

Naive-Bayesian Classification for Bot Detection in Twitter Notebook for PAN at CLEF 2019

Título Naive-Bayesian Classification for Bot Detection in Twitter Notebook for PAN at CLEF 2019

Autores [Pablo Gamallo](#), Sattam Almatarneh

Tipo Comunicación para congreso

Fonte  [Conference and Labs of the Evaluation Forum](#), Lugano (Switzerland), 2019.

ISSN 1613-0073

Abstract This article describes a system that participated in the Bots and Gender Profiling shared task at PAN 2019. The first objective of the task is to detect whether the author of a Twitter account is a bot or a human; and in case of human, the second objective is to identify the gender of the user account. For this purpose, we present a Bayesian strategy based on features, including specific content of tweets and automatically built lexicons. The best configuration of features reached 0.88 accuracies in the official Spanish test dataset and 0.81 in the English one for the bot/human classification. For gender profiling, the scores we obtained were lower, around 0.70.

Palabras clave Bot Detection, Gender Detection, Naive-Bayesian Classification

LIGAZÓNS

 [Versión da editorial](#)

DESCARGAS

 [Referencia BibTex](#)

 [Descargar versión do editor](#)