Presentation

Nils Schaetti

Author Verification in Stream of Text with Echo State Network-based Recurrent Neural Models

Abstract:
This paper evaluates a type of recurrent neural networks (RNN) named Echo State Network (ESN) on a NLP task referred as author verification. In this case, the model has to identify whether or not a given author has written a specific text. We evaluate these models on a difficult task where the goal is to detect the author in a noisy text stream being the result of a collaborative work of an unknown number of authors. We construct a new dataset (denoted SFGram) composed of science-fiction books, novels and magazines. From this dataset we select three authors, published between the 1953 and the 1974, and we evaluate the effectiveness of ESNs with word and character-based representations to detect these authors in a set of 91 science-fiction magazines (containing around 8M of words).

Biography: Nils Schaetti is a Ph.D. student in the field of Machine Learning and Natural Language Processing at the University of Neuchâtel. He received his Bachelor and Master degrees in Computer Science from the University of Franche-Comté (France) in addition to a CFC and an IT technician diploma from the CFPT in Geneva. From 2009 to 2016, he held full-time positions in various private companies and organizations such as the EPFL and the Kudelski Group as a computer engineer and HPC specialist. His research interests are in Machine Learning and Artificial Intelligence, Natural Language Processing and Authorship Attribution.

Organization: Université de Neuchâtel

Contact: nils.schaetti@unine.ch