



COMP-05

SunthorNLPhu: Semi-Automatic Natural Language Processing Poet

Jeerapat Suanthong, Chaturapatr Benjasatkul, Matin Kositchutima

Kamnoetvidya Science Academy

Supervisor: Prompong Pakawanwong

Email: jeerapat_s@kvis.ac.th

Thai octameter poem is one the most popular types of poem in Thailand. It features both the beauty of words and the stories as presented in well-known Thai literature. However, technological development resulted in rapid changes in communication to be faster. Words have become shorter and more concise. They are straying further from the nature of the octameter poem, which presents the beauty of words and rhyming scheme. This change makes Thai octameter poems limited only to a certain group with an interest in literature and not as well-known as in the past. Moreover, Thai students have yet to indulge in the true values and beauty of Thai poetry. Thai octameter poem only serves as an activity among Thai scholars.

Deep learning and natural language processing are increasing in popularity; they present a decent application in generating new text by learning and processing existing text. These new technologies allow machines to achieve almost the same poetry-writing capabilities as humans. As shown in generating the European Sonnets seen in Deep-speare. These artificial intelligence technologies have not yet been applied to the synthesis of Thai octameter poems. Thus, we wish to develop such artificial intelligence that can generate Thai octameter poems using transformer-based models along with WangchanBERTa, a Thai NLP model developed on a large Thai language corpus. Additionally, we integrated rule descriptors to each wak to increase the models' efficiency in learning rhyming schemes and the complex rhyming pairs in the Thai language. The model is trained with Thai octameter poems by Sunthorn Phu, a world-renowned Thai poet.

Keywords: Natural Language Processing, Thai Octameter Poem, Bidirectional Encoder Representations from Transformers