

氏名	SILAA VICTOR ALEX		
授与学位	博士(工学)		
学位記番号	博甲第211号		
学位授与年月日	令和5年9月5日		
学位授与の要件	学位規則第4条第1項		
学位論文題目	Web-based Safari review system development using Microblog analyzed data (マイクロブログの解析データを利用したWebベースのSafariレビューシステム開発)		
論文審査委員	主査 教授	梶井 文人	
	准教授	フダシスキ ミハウ エト`ムト`	
	教授	前田 康成	
	准教授	澁谷 隆俊	
	准教授	館山 一孝	

学位論文内容の要旨

In this thesis, I propose the use of online microblogs as review supplements and demonstrate their applicability through a designed tourist support system that aims to provide additional opinions and up-to-date points of interest to the less-known tourist spots.

The proposed approach folds into three. First, through the use of geotagged tweets. Tweets that contain geolocation information are considered geotagged and therefore treated as possible tourist on-spot opinions. The main challenge, however, is to confirm the authenticity of the extracted tweets. This approach includes the use of location clustering and classification technique. Specifically, extracted geotagged tweets are clustered by using location information and then annotated taking into consideration specific features applied to machine learning-based classification techniques.

As for the machine learning (ML) algorithms, I adopt a fine-tuned transformer neural network-based BERT model which implements the information of token context orientation.

Second, I further studied geolocation estimation from untagged tweets that have no geotagged attached. To accomplish this, I proposed the use of a two-stage process, a transformer model for the classification of primary tweets, and a combination of indicative words like location mention or event mention for location inferring.

The problem here is that Twitter data is typically noisy and consists of ungrammatical or informal phraseology and non-standard vocabulary, which additionally causes the feature sparsity problem, resulting in low classifier performance. To address this, I specifically evaluate a range of pre-processing techniques for text categorization to accurately obtain a proper set that collectively contributes to the improvement of prediction accuracy.

A classification framework created here relies on a fine-tuned transformer neural network model which learns from tweet contents and predicts the locations from which those tweets were sent - with a limited application in the detection of widely known general locations - such as tourist spots. I evaluated the combined impact of indicative words like location mention, and event mention on the model geolocation estimation, and model accuracy improvement when indicative words are involved compared to when they are not.

Thirdly, I demonstrate the applicability of these two approaches by designing a tourist support system and mapping extracted opinions to their respective tourist spots as touristic information.

審査結果の要旨

SNS サービスを利用した観光関連のロコミ情報の活用が普及している。しかしながら、投稿者が位置情報を付与する確率は低く、投稿情報を地図情報に対応づけるマッピングサイトやナビゲーションサービスにおいては、これが誤差や誤りを発生させる要因となっている。

著者は、AI 技術を用いて、現存する位置情報付きの投稿テキストを用いて位置情報推論モデルを構築し、位置情報無しの投稿テキストから位置情報を高い精度で推論する機構を構築した。特に、これまでに試みられた報告がない Transformer モデルを用い、従来の類似手法よりも高い推論性能が発揮できることを明らかにした。

第一に、観光スポットを訪れたユーザが投稿した大量の SNS テキストデータを収集し、これをクリーニングする過程、さらに、BERT や DistilBERT, Bertweet といった複数の Transformer モデルを用いた検証実験を実施し、複数の指標を用いて詳細な評価を行い、DisilBERT が最も良い性能を発揮することを明らかにするとともに、その性能限界についても議論している。

第二に、著者は、実装した機構を応用してタンザニア国立公園を対象とした SNS マッピングシステムを構築、これがユーザビリティやサービスの観点から有効に機能することを論理的に明らかにした。

これを要するに、著者は、生成系 AI 技術を駆使した位置情報の推論と自動付与による新たな支援技術に関する新知見を得たものである。これは、観光情報学分野および関連分野において貢献するところ大なるものがある。よって著者は、北見工業大学博士（工学）の学位を授与される資格があると認める。