Talks@DI

ZERO-SHOT HASHTAG SEGMENTATION FOR MULTILINGUAL SENTIMENT ANALYSIS

Speaker
Marcelo Inuzuka and Ruan Rodrigues

Affiliation
Federal University of Goiás, Brazil

When
December 6th, 14h00

Where
C6.3.27 and
https://videoconf-colibri.zoom.us/j/86353034907

Abstract

For sentiment analysis in medium to low-resourced languages, a multilingual approach that resorts to machine translation can be competitive with standard approaches to the task. We develop a zero-shot hashtag segmentation framework and demonstrate how it can be used to improve the accuracy of multilingual sentiment analysis pipelines. Besides, also being useful for accelerating data annotation, our zero-shot framework establishes a new state-of-the-art for hashtag segmentation datasets, surpassing even previous approaches that relied on feature engineering and language models trained on in-domain data.

Short bio

Marcelo is a faculty member of Federal University of Goiás, Brazil, and he is currently doing research on Argument Mining as a doctoral student at the NLX Group of the Dept of Informatics, Faculty of Sciences of the University of Lisbon.

Ruan is a Machine Learning Engineer with experience in projects and challenges in the field of Natural Language Processing both in academia and industry. He holds a Computer Science degree from the Federal University of Goiás and is currently a student at the Erasmus Mundus European Masters Program in Language and Communication Technologies.