

先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	知能コミュニケーション (中村 哲 (教授))		
学籍番号	2211176	提出日	令和 6年 1月 19日
学生氏名	田中 康紀		
論文題目	自然言語処理モデルにおける追加学習の最適化に関する研究		
要旨			
<p>ニューラルネットにおける自然言語処理は、大規模な事前学習とfine-tuningを組み合わせることで、下流タスクにおいて高精度なモデル構築を実現してきた。特にTransformer構造を持ったモデルはパラメータ数・データ量・計算回数に対してその性能がスケールすることが知られている。fine-tuningは既存の事前学習モデルをもとにすることから、パラメータ数が固定される。従って、fine-tuning後の性能は、計算回数あるいはデータ量が多いものが一般に性能が高くなる。本研究では、Transformerモデルの性能を支配する、計算回数に対する削減とデータ量の拡張を行い、そのモデルへの影響を検証する。</p> <p>計算回数に関しては、バイトレベルモデルのAttention情報から、重要度を計算し、文字結合を行う。これにより、系列長の削減による計算速度の向上と、文字結合による意味情報の保持を検証する。テキスト分類実験において、基盤モデルをfine-tuningしたものに比べ1.78倍の高速化、ランダムに結合する手法に比べ1.8ポイントの分類精度向上を示した。</p> <p>データ量に関しては、機械翻訳タスクにおいて、単言語コーパスと学習済み機械翻訳モデルを用い、その翻訳結果を品質推定器により評価し、この評価を最大化することで、対訳データ無しに機械翻訳器を改善するデータ拡張手法を提案する。翻訳精度の改善には至らなかったが、報酬と方策関数の改善の余地を示した。</p>			