未知語分散表現を用いた品詞解析

三林亮太 † 五十川真生 ‡ 永田亮 † 荒瀬由紀 ‡ 梶原智之§ † 中南大学 ‡ 大阪大学院情報科学研究科 § 大阪大学データビリティフロンティア機構

リサーチクエスチョン

Q. 未知語をより考慮した分散表現を用いれば 品詞解析の性能はどのくらい向上するのか?

しかし



背景

品詞解析には綴り誤りにより解析ミスを引き起こすという問題がある.



*Becose これは固有名詞だ.

(正:Because 接続詞)

単語内の文字列に基づき 単語の情報をベクトルに変換 似たベクトルになる

それらを用いることにより綴りが 誤っている単語を補完することが でき品詞解析の性能が改善された.*1

未知語をより考慮した
そこで分散表現を用いることで、
品詞解析の性能は向上する
のではないだろうか

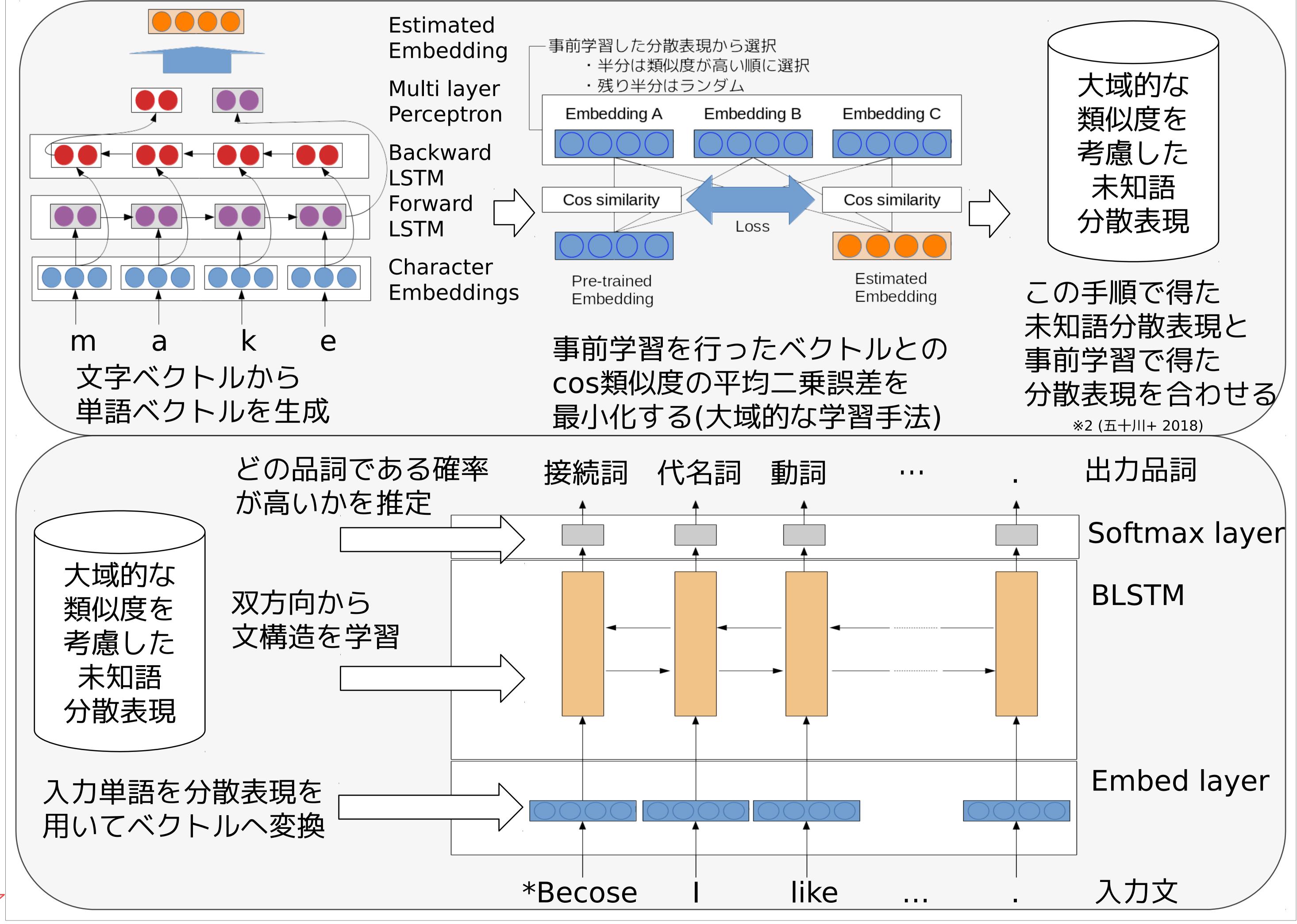
連絡先 三林亮太:s1671117@s.konan-u.ac.jp



※1(永田亮+ 2018)

提案手法

大域的な類似度を考慮した未知語分散表現※2を用いて品詞解析を行う.



今後の予定

提案手法に基づき学習を行い、品詞解析結果を考察する.

未知語分散表現をCNNを用いて生成したものに変更するなど他の検証を行う.