

Author Profiling: Prediction of Learners' Gender on a MOOC Platform Based on Learners' Comments

Authors : Tahani Aljohani, Jialin Yu, Alexandra. I. Cristea

Abstract : The more an educational system knows about a learner, the more personalised interaction it can provide, which leads to better learning. However, asking a learner directly is potentially disruptive, and often ignored by learners. Especially in the booming realm of MOOC Massive Online Learning platforms, only a very low percentage of users disclose demographic information about themselves. Thus, in this paper, we aim to predict learners' demographic characteristics, by proposing an approach using linguistically motivated Deep Learning Architectures for Learner Profiling, particularly targeting gender prediction on a FutureLearn MOOC platform. Additionally, we tackle here the difficult problem of predicting the gender of learners based on their comments only – which are often available across MOOCs. The most common current approaches to text classification use the Long Short-Term Memory (LSTM) model, considering sentences as sequences. However, human language also has structures. In this research, rather than considering sentences as plain sequences, we hypothesise that higher semantic - and syntactic level sentence processing based on linguistics will render a richer representation. We thus evaluate, the traditional LSTM versus other bleeding edge models, which take into account syntactic structure, such as tree-structured LSTM, Stack-augmented Parser-Interpreter Neural Network (SPINN) and the Structure-Aware Tag Augmented model (SATA). Additionally, we explore using different word-level encoding functions. We have implemented these methods on Our MOOC dataset, which is the most performant one comparing with a public dataset on sentiment analysis that is further used as a cross-examining for the models' results.

Keywords : deep learning, data mining, gender predication, MOOCs

Conference Title : ICADMA 2020 : International Conference on Advanced Data Mining and Applications

Conference Location : Sydney, Australia

Conference Dates : January 30-31, 2020