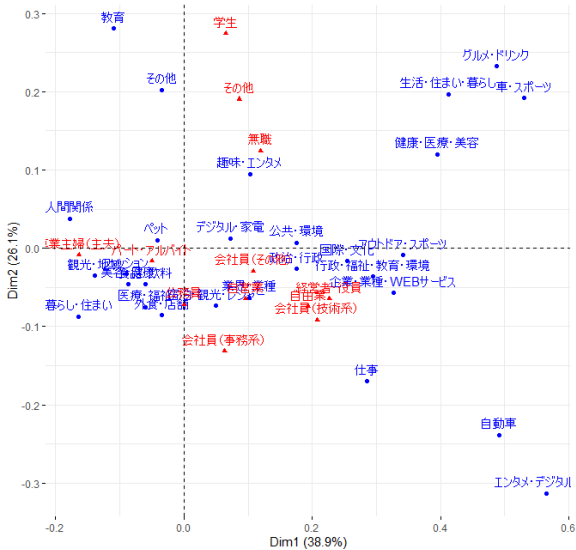
研究目的

何の課題に対して「誰が困っているのか？」を特定する
格差を表すデモグラフィクスの偏りがある不満要因を検出

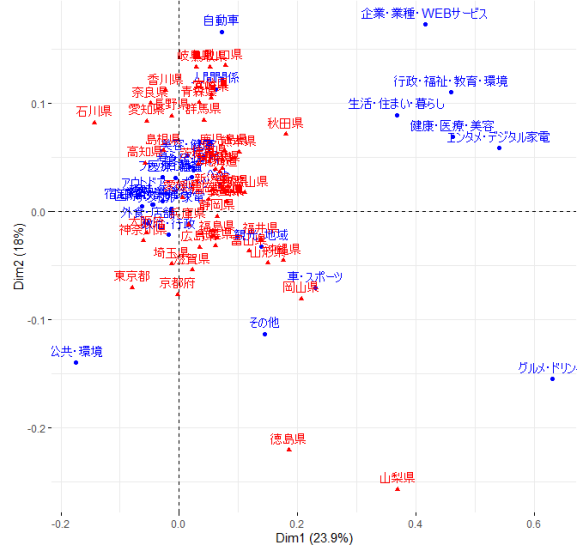


予備解析 (不満投稿の傾向を対応分析)

不満ジャンル×職業



不満ジャンル×地域



不満ジャンルに傾向があることが確認できる！

今後は「ジャンルのカテゴリの細分化」「トピックの検出」「性別・年代・収入など他属性の解析」などが考えられる

解析方法について

- 1) BERTopicによるトピックの獲得と文書分散表現の獲得
- 2) 文書の埋め込みベクトルをクラスタリング
- 3) 構成比に偏りが生じるクラスタを検出
- 4) XAIにより偏りに影響するトピックを検出

どの特徴量 (トピック) が
クラスタあるいはデモグラを
特定するのに寄与するか？



不満コメント

不満のクラスタ

不満の
クラスタ

不満の
クラスタ

不満の分散表現をクラスタリング

- ・ 疑似ラベルをデータに付与
- ・ モデルを訓練
(ランダムフォレストやニューラルネット)
- ・ モデルにSHAPを適用する

疑似ラベルにSHAPを適用

参考文献

- [1] M. SUZUKI et al. "An Analysis on the Customer Logistic Satisfaction based on Word Clustering," Innovation and Supply Chain Management, Vol. 17, No. 1, pp.11-16, (2023)
- [2] 佐々木ら, "疑似ラベルに基づいたクラスタリングに対する局所説明の付与," 電子情報通信学会論文誌D, Vol.J108-D, No.06, pp.402-413 (2025)