

good_kensaku();

midterm presentation

Fujita, Masui, Maeda,
Nagata, Nakamura, Uchida

進捗(1/4): スクレイピング

現在、スクレイピング対象サイトを拡大中
悪い評価のサイトデータが少ないため、増やしている
得られたデータは、既に学習モデルに組み込み済み



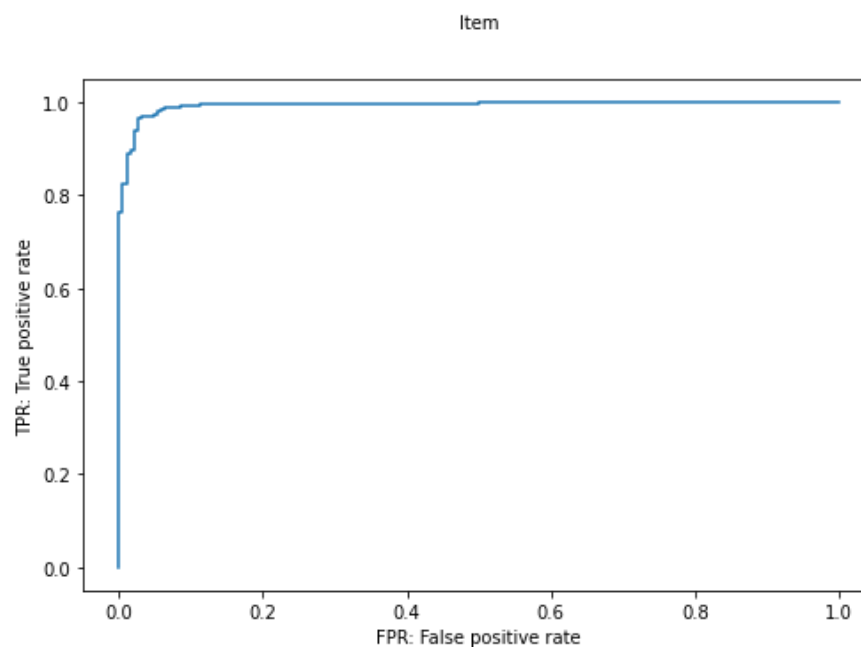
進捗(2/4): 学習モデル

どのモデルを使用するかが決まった

Bag of Words + lightGBM

<理由>

- モデル作成が高速
- 精度がかなり出ている



補足: 学習モデル

その他検討していたモデルについて

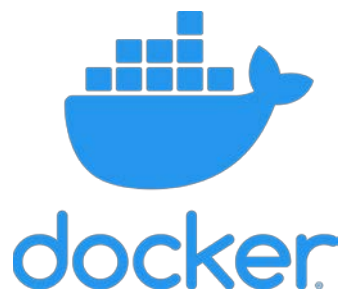
- BERT

文章の切り方などを工夫
未知の文章に対する精度が
出なかったため、断念

- FastText+RNN

Bag of Wordsに比べ、精度が
出なかったため、断念

進捗(3/4): システム化



開発環境の決定



メモリ保存による高速化

進捗(4/4): Chrome拡張機能

拡張機能で表示されるアイコンの作成(仮)

UIのバグ修正

システムに分類器を組み込んだ

バグが見つかったため、現在デバッグ中