Title	Emotion Analysis Model Using Dialect Corpus and Proposal of Flaming and Cyberbullying Detection Method
Author(s)	加藤, 大造
Citation	
Issue Date	2023-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/18753
Rights	
Description	Supervisor: NGUYEN, Minh Le, 先端科学技術研究科, 修士(情報科学）

Emotion Analysis Model Using Dialect Corpus and Proposal of Flaming and Cyberbullying Detection Method

2130006 KATO, Taizo

It has been a long time since users were able to post their own feelings and thoughts freely and easily on the Internet through features such as posting to social networking services (SNS) and bulletin boards or commenting on video sharing services. With 74.2% of individuals using SNS in modern society, online conversations have become a part of daily life. These conversations are typically spoken language and are often spoken with regional dialects reflecting the user's place of residence or born and raised. The volume of such dialect-infused text data is on the rise, a natural language processing (NLP) models that can understand these dialects is required. In this study, we hypothesize that text containing dialects more strongly reflects the writer's emotions. We built a dialect corpus of approximately 320,000 instances gathered from dialect dictionaries and Twitter to train a variation of the BERT language model, which is named "DialectBERT". We fine-tuned this model for analyzing the intensity of eight emotions (joy, sadness, anticipation, surprise, anger, fear, disgust, and trust) and their polarities. As a result, we confirmed that DialectBERT could correctly analyze six out of the eight emotions (joy, sadness, anticipation, surprise, anger, disgust) more accurately than existing models. DialectBERT also outperformed in terms of sentiment polarity analysis. Further, we demonstrated that comparable accuracy can be achieved with between 100,000 and 150,000 training instances.

The use of SNS becomes commonplace, problems such as online flaming and cyberbullying have become social issues. To address this, we collected conversational data from Twitter containing words related to flaming and cyberbullying, and then labelled data where these issues occurred. Using these data and the eight-emotion analysis model, we analyzed the emotional intensity of each conversation and created emotional vectors. These vectors were then used to detect incidents of flaming and cyberbullying through vector similarity and machine learning algorithms. In all cases, models using DialectBERT yielded better detection accuracy. This study demonstrated that a BERT model trained with a dialect corpus can more accurately analyze emotional intensity, and that this model can effectively detect online flaming and cyberbullying incidents.