

Détection automatique de marqueurs de registres littéraires profonds

Véronique Magri¹, Laurent Vanni²

¹Université Côte d'Azur, CNRS, BCL – Veronique.Magri@univ-cotedazur.fr

² Université Côte d'Azur, CNRS, BCL – Laurent.Vanni@univ-cotedazur.fr

Abstract

This contribution proposes a method for identification of markers of literary registers in a contrastive corpus: travel contrastive corpus: travelogues are contrasted with fiction. The corpus consists of texts by an author who has written both travelogues and fiction make up the corpus. The travelogue, a factual text, is associated with the predominantly didactic register - the writer wants to transmit knowledge. The traditional approach of statistical analysis of textual data textual data is compared/complemented by a deep neural network learning approach neural networks dedicated to the task of automatically classifying (predicting) the(prediction) of the literary register. The extraction of textual detected by the AI is analysed as characteristic features of each literary each literary register. More specifically, the formal contrasts are postulated as markers of fictionality, or even of literary literariness as a register for fiction.

Keywords: motifs, pattern, literary registers, CA, z-score, deep learning

Résumé

Cette contribution propose une méthode de repérage et d'identification de marqueurs de registres littéraires dans un corpus contrastif : des récits de voyage sont mis en contraste avec des fictions. Les textes d'un auteur qui a écrit à la fois des récits de voyage et des fictions composent le corpus. Le récit de voyage, texte factuel, est associé au registre prédominant du didactique - l'écrivain veut transmettre un savoir. L'approche traditionnelle de l'analyse statistique de données textuelles est comparée/complétée par une approche d'apprentissage de réseaux de neurones profonds dédiée à la tâche de classification automatique (prédiction) du registre littéraire. L'extraction des saillances textuelles détectées par l'IA est analysée comme traits caractéristiques de chaque registre littéraire. Plus particulièrement, les contrastes formels décelés sont postulés comme marqueurs de fictionnalité voire de littéarité envisagée comme registre pour les fictions.

Mots clés : motifs profonds, registres littéraires, analyse arborée, AFC, deep learning