**Authorship Attribution Using Natural Language Processing**

**Abstract:**

Authorship attribution is the task of identifying the author of a given piece of writing from a pool of known candidates, based on stylistic and linguistic cues. This project investigates the use of Natural sattribution, focusing on both classical machine learning and deep learning approaches. The central idea behind authorship attribution is that each individual has a distinctive writing style, which can be captured through a variety of textual features. These features include lexical elements (e.g., word frequency, vocabulary richness), syntactic structures (e.g., sentence length, use of punctuation), and structural patterns (e.g., paragraph formatting and n-gram usage). In this project, a pipeline is developed that begins with text preprocessing including tokenization, stop-word removal, stemming, and lemmatization followed by feature extraction techniques such as Term Frequency-Inverse Document Frequency (TF-IDF) and stylometric analysis. Model performance is evaluated using standard metrics including accuracy, precision, recall, and F1-score. Comparative results reveal that deep learning models tend to outperform classical approaches, especially when dealing with larger datasets and more nuanced writing styles. Overall, the project demonstrates that NLP techniques offer a powerful framework for authorship attribution, with wide-ranging applications in fields such as security, literature, education, and journalism. By combining linguistic theory with modern computational methods, this work contributes to the growing body of research aimed at understanding and modeling human language through machine learning.