

Humain Contre Robot : Différencier l'expression des IA génératives de celle d'auteurs humains en utilisant l'Analyse de Données Textuelles

Jean Moscarola¹, Zakari Chauhan²

¹Université Savoie Mont Blanc - jmoscarola@gmail.com

²University of Toronto - zakarichauhan@outlook.com

Abstract

This paper focuses on the differences between texts written by humans, or generated by artificial intelligence. The aim is to understand the capabilities and limitations of a chatbot compared with human output. To this end, a corpus of 206 academic assignments, comprising pairs of essays written on identical topics by humans and by ChatGPT, was analyzed. Textual data analysis and artificial intelligence were used to show that humans and the robot have distinct and recognizable styles. This research also explores the effectiveness of available methods for detecting misuse, and proposes recommendations for ethically integrating AI tools into education.

Keywords : author attribution, classification, thesaurus, Sphinx, ChatGpt

Résumé

Cette communication porte sur les différences entre textes écrits par des humains et générés par une intelligence artificielle. L'objectif est de comprendre les capacités et limites d'un chatbot par rapport à la production humaine. Un corpus de 206 devoirs académiques, composé de paires d'essais écrits sur des sujets identiques par des humains et par ChatGPT, a été analysé. L'analyse de données textuelles et l'intelligence artificielle ont été utilisées pour montrer que les humains et le robot ont des styles distincts et reconnaissables. Cette recherche explore également l'efficacité des méthodes disponibles pour détecter les utilisations abusives et propose des recommandations pour intégrer éthiquement les outils de l'IA dans l'éducation.

Mots clés : attribution d'auteurs, classification, thesaurus, Sphinx, ChatGpt