

論文 / 著書情報  
Article / Book Information

題目(和文)	
Title(English)	Improving Word Representations for Language Modeling
著者(和文)	FENGYukun
Author(English)	Yukun Feng
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12394号, 授与年月日:2023年3月26日, 学位の種別:課程博士, 審査員:奥村 学,熊澤 逸夫,中山 実,篠崎 隆宏,船越 孝太郎
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12394号, Conferred date:2023/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

## 論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	Feng Yukun	
論文審査 審査員		氏名	職名	氏名	職名
	主査	奥村学	教授	船越孝太郎	准教授
	審査員	熊澤 逸夫	教授		
		中山実	教授		
篠崎隆宏		准教授			

### 論文審査の要旨 (2000 字程度)

本論文は「Improving Word Representations for Language Modeling」と題し、英文全5章より構成されている。

第1章「Introduction」では、本研究の背景を説明するとともに、本論文の目的を述べている。近年成功を取めている文字レベルのニューラル言語モデルには、依然として、低頻度語間の意味的関係を学習するのに有効でないこと、単語レベルの情報を持ち込む効果的な方法が明らかでないことが課題として残っていることをまず述べている。そして、文字レベルの言語モデルを改良する2つの提案手法について、その概要を説明している。1つ目の課題を解決する1つ目の研究では、Continuous Bag-of-Words (CBOW)モデルに単語クラスタを導入する単純でかつ効果的な手法を提案することで、より高度な単語埋め込みを作成できること、2つ目の課題を解決する2つ目の研究では、従来とは異なり、単語の系列の情報を出力層において持ち込むことで、文字レベルの言語モデルと単語レベルの言語モデルの組み合わせを実現できることを述べている。

第2章「Related Work」では、本研究に関連する、単語埋め込みに関する研究、単語クラスタの応用に関する研究、単語レベルの情報を文字レベルの言語モデルに持ち込む先行研究について説明している。

第3章「Incorporating Word Clusters into CBOW」では、まずベースとなるCBOWモデルについて述べた後、低頻度語を単語クラスタと置換する提案手法について述べている。提案手法には、1) 多くの低頻度語が単語クラスタと置換されることで、劇的に少ないパラメタ数で十分であること、2) 低頻度語の埋め込みが、単語クラスタの埋め込みで初期化されることで、言語モデルのfine-tuning時にはより良いものとなること、3) 低頻度語が置換され文脈中のノイズが減ることで高頻度語の埋め込みもより良くなることなどの利点があることを述べている。提案手法の評価実験は、10種類のlanguage modelingデータセットと2種類の機械翻訳用データセットにおいて、類型学的に様々な言語について行われ、従来手法を上回る性能を発揮することを確認している。

第4章「Incorporating Word-level Information into Character-aware NLMs」では、従来手法とは異なり、単語の系列の情報を出力層において文字レベルの言語モデルに持ち込む提案手法について述べている。提案手法は、入力側で単語の情報を持ち込む従来手法と比べ優れていることについて説明するとともに、従来手法と組み合わせても利用可能であり、更なる性能向上が実現可能であることについて述べている。提案手法の評価実験は、14の類型学的に異なる言語について行われ、従来手法を上回る性能を発揮することを確認している。また、従来手法と組み合わせることで、更なる性能向上が得られることを示している。最後に、文字レベルの言語モデルにおける単語レベルの情報の有効性について、どのような単語（例えば、低頻度語）の情報を持ち込むことが効果的であるかを分析している。

第5章「Conclusion and Future Work」では、本研究の結論と、第3章と第4章で提案した手法に対する課題と今後の展望について述べている。

以上を要するに、本論文は、文字レベルの言語モデルの高度化のために、2つの方向性で新しい手法を提案し、その有効性を検証しており、工学上貢献するところが大きい。よって博士（工学）の学位を授与するに十分な価値を持つものと認められる。