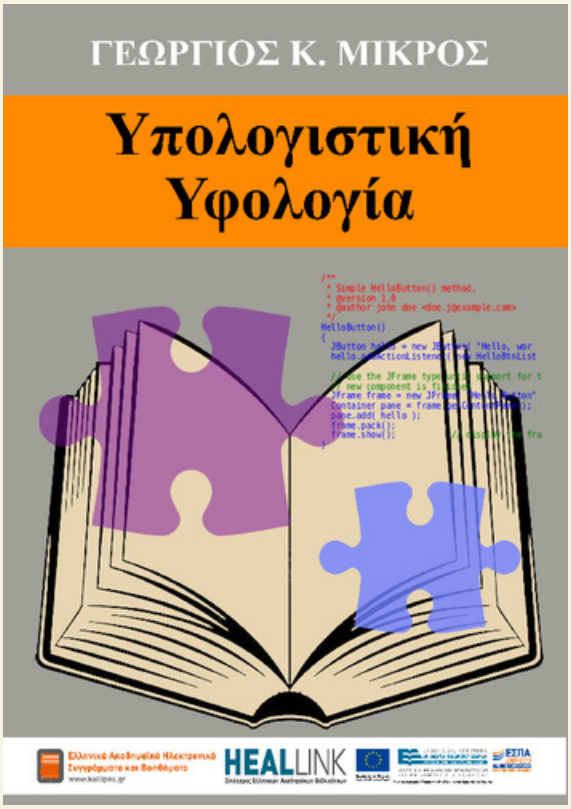

METADATA

Title: Computational Stylistics

Other Titles: -

Language: Greek

ISBN: 978-960-603-122-9

Subject: HUMANITIES AND ARTS

Keywords: Computational Linguistics / Natural Language Processing / Stylometry / Aytomatic Authorship Attribution / Machine Learning

Bibliographic Reference: Mikros, G. (2015). Computational Stylistics [Undergraduate textbook]. Kallipos, Open Academic Editions. http://dx.doi.org/10.57713/kallipos-546

Abstract

The book provides a comprehensive examination of stylometry, focusing on the automated identification of an author's writing style. This interdisciplinary field blends insights from Linguistics, Natural Language Processing, Literary Analysis, Statistics, Information Retrieval, and Machine Learning. The text begins with a historical overview of stylometry, tracing its roots back to the 18th century with studies on Shakespeare's works. Moving into the 20th century, the book highlights landmark studies, including Mosteller and Wallace's work on the authorship of The Federalist Papers, showcasing the importance of function words in determining writing style. It also covers John Burrows' contributions, particularly the use of Principal Component Analysis (PCA) in stylometry, which revolutionized the field by enabling more sophisticated statistical analyses. The book further analyzes various stylistic features, categorized into phonological/graphical, morphological/lexical, syntactic, semantic, and extralinguistic domains. Each category is examined in detail, with examples and explanations of how different features can be measured and analyzed. One of the key applications discussed is the stylometric attribution of authorship, particularly in cases involving multiple potential authors. The book outlines machine learning algorithms such as Logistic Regression and Support Vector Machines (SVMs) for classifying texts based on stylistic features. It also explores the challenges and limitations of these methods, emphasizing the need for careful selection and quantification of features to ensure accurate results. The author also addresses the use of stylometry in identifying demographic and psychological characteristics of authors, as well as in detecting stylistic consistency in collaborative digital media, such as identifying plagiarism or malicious edits in platforms like Wikipedia. The book concludes by discussing the future prospects of computational stylistics, advocating for the use of open-source tools for text analysis, and emphasizing the importance of interdisciplinary collaboration in advancing the field.

The Project is funded by the National Development Programme 2021-2025 of the Ministry of Education and Religious Affairs and implemented by the Special Account for Research Funds of the National Technical University of Athens and the Hellenic Academic Libraries Link.