AbjadAuthorID: Authorship Identification for Arabic-Script Languages at AbjadNLP 2026

Shadi Abudalfa[1], Saad Ezzini[1], Ahmed Abdelali[2], Mustafa Jarrar[3,4], Mo El-Haj [5], Nadir Durrani [3], Hassan Sajjad[6], Farah Adeeba[7], Sina Ahmadi[8]
[1]King Fahd University of Petroleum & Minerals, [2]Humain,
[3]Hamad Bin Khalifa University, [4]Birzeit University, [5]VinUniversity,
[6]Dalhousie University, [7]University of Engineering & Technology, [8]University of Zurich

Abstract

Authorship identification is a core problem in Natural Language Processing and computational linguistics, with applications spanning digital humanities, literary analysis, and forensic linguistics. While substantial progress has been made for English and other high-resource languages, authorship attribution for languages written in the Arabic (Abjad) script remains underexplored (Al-Khalifa et al., 2025). In this paper, we present an overview of **AbjadAuthorID**, a shared task organised as part of the AbjadNLP workshop at EACL 2026, which focuses on multiclass authorship identification across Arabic-script languages.

The shared task covers Modern Standard Arabic, Urdu, and Kurdish, and is formulated as a closed-set multiclass classification problem over literary text spanning multiple authors and historical periods. We describe the task motivation, dataset construction, evaluation protocol, and participation statistics, and report official results for the Arabic track. The findings highlight both the effectiveness of current approaches in controlled settings and the challenges posed by lower participation and resource availability in some language tracks. AbjadAuthorID establishes a new benchmark for multilingual authorship attribution in morphologically rich, underrepresented languages.

1 Introduction

Authorship identification seeks to determine the author of a given text based on linguistic and stylistic cues. It is a long-standing problem in NLP, with established applications in literary studies, plagiarism detection, and forensic analysis (Abudalfa et al., 2025b; Mosteller and Wallace, 1963; Lagutina et al., 2019). Traditional approaches have relied on stylometric features and classical classifiers, while more recent work has leveraged neural representations and transformer-based models to capture higher-level stylistic patterns and improve attribution performance (Devlin et al., 2019; Huang et al., 2025).

Despite this progress, most existing benchmarks and evaluations focus on English or other languages written in the Latin script. Languages that use the Arabic script, such as Arabic, Urdu, and Kurdish, pose distinct challenges due to rich morphology, orthographic ambiguity, and substantial variation across language families that nevertheless share a common writing system (Durrani, 2007). These properties complicate feature extraction and model generalisation, particularly in multiclass settings involving a large number of candidate authors (Alqahtani and Dohler, 2023; El-Haj et al., 2018). Dialectal variation, inconsistent spelling, and omitted diacritics further increase ambiguity, limiting the transferability of methods developed for Latin-script languages and motivating dedicated benchmarks for Arabic-script languages.

The AbjadAuthorID shared task builds on earlier work introduced in the AraGenEval shared task at the Third Arabic Natural Language Processing Conference (ArabicNLP) in 2025 (Abudalfa et al., 2025a). AraGenEval provided the first large-scale benchmark for Arabic authorship analysis, including authorship identification, authorship style transfer, and AI-generated text detection. While its results demonstrated strong performance for Arabic authorship identification in controlled settings, they also highlighted the need for broader multilingual evaluation and deeper analysis across different Arabic-script languages.

AbjadAuthorID extends this line of work by framing authorship identification as a multilingual, multiclass problem across Arabic, Urdu, and Kurdish. By focusing on literary text drawn from multiple authors and historical periods, the task aims to advance research on robust authorship attribution methods for morphologically rich, underrepresented languages that share the Abjad writing system (El-Haj and Ezzini, 2024).

2 Related Work

Authorship Identification concerns the problem of attributing a given text to its correct author from a predefined set of candidates (Mosteller and Wallace, 1963). The area has its origins in **stylometry**, which assumes that authors exhibit distinctive and measurable writing habits that can be exploited for attribution (Mosteller and Wallace, 1963; Lagutina et al., 2019). Early research relied heavily on manually crafted lexical, syntactic, and structural features, such as word usage patterns, sentence length distributions, and punctuation statistics, combined with classical machine learning classifiers including Naive Bayes, logistic regression, and support vector machines (Aborisade and Anwar, 2018; Bacciu et al., 2019).

The introduction of deep learning substantially reshaped the field by reducing dependence on explicit feature engineering and enabling models to learn stylistic representations directly from data (Bauersfeld et al., 2023; Huang et al., 2025). A range of neural architectures has since been explored, including recurrent neural networks (Bagnall, 2015), long short-term memory models (Qian et al., 2017), and convolutional neural networks operating at the character and word levels (Ruder et al., 2016; Shrestha et al., 2017). More complex designs, such as Siamese architectures and attention-based models, have been proposed to capture inter-text similarity and author-specific patterns more explicitly (Boenninghoff et al., 2019; Saedi and Dras, 2021).

With the emergence of large-scale pre-trained language models, transformer-based approaches have become the dominant paradigm for authorship identification. Models based on BERT and its extensions (Devlin et al., 2019; Fabien et al., 2020; Huertas-Tato et al., 2022) consistently outperform earlier neural methods, particularly when combined with techniques such as supervised contrastive learning (Khosla et al., 2020). Despite these gains, challenges remain, notably in terms of cross-domain robustness and the interpretability of learned stylistic features (Rivera-Soto et al., 2021). More recently, large language models (LLMs) have been investigated as tools for representation learning, data annotation, and even direct end-to-end attribution, showing encouraging results in domain adaptation and explainability (Brown et al., 2020; Huang et al., 2024, 2025; Fanar-Team et al., 2025).

In the context of Arabic NLP, authorship identification has been studied across a wide range of genres, including classical texts, poetry, religious writing, and contemporary online content (El-Haj, 2020, 2025a). Early evaluation efforts, such as PAN/CLEF shared tasks on author profiling (Rosso, 2017) and AraPlagDet on plagiarism detection (Bensalem et al., 2015), provided useful resources but did not explicitly target multiclass authorship attribution for Arabic. A comprehensive survey of Arabic authorship studies reports substantial variation in performance, largely attributable to differences in genre, feature representation, and dataset scale, and highlights the additional complexity introduced by Arabic morphology and diglossia (Alqahtani and Dohler, 2023; El-Haj et al., 2018).

Recent work has demonstrated the benefits of Arabic-specific pre-trained models, including AraBERT (Antoun et al., 2020a), AraELECTRA (Antoun et al., 2020b), and CAMeLBERT, which outperform multilingual alternatives on a range of authorship-related tasks, such as attribution of classical poetry and legal texts (AlZahrani and Al-Yahya, 2023; Alqurashi et al., 2025). However, generalisation across domains remains difficult, with models trained on informal or contemporary data often failing to transfer effectively to literary or historical text. The absence of large, unified benchmarks further complicates systematic comparison (Abdelali et al., 2024). AraGenEval was introduced to address this limitation by offering a controlled, multi-author benchmark for Arabic authorship analysis, a gap that the AbjadAuthorID shared task extends to a broader set of Arabic-script languages.

3 Task Description

Hosted as part of the AbjadNLP workshop at EACL 2026 (El-Haj, 2025b, 2026), AbjadAuthorID is formulated as a closed-set multiclass classification task. Given a text excerpt written in the style of a particular author, systems are required to predict the correct author from a predefined set of candidates.

The shared task is organised into three language-specific tracks, each evaluated independently using the Codabench platform.

3.1 Arabic Authorship Identification

This track targets Modern Standard Arabic. The dataset consists of literary text from 21 authors, with ten publicly accessible books per author. Each

book is segmented into semantically coherent paragraphs. Selected paragraphs are rephrased into a standardised formal style using an automated paraphrasing process, resulting in stylistically consistent inputs while preserving author-specific characteristics. The data is split into training, validation, and test sets.

3.2 Urdu Authorship Identification

The Urdu track follows the same task formulation and dataset construction methodology as the Arabic track. It enables investigation of authorship attribution in a lower-resource setting, where stylistic variation and limited training data pose additional challenges. The dataset is likewise divided into training, validation, and test splits.

3.3 Kurdish Authorship Identification

The Kurdish track, focusing on authors of Central Kurdish (Sorani), extends the task to another Arabic-script language with distinct linguistic properties. As with the other tracks, the dataset comprises literary text from multiple authors and is organised into training, validation, and test partitions. This track is intended to encourage exploration of authorship attribution in even lower-resource contexts.

3.4 Input and Output

For all tracks, the input to the system is a text segment, typically a paragraph, written in the style of a specific author. The output is the predicted author name, returned exactly as it appears in the dataset. Systems are evaluated against gold-standard author labels provided in the data.

3.5 Evaluation Metrics

Performance is evaluated primarily using the macro-averaged F1 score, which accounts for class imbalance across authors. Accuracy, precision, and recall are reported as secondary metrics. Additional qualitative analysis is encouraged to assess robustness across text lengths and stylistic variation.

4 Data

4.1 Corpus Collection for Arabic

We compiled a corpus drawn from the writings of 21 different authors, all of which are available in the public domain. For each author, ten books were selected. These works were segmented into logically consistent paragraphs. This procedure resulted in aligned source–target paragraph pairs covering the following authors: A. Amin, A. T. Pasha, A. Shawqi, A. Rihani, T. Abaza, G. K. Gibran, J. Zaydan, H. Hanafi, R. Barr, S. Moussa, T. Hussein, A. M. Al-Aqqad, A. G. Makawi, G. Le Bon, F. Zakaria, K. Kilani, M. H. Heikal, N. Mahfouz, N. El Saadawi, W. Shakespeare, and Y. Idris.

4.2 Corpus Collection for Urdu

We curated an Urdu literary corpus by crawling publicly available textual content from the Rekhta digital library. The corpus consists of prose articles and short literary texts authored by a diverse group of prominent Urdu writers, covering multiple literary movements and stylistic traditions (Hussain et al., 2005).

Texts were collected for the following authors: Qurat-ul-Ain Haider, Saadat Hassan Manto, Rajinder Singh Bedi, Ghulam Abbas, Ismat Chughtai, Prem Chand, Krishan Chander, Mumtaz Mufti, Muhammad Hameed Shahid, and Ahmad Nadeem Qasmi.

For each author, multiple articles were extracted to ensure adequate thematic coverage and linguistic diversity. The collected texts were preprocessed to remove metadata, formatting artifacts, and non-content elements, resulting in a clean corpus containing only Urdu text. After cleaning and other filtering steps, a subset of approximately $10K$ sentences per author were released from which train, dev and test splits were created.

4.3 Corpus Collection for Kurdish

We assembled a Kurdish literary dataset by harvesting texts from openly accessible online sources. The collection brings together works produced by a wide range of well-known Kurdish authors. Materials were gathered from the writings of 16 individuals: Hejar, Hêmin, Cemîl Sa'îb, Ehmed Muxtar Caf, Melay Gewre, Zêwer, Mela Mihemedî Çirustanî, 'Elaeddîn Seccadî, Cemall Nebez, Siware Îlxanîzade, Hesen Qizillcî, Îbrahîm Ehmed, Kerîm Begî Caf, Ehlam Mensûr, Mela Kerîm Sarde Kosanî and Elî Hesenyanî. After preprocessing using KLPT (Ahmadi, 2020) and filtering steps, number of total sentences are about 10K divided into trining, val, test as 70 %, 10%, and 20%.

5 Results

This section reports the official results released on the Codabench platform for the AbjadAuthorID

shared task.

5.1 Arabic Track Results

Table 1 summarises the top-performing systems for the Arabic authorship identification track. The results indicate strong performance by the leading system, with a noticeable performance gap between the top-ranked and lower-ranked submissions.

Participant ID	Macro-F1	Accuracy
zaghloul2012	0.93211	0.96339
grkurdi	0.88972	0.9244
33_tree	0.86958	0.90503
HCMUS_PrisonDilemma	0.84493	0.87674
mayar_boghdady	0.84002	0.88042
shahadsuh	0.83635	0.86913
Ali Al-Laith	0.79183	0.84785
hurryte	0.79011	0.83002

Table 1: Results for the Arabic authorship identification track.

The Arabic track attracted 15 registered participants, with a total of 68 submissions evaluated during the development and final phases.

5.2 Urdu Track Results

The authorship attribution task for Urdu attracted two teams in total, but only one of them submitted a system that met the evaluation requirements. As a result, the final leaderboard includes a single entry: the participant "shahadsuh," which achieved an F1 score of 0.39512 and an accuracy of 0.35464.

5.3 Kurdish Track Results

In the Kurdish track, two teams initially enrolled, but only a single acceptable entry was submitted at the final evaluation stage. The system achieved an F1-score of 0.59643 and an accuracy of 0.750623, and the submission was produced by the participant "rania-azad"

6 System Overview

Across the submitted system papers, authorship attribution is uniformly treated as a closed-set, multi-class classification task. However, the ways in which systems encode and exploit stylistic information differ considerably.

A number of approaches rely on transformer-based architectures including AraBERT, XLM-RoBERTa, and LLMs which are typically fine-tuned to model stylistic patterns rather than relying solely on surface-level lexical features. To cope with real-world challenges such as excessive document length, teams frequently adopt sliding-window segmentation combined with various pooling mechanisms. In several cases, transformers are enhanced through architectural constraints, such as layered classification heads or dual-dropout schemes, or are combined with conventional machine learning components within ensemble frameworks.

Alongside these neural approaches, character-level n-gram features paired with linear SVM classifiers continue to serve as competitive and widely used baselines. Such models are particularly effective at capturing subtle orthographic and morphological cues, which are especially informative in languages with complex morphology. Additional techniques—including confidence calibration, selective pseudo-label generation, and result reranking—are applied in some systems, most notably when LLMs are used in few-shot configurations, to mitigate issues related to class imbalance and domain heterogeneity.

7 Discussion

A recurring insight with this work is the challenge to the common belief that increasing model size or semantic capacity automatically leads to better stylometric performance. Several contributions demonstrate that leaner or less complex approaches can surpass LLMs, especially in Arabic authorship attribution. In particular, AraBERT-base shows more reliable generalization to unseen data than its larger counterpart, AraBERT-large, while character-level n-gram SVMs achieve markedly stronger results than deep neural architectures in the top-performing Arabic system.

Taken together, these outcomes imply that stylistic signals in Arabic-script languages are frequently grounded in surface-level features—such as morphology and orthography—rather than in abstract semantic representations. This interpretation is supported by class-wise evaluations: systems perform well on translated or contemporary prose, yet struggle with genres governed by strict or shared formal conventions, notably classical poetry, where stylistic variation is constrained.

Results across languages further underline that no single modeling strategy is universally optimal. Fine-tuned transformer models work well for Arabic and Kurdish, whereas more conventional lexical approaches remain more stable for Urdu. Although

LLMs tend to underperform in zero-shot scenarios, they exhibit potential when incorporated as reranking components in few-shot frameworks, suggesting their comparative reasoning abilities may be more valuable than direct prediction in stylometric classification.

8 Limitations

Although the reported experiments yield encouraging outcomes, important constraints remain. To begin with, many approaches depend on narrowly tailored mechanisms—such as window-based segmentation, aggregation schemes, or after-the-fact calibration—that tend to be brittle when transferred to new datasets or domains. Moreover, uneven results across genres and author categories point to persistent difficulties in separating idiosyncratic authorial traits from broadly shared stylistic norms, a problem that is especially pronounced in rigid or highly conventionalized literary settings.

In addition, the scarcity of resources for several of the studied languages limits the effective use of large-scale models, both because of insufficient training material and an elevated risk of overfitting. Lastly, while ensemble-based and hybrid methods deliver the strongest results, their increased architectural complexity and computational demands raise concerns about scalability and real-world applicability.

9 Conclusion

The findings from this work make clear that advances in stylometric modeling for Arabic-script languages do not emerge from a uniform pattern of scale-driven improvement. Strong performance is instead achieved by approaches that align model complexity with language-specific features, frequently privileging representations informed by morphology rather than relying exclusively on semantic abstraction. In this respect, the results complicate the assumption that LLMs alone guarantee better outcomes and underscore the continued relevance of classical techniques when they are judiciously integrated with contemporary neural methods.

This shared task broaden the scope of authorship attribution by moving past Arabic to incorporate relatively neglected languages such as Urdu and Kurdish. In doing so, they introduce new evaluation standards and illuminate trade-offs that are highly dependent on linguistic context. Ongoing research is likely to advance through closer engagement with linguistic structure, systematic testing across genres, and a more measured use of LLMs as supportive tools rather than central drivers within authorship attribution systems.

Acknowledgements

We thank all the participating teams for their hard work and contributions. We also acknowledge the support of the AbjadNLP workshop organizers and the EACL 2026 conference for hosting this shared task. Sina Ahmadi gratefully thanks the support of the UZH Postdoc Grant (reference number 269093).

References

Ahmed Abdelali, Hamdy Mubarak, Shammur Chowdhury, Maram Hasanain, Basel Mousi, Sabri Boughorbel, Samir Abdaljalil, Yassine El Kheir, Daniel Izham, Fahim Dalvi, Majd Hawasly, Nizi Nazar, Youssef Elshahawy, Ahmed Ali, Nadir Durrani, Natasa Milic-Frayling, and Firoj Alam. 2024. LAraBench: Benchmarking Arabic AI with large language models. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 487–520, St. Julian's, Malta. Association for Computational Linguistics.

Opeyemi Aborisade and Mohd Anwar. 2018. Classification for authorship of tweets by comparing logistic regression and naive bayes classifiers. In *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, pages 269–276. IEEE.

Shadi Abudalfa, Saad Ezzini, Ahmed Abdelali, Hamza Alami, Abdessamad Benlahbib, Salmane Chafik, Mo El-Haj, Abdelkader El Mahdaouy, Mustafa Jarrar, Salima Lamsiyah, and Hamzah Luqman. 2025a. The AraGenEval shared task on Arabic authorship style transfer and AI generated text detection. In *Proceedings of The Third Arabic Natural Language Processing Conference: Shared Tasks*, pages 1–13, Suzhou, China. Association for Computational Linguistics.

Shadi Abudalfa, Motaz Saad, and Samhaa El-Beltagy. 2025b. Emerging techniques in arabic natural language processing. *Frontiers in Artificial Intelligence*, 8:1715520.

Sina Ahmadi. 2020. KLPT – Kurdish language processing toolkit. In *Proceedings of Second Workshop for NLP Open Source Software (NLP-OSS)*, pages 72–84, Online. Association for Computational Linguistics.

Shahad Al-Khalifa, Nadir Durrani, Hend Al-Khalifa, and Firoj Alam. 2025. The landscape of arabic large

language models (allms): A new era for arabic language technology. *Communications of the ACM*. Online First.

Fatimah Alqahtani and Mischa Dohler. 2023. Survey of authorship identification tasks on arabic texts. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 22(4):1–24.

Lama Alqurashi, Serge Sharoff, Janet Watson, and Jacob Blakesley. 2025. Bert-based classical arabic poetry authorship attribution. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 6105–6119.

Fetoun Mansour AlZahrani and Maha Al-Yahya. 2023. A transformer-based approach to authorship attribution in classical arabic texts. *Applied Sciences*, 13(12):7255.

Wissam Antoun, Fady Baly, and Hazem Hajj. 2020a. AraBERT: Transformer-based model for Arabic language understanding. In *Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools, with a Shared Task on Offensive Language Detection*, pages 9–15, Marseille, France. European Language Resource Association.

Wissam Antoun, Fady Baly, and Hazem Hajj. 2020b. Araelectra: Pre-training text discriminators for arabic language understanding. *arXiv preprint arXiv:2012.15516*.

Andrea Bacciu, Massimo La Morgia, Alessandro Mei, Eugenio Nerio Nemmi, Valerio Neri, Julinda Stefa, and 1 others. 2019. Cross-domain authorship attribution combining instance-based and profile-based features notebook for pan at clef 2019. In *CEUR WORKSHOP PROCEEDINGS*, volume 2380. CEUR-WS.

Douglas Bagnall. 2015. Author identification using multi-headed recurrent neural networks. *arXiv preprint arXiv:1506.04891*.

Leonard Bauersfeld, Angel Romero, Manasi Muglikar, and Davide Scaramuzza. 2023. Cracking double-blind review: authorship attribution with deep learning. *Plos one*, 18(6):e0287611.

Imene Bensalem, Imene Boukhalfa, Paolo Rosso, Lahsen Abouenour, Kareem Darwish, and Salim Chikhi. 2015. Overview of the araplagdet pan@ fire2015 shared task on arabic plagiarism detection. In *FIRE workshops*, pages 111–122.

Benedikt Boenninghoff, Steffen Hessler, Dorothea Kolossa, and Robert M Nickel. 2019. Explainable authorship verification in social media via attention-based similarity learning. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 36–45. IEEE.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, and 1 others. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Nadir Durrani. 2007. Typology of word and automatic word segmentation in Urdu text corpus. Master's thesis, National University of Computer and Emerging Sciences, Lahore, Pakistan, August.

Mahmoud El-Haj. 2020. Habibi-a multi dialect multi national arabic song lyrics corpus. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 1318–1326.

Mahmoud El-Haj, Paul Rayson, and Mariam Aboelezz. 2018. Arabic dialect identification in the context of bivalency and code-switching. In *Proceedings of the 11th International Conference on Language Resources and Evaluation, Miyazaki, Japan.*, pages 3622–3627. European Language Resources Association.

Mo El-Haj. 2025a. Arabjobs: A multinational corpus of arabic job ads. In *Proceedings of The Third Arabic Natural Language Processing Conference*, pages 16–25.

Mo El-Haj. 2025b. Proceedings of the 1st workshop on nlp for languages using arabic script (abjadnlp 2025). In *Proceedings of the 1st Workshop on NLP for Languages Using Arabic Script, held at COLING 2025*, Abu Dhabi, United Arab Emirates.

Mo El-Haj. 2026. Proceedings of the 2nd workshop on nlp for languages using arabic script (abjadnlp 2026). In *Proceedings of the 2nd Workshop on NLP for Languages Using Arabic Script, held at EACL 2026*, Rabat, Morocco.

Mo El-Haj and Saad Ezzini. 2024. The multilingual corpus of world's constitutions (mcwc). In *Proceedings of the 6th Workshop on Open-Source Arabic Corpora and Processing Tools (OSACT) with Shared Tasks on Arabic LLMs Hallucination and Dialect to MSA Machine Translation@ LREC-COLING 2024*, pages 57–66.

Maël Fabien, Esau Villatoro-Tello, Petr Motlicek, and Shantipriya Parida. 2020. Bertaa: Bert fine-tuning for authorship attribution. In *Proceedings of the 17th International Conference on Natural Language Processing (ICON)*, pages 127–137.

Fanar-Team, Ummar Abbas, Mohammad Shahmeer Ahmad, Firoj Alam, Enes Altinisik, Ehsannedin Asgari, Yazan Boshmaf, Sabri Boughorbel, Sanjay Chawla,

Shammur Chowdhury, Fahim Dalvi, Kareem Darwish, Nadir Durrani, Mohamed Elfeky, Ahmed Elmagarmid, Mohamed Eltabakh, Masoomali Fatehkia, Anastasios Fragkopoulos, Maram Hasanain, and 23 others. 2025. Fanar: An arabic-centric multimodal generative ai platform. *Preprint*, arXiv:2501.13944.

Baixiang Huang, Canyu Chen, and Kai Shu. 2024. Can large language models identify authorship? In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 445–460, Miami, Florida, USA. Association for Computational Linguistics.

Baixiang Huang, Canyu Chen, and Kai Shu. 2025. Authorship attribution in the era of llms: Problems, methodologies, and challenges. *ACM SIGKDD Explorations Newsletter*, 26(2):21–43.

Javier Huertas-Tato, Alvaro Huertas-Garcia, Alejandro Martin, and David Camacho. 2022. Part: Pre-trained authorship representation transformer. *arXiv preprint arXiv:2209.15373*.

Sarmad Hussain, Nadir Durrani, and Sana Gul. 2005. Pan localization: Survey of language computing in asia. *Center for Research in Urdu Language Processing, Lahore, Pakistan.*

Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. 2020. Supervised contrastive learning. *Advances in neural information processing systems*, 33:18661–18673.

Ksenia Lagutina, Nadezhda Lagutina, Elena Boychuk, Inna Vorontsova, Elena Shliakhtina, Olga Belyaeva, Ilya Paramonov, and PG Demidov. 2019. A survey on stylometric text features. In *2019 25th Conference of Open Innovations Association (FRUCT)*, pages 184–195. IEEE.

Frederick Mosteller and David L Wallace. 1963. Inference in an authorship problem: A comparative study of discrimination methods applied to the authorship of the disputed federalist papers. *Journal of the American Statistical Association*, 58(302):275–309.

Chen Qian, Tianchang He, and Rao Zhang. 2017. Deep learning based authorship identification. *Report, Stanford University*, pages 1–9.

Rafael A Rivera-Soto, Olivia Elizabeth Miano, Juanita Ordonez, Barry Y Chen, Aleem Khan, Marcus Bishop, and Nicholas Andrews. 2021. Learning universal authorship representations. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 913–919.

Paolo Rosso. 2017. Author profiling at PAN: from age and gender identification to language variety identification (invited talk). In *Proceedings of the Fourth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial)*, page 46, Valencia, Spain. Association for Computational Linguistics.

Sebastian Ruder, Parsa Ghaffari, and John G Breslin. 2016. Character-level and multi-channel convolutional neural networks for large-scale authorship attribution. *arXiv preprint arXiv:1609.06686*.

Chakaveh Saedi and Mark Dras. 2021. Siamese networks for large-scale author identification. *Computer Speech & Language*, 70:101241.

Prasha Shrestha, Sebastian Sierra, Fabio González, Manuel Montes, Paolo Rosso, and Thamar Solorio. 2017. Convolutional neural networks for authorship attribution of short texts. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 669–674, Valencia, Spain. Association for Computational Linguistics.