

“役立つ不満” 抽出のためのコーパス構築とトピック分類

株式会社 Insight Tech(2017): 不満調査データ. 国立情報学研究所情報学研究データリポジトリ. (データセット) (2016). <https://doi.org/10.32130/idr.7.1>

“役立つ不満” とはどのようなことか

- 不満自体が悪いものではない
- よし悪しの判断が難しい

[例] 会場が狭い
 学校 : 改善不可
 ライブ会場: 会場変更可能

トピックに分けることで判断しやすく

評価基準

改善案が思いつく	役立つ不満
改善案が思いつかず 第三者が閲覧して不快に思う	参考に しにくい不満
改善案が思いつかないが 第三者が不快に思わない	一般的な不満

BERTopic でのトピック分類

データ: 不満調査データのサブカテゴリ
 「ライブ・コンサート・カラオケ」
 モデル: 学習済みの言語モデルをイベント不満で
 Fine-Tuning し、イベント不満に特化した
 言語モデルを構築したものを使用

作成されたトピックの一部をアノテーションに使用

“役立つ不満” のアノテーション



“役立つ不満” サブデータセット

Topic: ライブの座席やチケット代に関する不満

text	ans1-1	ans1-2	ans2-1	ans2-2	ans3-1	ans3-2
コンサートで後ろの席でも同じ値段なこと	1	3	1	3	1	3
ライブに行くけど周りがほとんど若い子...	2	6	1	3	2	6
地方での公演が少なすぎる。地方でもや...	2	6	1	3	2	6
見たいコンサートがないので残念です。	2	6	1	4	2	6
思春期の子供に悪影響だと思う。	2	5	2	5	2	6

Topic: ファンクラブに関する不満

text	ans1-1	ans1-2	ans2-1	ans2-2	ans3-1	ans3-2
ファンクラブに入らなくても買えるよう...	2	6	1	3	1	3
転売目的でチケットを買う人に重い刑罰...	1	4	1	4	1	3
ファンクラブに入ってもチケット買え...	2	5	1	3	2	6
ファンクラブの特典がほとんど無いに等...	2	6	1	3	2	6
入ってない子にチケット譲ると言われ...	2	6	2	6	2	6

Topic: カラオケの部屋や採点に関する不満

text	ans1-1	ans1-2	ans2-1	ans2-2	ans3-1	ans3-2
隣の人の声が聞こえてて自分が歌いにく...	1	3	1	3	2	6
歌ってないときも何やら反応して気持ち...	1	4	1	3	2	6
カラオケで部屋を覗いてくる人が多い。	1	4	1	4	2	6
コンピュータなんかは俺の歌の良さがわ...	2	5	2	6	2	6
大人数なのに狭い部屋に案内された。	2	6	1	3	1	3

1人分のラベル 1人分のラベル 1人分のラベル

ラベルの種類

- 1 ... 改善案が思いつく
 - 2 ... 改善案が思いつかない
 - 3 ... 主催者側で解決可能
 - 4 ... 他者の働きが必要
 - 5 ... 不快に感じる
 - 6 ... 不快に感じない
- ans1 (1, 2)
ans2 (3, 4, 5, 6)

考察・まとめ

- 不満を先にトピックに分類することで不満の分類が可能になる
- 不満にアノテーションすることでサブデータセットを構築した

◎色々な研究課題に応用できる可能性