

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)	自然言語処理学 (渡辺 太郎 (教授))		
学籍番号	2411029	提出日	令和 8年 1月 19日
学生氏名	岩國 巧		
論文題目	誤り単語の混入による単語単位翻訳誤り検出の頑健性評価		
要旨			
機械翻訳の発展により翻訳精度や流暢さは大きく向上しているが、固有表現や数値といった重要語の誤訳は依然として残存している。 特に医療や金融などの専門領域では、こうした誤訳が意思決定や実務に影響を及ぼす可能性があるため、単語レベルでの誤りを高精度に検出できる翻訳評価指標が求められている。 本研究では、ニューラルベースの翻訳評価指標 XCOMET における単語レベルの誤り検出性能を調査するため、大規模言語モデルを用いて言い換え文を生成し、数値および固有表現を書き換えた擬似誤訳文から単語単位の誤訳検出データセットを構築した。 金融ドメインの英日翻訳コーパス Timely Disclosure Documents Corpus を用いた実験の結果、XCOMET は誤訳に対して高い感度を示す一方で、誤りを含まないトークンに対する過検出が多く、特に文頭付近で顕著であることが明らかとなった。 これらの結果から、XCOMET の単語誤り検出能力には一定の有効性が認められる一方で、誤り位置の正確な特定や局所的な過検出の抑制に関しては改善の余地があることが示された。			