

深層学習を利用した FAQ検索システムへの説明可能AIの適用

富士通株式会社

飯塚 健男

iizuka.takeo@fujitsu.com

FAQ検索システムにおける課題

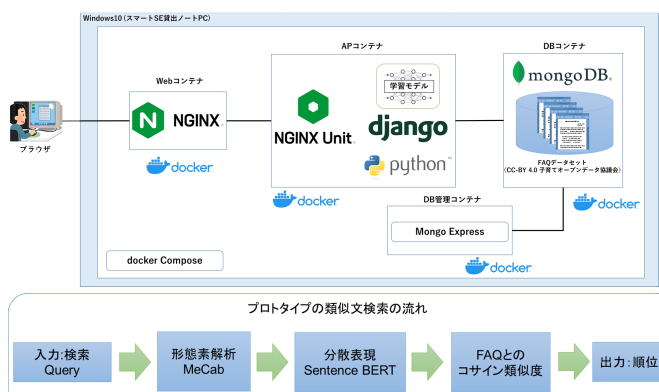
FAQの検索結果(ランキング)に対し、検索結果理由について、な数値的根拠を示すことが困難

FAQ検索結果をお客様が想定するランキングにするための、対応方法(教師データの追加根拠)が不明確

手法・ツールの適用による解決

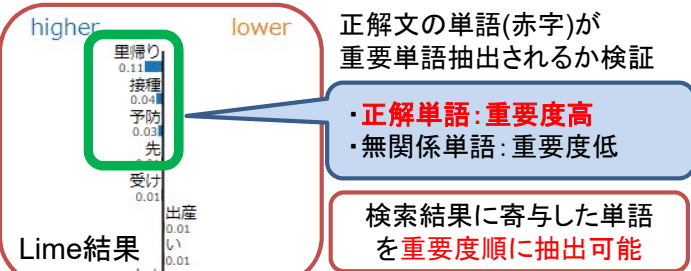
深層学習を使用したFAQ検索システムのプロトタイプを作成。
プロトタイプに対し説明可能AI(Lime)を適用し
Limeによる重要語の抽出結果による、FAQ検索結果の説明性および、想定結果実現に向けた教師データへの言及性について検討する

FAQ検索プロトタイプ構成



検索結果の重要語抽出結果

検索文: 近所に親類がいないため、**里帰り**出産を考えています。
里帰り先ではしばらく滞在を予定していますが
そこで**予防接種**は受けられるでしょうか



教師データ比較検証

以下の教師データを追加した学習モデルを比較

- ・検索文の原文追加モデル
- ・重要語使用の短文追加、重要語追加モデル

検索結果 (順位)		Lime結果	
順位	問い合わせ文	higher	lower
1	前期に保育利用料軽減の申請をしたが、9月以降は申請できませんか?	申請	軽減
79	療育医療はどんな時に申請できるのですか?	療育	医療

原文追加モデル

検索結果 (順位)		Lime結果	
順位	問い合わせ文	higher	lower
1	病児・病後児利用料補助	病児	補助
62	療育医療はどんな時に申請できるのですか?	療育	医療

重要語追加モデル

検証結果

- ・Lime結果
どちらも重要語「結核」が抽出されたが**値は同等**
- ・順位: **同等**
※追加前: 140位
- 原文追加: 79位
- 重要語追加: 62位

若干重要語モデルの順位が高いが、**ほぼ同等の結果**
重要語追加モデルで、より想定通りの結果が得られるという仮説を裏付けできなかった

考察・今後の課題

■教師データの追加についての考察

重要語を使用した短文を教師データとして追加することにより、重要語をより強調する(重要度を高くする)試みについては仮説を裏付けることができなかった。これは本システムで使ったSentence BERTが学習データ全体から大域的な特徴を学習しているため、局所的な情報による追加では、補正にとどまり、ランキングが大きく変動するような学習が困難であったと推測される。

■今後の課題・検討事項

Sentence BERTのような大域的な特徴により類似度が決定するような、大域的な特徴を考慮可能な説明AIの検討

学習モデルだけでなく、ルールベースを取り入れたハイブリッドAIといった手法の検討も併せて進めていきたい。