INFOTEC-LaBD at PoliticES-IberLEF2023: Explainable Non-Linear Low-Dimensional Projections

Hiram Cabrera-Pineda[1,*,†], Eric Sadit Tellez[1,2,3] and Sabino Miranda[1,3,4]

[1] *INFOTEC Aguascalientes, México*
[2] *CICESE, Ensenada, Baja California, México*
[3] *CONAHCYT, México*
[4] *UPIITA-IPN, Ciudad de México, México*

Abstract

This paper reports our algorithm for user profiling based on a non-linear low-dimensional representation of term distribution entropy applied to the PoliticES 2023 challenge. The proposed algorithm learns a 3D-dimensional model of text data that captures the important features for user profiling while can provide insightful information through cluster analysis and visualizations. The method uses a bag of words and weighting schemes based on the term's distribution entropy. We evaluated the proposed algorithm on the PoliticES 2023 dataset in its four user profiling subtasks: gender identification, profession identification, and binary and multiclass political ideology. The results show that our proposed solution achieved competitive results on all four sub-tasks. The proposed algorithm is also explainable, which means it can provide insights into why it makes certain predictions. This makes the proposed algorithm a valuable tool for user profiling, as it can be used to understand the factors influencing user behavior.

Keywords

explainable user-profiling, low-dimensional representations, political parties identification, term's distribution entropy weighting

1. Introduction

User profiling is the process of collecting and analyzing data about users to understand their interests, demographics, and behaviors. This information has many applications, like modeling users to improve user experience, target advertising, detect malicious activity, and measure people's preferences. These preferences can span widely, but in this manuscript, we focus on political preferences as participation in the PoliticES challenge of the Iberlef 2023 forum [1]. The task consists of predicting user demographics like gender, profession, and political ideology from 43,760 Twitter messages in the Spanish language, primarily generated in Spain [2].

IberLEF 2023, September 2023, Jaén, Spain

[*] Corresponding author.

hiram.cabrerap@infotec.mx (H. Cabrera-Pineda); eric.tellez@ieee.org (E. S. Tellez); smiranda@ieee.org (S. Miranda)
0000-0003-1419-761X (H. Cabrera-Pineda); 0000-0001-5804-9868 (E. S. Tellez); 0000-0002-9421-8566 (S. Miranda)

CEUR Workshop Proceedings (CEUR-WS.org)

Tweets can be about anything, from news and current events to personal thoughts and experiences. Twitter is a valuable data source for user profiling, providing information about users' interests, demographics, and behaviors. These tasks have practical applications in various domains, including personalized marketing, content recommendation, and social science research [3].

Accurately determining user demographics provides valuable insights into individual characteristics, preferences, and ideological orientations. In this sense, several competitions have been run contests for author profiling tasks, such as a series of PAN@CLEF and FIRE [4, 5, 6, 7, 8]. These forums address different problems such as age, gender, language variety identification, personality recognition in several languages and genres, including blogs, reviews, social media, and Twitter, among others; MEX-A3T (type of occupation and place of residence) [9], IberLEF@SEPLN 2022 (gender, profession, and political ideology) [10], and others.

1.1. Related work

The idea behind user profiling is that people publish messages regularly, establishing their position about some point of interest. A knowledge database is then created, collecting user messages and labeling users to learn models on this data to predict the labels of interest for users who were never seen.

Several different techniques can be used for user profiling. Traditional methods, such as bag-of-words models, can extract features from tweets. These techniques rely on lexical features that should be determined through some weighting scheme such as TF, TF-IDF, Entropy-based, etc. For instance, in [11], the vector representation uses an entropy weighting scheme to classify gender and language variety; it represents the distribution of each term (token) over the available classes, i.e., large entropy values, terms uniformly distributed along all classes, have a low weight.

In recent years, deep learning techniques have been used to learn language models that can boost many natural language processing tasks with outstanding performance. The general strategy uses pre-trained language models to obtain high-dimensional dense vectors (embeddings) representing users. Once user vectors are computed, different algorithms can be used to learn a classifier that profiles users in some task. For instance, In [12], a transformers-based system combines the Spanish pre-trained BERT (BETO) and RoBERTa models. Both architectures are used for document-level characteristics extraction combining an MLP for label decoding. In [13], the authors use two-stage domain adaptation on BETO to teach the structure of language from Twitter and further generalize to the general political language of newspapers. It consists in continuing BETO's pretraining through the Masked Language Modelling task for more epochs.

1.2. Roadmap

This section describes the author profiling task and contextualizes its application to political domains, particularly the PoliticES challenge. Section 2 explains our approach. Our methodology is detailed in Section 3; this section briefly analyzes our dataset and models. Section 4 is devoted to experimental results and their discussion. We discuss our approach achievements and their implications in Section 5.

2. Our approach

We are primarily interested in creating powerful explainable representations, i.e., that achieve good performances and can be audited to understand the reasons behind the model's label predictions. Instead of selecting a particular set of features, we produce 3D maps where the spatial structure resembles the similarity notion of the original high-dimensional representation. In this spatial representation, it is possible to inspect similar clusters of texts by directly inspecting their messages or analyzing their shared vocabulary.

We introduced our general approach in [14]. In this edition, our approach becomes simpler since we reduced the set of hyperparameters. More specifically, our approach to user profiling consists of three main modules:

Vector spaces models. We implemented various preprocessing steps to prepare the data; for instance, we converted all messages to lowercase, normalized blank spaces to a single space, and removed diacritic marks. Token numbers 1-9 were preserved to capture important information on small numbers, while other numbers were replaced by 0 (to reduce dimensionality). We employed three types of tokens: unigrams, bigrams, and character q-grams of size four. Each token is modeled as a distribution along classes, and we compute the token's weight based on the distribution's entropy using the formulation of [14], that is, for each token t

$$\textsf{entweight}(t) = 1 + \frac{\sum_{c \in L} p_c^t \log p_c^t}{\log \# L}$$

where L is the set of tables and p_c^t is the probability of token t in class c. Note that the numerator's $\log$ produces negative numbers. Therefore, the weight is bounded between 0 (tokens with low-discrimination power) and 1 (high-discriminant tokens). This formulation ignores smoothing constants as needed by our original formulation. Instead, we reject tokens from the vocabulary if they do not occur in at least M clusters of texts. This change reduces the memory required by our models and speedups computations.

Non-linear dimensional reduction. The non-linear dimensional reduction module uses the Uniform Manifold Approximation and Projection (UMAP) [15] to produce a low-dimensional vector space (UMAP model) from the resulting vector space in the previous module. We made three-dimensional projections of the data for visualization and as the input of classifiers. UMAP projection requires the k nearest neighbor graph (using our high-dimension vector space and cosine similarity) to capture the dataset's structure. More detailed, a fuzzy smoothed version of the graph is created, and then, the eigenvectors of the normalized Laplacian matrix are computed to initialize low-dimensional embedding vectors that are finally optimized to preserve the k nearest neighbor graph in the low-dimensional projection. This procedure captures similar properties to the spectral clustering [16] while producing insightful visualizations. The explainability aspect involves leveraging these UMAP projections to create visualizations that aid in understanding the data distribution and the relationships between different clusters and provide an intuitive representation of the clusters, highlighting their separability, proximity, and any discernible patterns or trends.

Classification. The supervised learning stage uses the low-dimensional vectors to train a classifier to predict the user's political ideology, gender, or profession. We use SVM classifiers with linear and non-linear kernels in the sklearn library [17]. We perform a model selection procedure for tuning each classifier.

3. Methodology and model analysis

This manuscript tackles the problem of profiling political preferences as participation in the PoliticES challenge of the Iberlef 2023 forum [1]. The objective of this task is to predict user demographics and political ideology based on a collection of Spanish texts authored by various individuals. To ensure privacy and ethical considerations, an automated clustering approach was employed, grouping 80 tweets from different users who shared all the evaluated traits [18].

The provided training corpus consists of 180,000 messages collected from Twitter[2], aiming to extract the author's political ideology information from Spanish texts. This shared task entails gathering demographic traits, i.e., gender and profession, and political ideology as a psychographic trait.

The training dataset includes tweets from 2,250 clusters, each contributing 80 tweets. However, it is important to note that the dataset exhibits varying degrees of class imbalance across different class categories, as shown in Figure 1.

As described in Section 2, we created low-dimensional projections to provide a simple way of visualizing a dataset and its labels and to help discover the properties of each group so a latent cluster structure can also arise. Please recall that it uses a k nearest neighbor graph constructed with a vector space built from the entropy weighting model, but we discard low-frequent tokens (tokens appear at least in M cluster of texts). Therefore, our primary hyperparameters are k and M.

We use a set of $k = [10, 20, 30, 40, 50]$ and $M = [3, 5, 7, 10, 15, 25, 35]$, and fixed the UMAP method to create three-dimensional embeddings using spectral layout for embedding initialization and then optimized with 100 epochs and three negative samples per point (vector of tweets cluster).

From the given values of k and M, we generated 35 different UMAP projection visualizations for each task in the challenge, including gender, profession, binary ideology, and multiclass ideology. Each visualization corresponds to a unique combination of k and M parameters. These parameters are crucial in determining the neighborhood size and minimum distance threshold used in the UMAP algorithm.

For each combination of k and M, we applied the UMAP algorithm to project the high-dimensional data onto a lower-dimensional space. This resulted in distinct visual representations that aimed to capture the underlying structure and relationships within the data.

Figure 2 overviews the varying parameter combinations and their corresponding effects on the data representation for k and M. Note that even when we used three dimensional projections in our models, our figures are two-dimensional projections to simplify visualization. The figures show centers computed with the Affinity Propagation (AP) clustering algorithm [19] over the low-dimension embedding. Results can help identify interesting patterns, separations, or overlaps in the data, facilitating further analysis and interpretation. Centers do not represent classes but

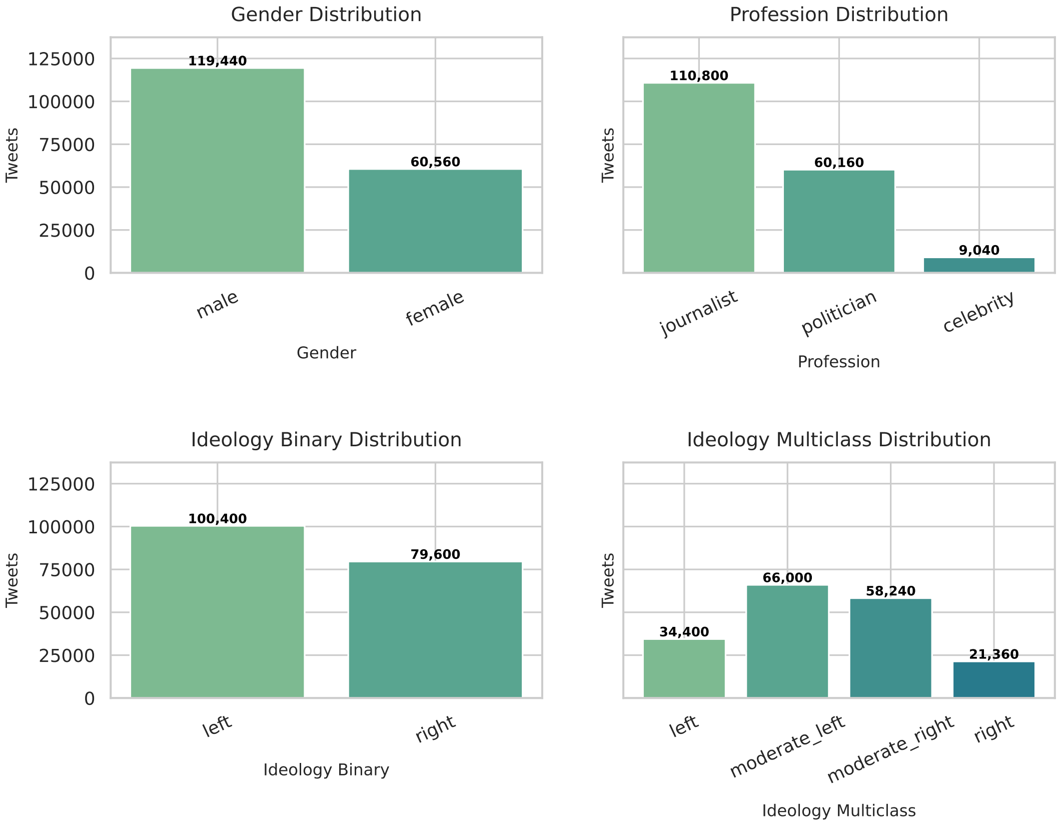


Figure 1: Class distribution of the PoliticES 2023 corpus.

groups of near clusters of texts.

We employ the silhouette score as a metric for the resulting clusters to evaluate the clustering quality within the projections. In the context of explainability, the use of silhouette scores helps in evaluating the clustering quality within the UMAP projections. By considering the silhouette scores alongside the visualizations, it becomes possible to gain a more comprehensive understanding of the clustering performance and the separability of different clusters in the lower-dimensional space.

Silhouette scores measure the cohesion and separation of groups in the projected space. Higher silhouette scores indicate well-defined clusters with distinct boundaries and good separation between clusters, indicating a clear structure within the data. On the other hand, lower silhouette scores suggest potential issues such as overlapping or poorly separated clusters, indicating a more ambiguous or complex data structure [20].

To gain further insights into the clustering performance, we generate heatmap visualizations based on the silhouette scores of the projections for each task in the challenge. These heatmaps represent the silhouette scores for each combination of k and M, with color intensity indicating the strength of the clustering.

By examining these visualizations, we can identify regions with high silhouette scores (indicat-

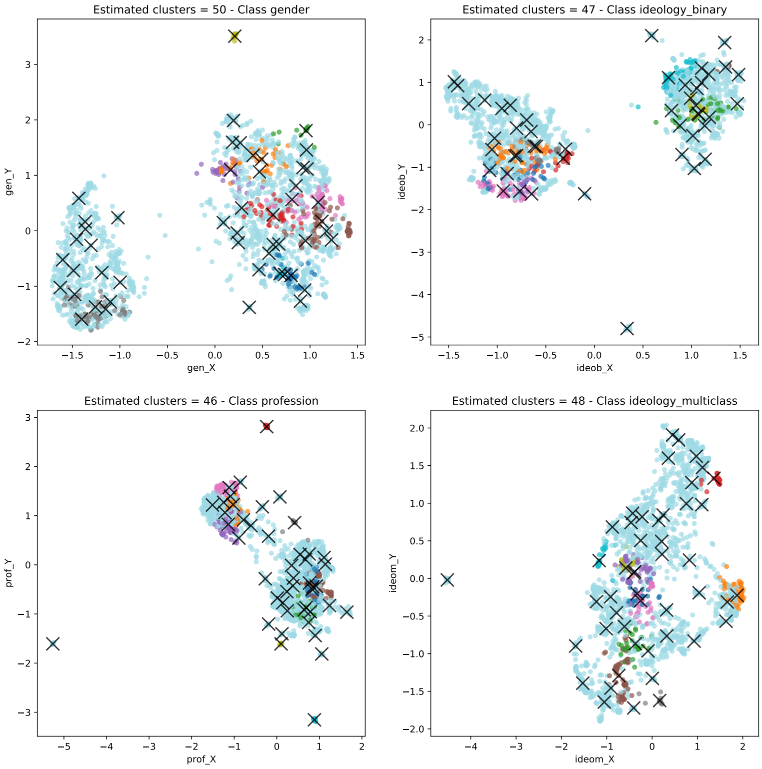


Figure 2: 2D UMAP projections for each task with $M = 5$ and $k = 20$.

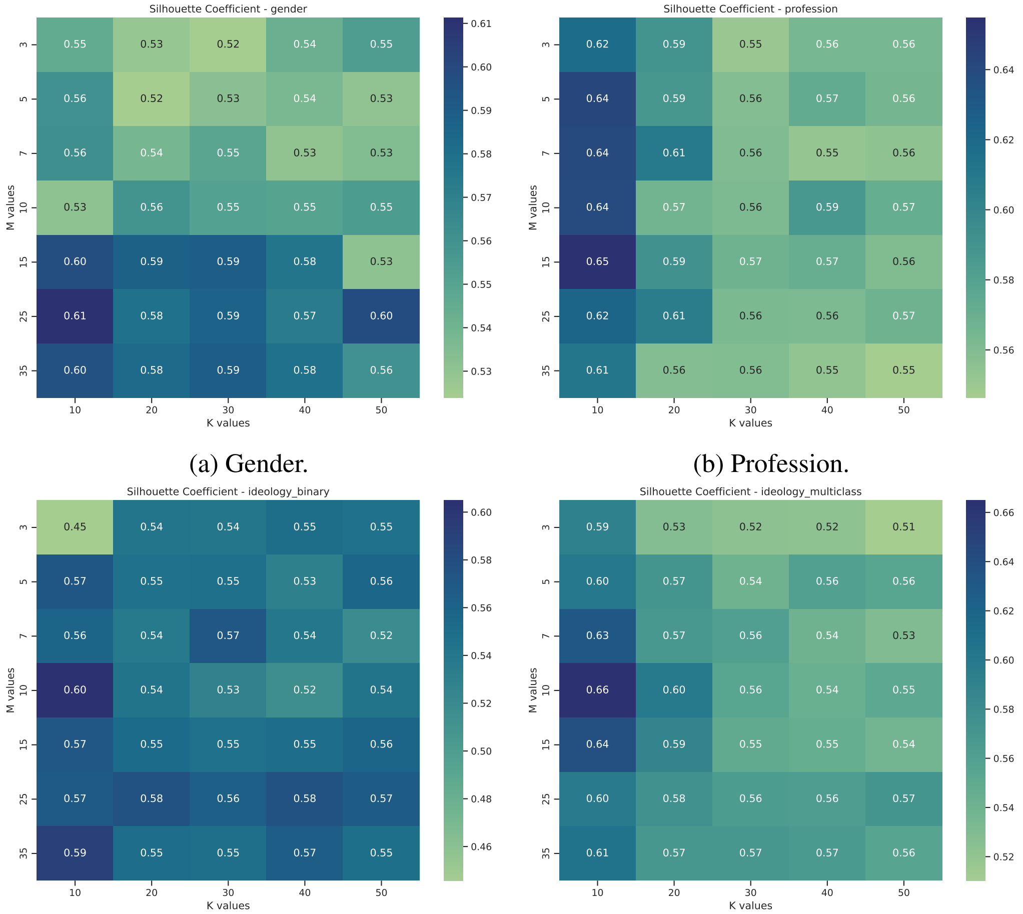
ing well-separated clusters) and regions with potential clustering ambiguities or overlaps (low scores). For visual references, please refer to Figure 3.

3.1. Model Selection

To participate in the PoliticES challenge, we developed several models using the training data provided by the organizers. We conducted a model selection using a Grid Search approach to select the most suitable models. The evaluation was performed using five-fold stratified cross-validation to maximize the macro-F1 score. We used SVM classifiers with linear and non-linear kernels for creating our classification models. We ran a hyperparameter optimization process to select those models that perform the best using five-fold cross-validation.

Each model is a distinct cluster representation (k and M) and classifiers with its hyperparameters; we also considered three-dimensional representations and the 12-dimensional concatenation of the four 3D maps of all subtasks.

This extensive evaluation allowed us to comprehensively assess the performance of various models and identify the most effective approaches for each subtask. We chose the parameter combination with the highest accuracy during the cross-validation for the final selection.



(a) Gender.

(b) Profession.

(c) Ideology binary.

(d) Ideology multiclass.

Figure 3: Silhouette heatmaps for evaluation of low-dimensional clustering.

4. Experimental results

This section presents the experimental results of our approach for the IberLEF 2023 *PoliticEs* task. Source code is available at https://github.com/hiramcp/PoliticEs2023. Our experiments were run on a four-core Laptop with 32 GB of RAM with an Intel Core i7-1165G7 @ 2.80GHz using the Windows 10 operating system. We compute the vector space using the `TextSearch.jl` Julia package available at https://github.com/sadit/TextSearch.jl which also implements all preprocessing functions, tokenizers, and the entropy-based weighting scheme. The UMAP projections were computed with the `SimSearchManifoldLearning.jl` Julia package obtained from https://github.com/sadit/SimSearchManifoldLearning.jl. The model selection and classification were performed with the Python scikit-learn package [17].

We experiment with different values of k and M to explore the trade-offs between local and global structure preservation and select the values that best suit the specific requirements of the challenge.

By analyzing the silhouette scores and heatmap visualizations, we could identify the combinations of k and M parameters that resulted in competitive and reliable projections for each task. We ensured that the subsequent training and classification tasks would be performed on data representations that exhibited clear and meaningful clusterings. This increased the chances of obtaining competitive results during the training process.

The values presented in Table 1 are the optimal combination of k and M parameters for each task in the challenge. These combinations were determined by evaluating projections using silhouette scores and affinity propagation.

Task	**k neighbors**	**M**
Gender	50	25
Profession	50	3
Ideology Binary	40	3
Ideology Multiclass	40	5

Table 1: Optimal combinations of k neighbors and M UMAP projection for each task in the challenge.

4.1. Non-supervised scores

In the Gender classification task, the combination of $k = 50$ neighbors and M of 25 yielded the second-highest silhouette score. This indicates that a larger neighborhood size and a relatively larger minimum distance effectively capture the gender clusters within the UMAP projection.

For the Profession task, the combination of $k = 50$ neighbors and a M of 3 resulted in a silhouette score of 0.56, in a range of $[0.55, 0.65]$. This indicates that the projection achieved a reasonably good separation and distinction of different professions. By using a larger neighborhood size of 50, the algorithm was able to capture more local information and capture the fine-grained differences between professions. Additionally, setting a smaller minimum distance of 3 ensured the resulting clusters were compact and well-defined, allowing for better separation between different professional groups. The silhouette score of 0.56 suggests that the projection successfully captured the underlying structure and patterns related to different professions, making it an effective choice for this task.

For the Ideology binary task, the combination of $k = 40$ neighbors and a M of 3 resulted in a silhouette score of 0.55. This score indicates that the projection achieved a reasonable level of separation and distinction between the binary ideology groups. By choosing a moderate neighborhood size of 40, the UMAP algorithm effectively captured local information and the subtle differences among different ideology categories. Additionally, setting a small minimum distance of 3 ensured that the resulting clusters were compact and well-defined, facilitating improved separation of the binary ideology groups. While the silhouette score of 0.55 suggests that the clustering within the projection may not be as clear-cut as in the Profession task, it still demonstrates a significant level of separation and a discernible structure between the binary ideology groups.

A silhouette score of 0.56 suggests that the projection for Ideology multiclass achieved satisfactory clustering and separation among the multiclass ideology groups. By selecting a moderate neighborhood size of 40, the UMAP algorithm effectively captured the local relationships and discerned the distinguishing characteristics of the various ideology categories. Additionally, setting a slightly larger M of 5 ensured that the resulting clusters were well-defined and distinct, facilitating improved separation between the multiclass ideology groups. Although the silhouette score of 0.56 falls within a range of scores, it indicates a meaningful structure within the UMAP projection, highlighting discernible patterns and groupings among the multiclass ideology categories.

4.2. Prediction results

We chose the parameter combination in each task that had the highest macro F1 score during the cross-validation for the final selection, as described in Section 3. The best hyperparameters used to tackle every task are summarized in Table 2.

Task	Classifier	Vector Dim.	Standardized	Hyperparameters
Gender	Linear SVM	3-D	No	C =1, class_weight = None, dual = False, max_iter =12000, penalty = L1, random_state = 42
Profession	Linear SVM	12-D	No	C =10, class_weight = None, dual = False, max_iter =12000, penalty = L2, random_state = 42
Ideology Binary	Linear SVM	3-D	Yes	C =1, class_weight = balanced, dual = False, max_iter =12000, penalty = L1, random_state = 42
Ideology Multiclass	RBF SVM	12-D	No	C =1, class_weight = balanced, gamma = scale, kernel = rbf

Table 2: The best hyperparameters for the machine learning models

We utilized the best Machine Learning models from our previous evaluations to extract demographic and psychographic traits from Tweets in the Test dataset. Specifically, we focused on identifying self-assigned gender and profession as demographic traits, along with political ideology as a psychographic trait.

Our approach achieved a macro-average score of 0.765 in the final evaluation, ranking us 6th on the leaderboard. The results for each task is shown in Table 3.

The macro-average score reflects the overall performance of our models in accurately predicting the assigned gender, profession, and political ideology based on the provided Test dataset. It serves as a comprehensive measure of our approach's effectiveness in capturing and predicting various demographic and psychographic traits based on the provided data.

Task	Macro F1 Score
Gender	0.743540
Profession	0.791357
Ideology Binary	0.879303
Ideology Multiclass	0.647094

Table 3: Performance for the most suitable models using the Test partition.

5. Conclusions

The combination of different parameter values for k neighbors and M in provided valuable insights into the clustering and separation of different tasks. The findings demonstrated the effectiveness of considering neighborhood size and minimum documents in capturing and distinguishing gender, profession, and ideology groups within the UMAP projections. While specific parameter combinations were identified as optimal for each task, further research can explore different parameter settings and evaluate their impact on the quality of the projections.

Using UMAP projections in conjunction with Machine Learning models proved highly effective, as evidenced by the competitive macro F1 scores obtained. These findings highlight the potential of using dimensionality reduction techniques like UMAP in improving the performance of machine learning approaches for analyzing textual data and extracting meaningful insights about individuals' characteristics and ideologies.

Additionally, conducting a thorough error analysis can provide valuable insights into our models' limitations and areas for improvement. By identifying the specific challenges or patterns in misclassifications, we can refine our models and develop targeted strategies to address those issues.

Furthermore, we can explore the application of more advanced techniques in dimensionality reduction and data representation. While our approach has shown promising results in capturing the underlying structure of the data, it is worth exploring other dimensionality reduction algorithms or combinations of techniques to potentially improve the classification performance.

Acknowledgments

H. Cabrera-Pineda is grateful to CONAHCyT for "Becas Nacionales para Estudios de Posgrado" scholarship No. 899981.

References

[1] S. M. Jiménez-Zafra, F. Rangel, M. Montes-y Gómez, Overview of IberLEF 2023: Natural Language Processing Challenges for Spanish and other Iberian Languages, in: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2023), co-located with the 39th Conference of the Spanish Society for Natural Language Processing (SEPLN 2023), CEUR-WS.org, 2023.

[2] J. A. García-Díaz, R. Colomo-Palacios, R. Valencia-García, Psychographic Traits Identification Based on Political Ideology: An Author Analysis Study on Spanish Politicians' Tweets Posted in 2020, Future Generation Computer Systems 130 (2022) 59–74.

[3] C. I. Eke, A. A. Norman, L. Shuib, H. F. Nweke, A Survey of User Profiling: State-of-the-Art, Challenges, and Solutions, IEEE Access 7 (2019) 144907–144924. doi:10.1109/ACCESS.2019.2944243.

[4] E. Stamatatos, M. Potthast, F. Rangel, P. Rosso, B. Stein, Overview of the PAN/CLEF 2015 Evaluation Lab, in: Proceedings of the 6th International Conference on Experimental IR Meets Multilinguality, Multimodality, and Interaction - Volume 9283, CLEF'15, Springer-Verlag, Berlin, Heidelberg, 2015, p. 518–538. URL: https://doi.org/10.1007/978-3-319-24027-5_49. doi:10.1007/978-3-319-24027-5_49.

[5] F. M. R. Pardo, P. Rosso, M. M. y Gómez, M. Potthast, B. Stein, Overview of the 6th Author Profiling Task at PAN 2018: Multimodal Gender Identification in Twitter, in: CLEF, 2018.

[6] P. Rosso, F. Rangel, Author Profiling Tracks at FIRE, SN Computer Science 1 (2020) 72. URL: https://doi.org/10.1007/s42979-020-0073-1. doi:10.1007/s42979-020-0073-1.

[7] J. Bevendorff, B. Chulvi, G. L. D. L. P. Sarracén, M. Kestemont, E. Manjavacas, I. Markov, M. Mayerl, M. Potthast, F. Rangel, P. Rosso, E. Stamatatos, B. Stein, M. Wiegmann, M. Wolska, , E. Zangerle, Overview of PAN 2021: Authorship Verification, Profiling Hate Speech Spreaders on Twitter,and Style Change Detection, in: 12th International Conference of the CLEF Association (CLEF 2021), Springer, 2021.

[8] F. Rangel, G. L. D. L. P. Sarracén, B. Chulvi, E. Fersini, P. Rosso, Profiling Hate Speech Spreaders on Twitter Task at PAN 2021, in: G. Faggioli, N. Ferro, A. Joly, M. Maistro, F. Piroi (Eds.), CLEF 2021 Labs and Workshops, Notebook Papers, CEUR-WS.org, 2021.

[9] M. Á. Álvarez-Carmona, E. Guzmán-Falcón, M. Montes-y Gómez, H. J. Escalante, L. Villaseñor-Pineda, V. Reyes-Meza, A. Rico-Sulayes, Overview of MEX-A3T at IberEval 2018: Authorship and Aggressiveness Analysis in Mexican Spanish Tweets, in: Notebook Papers of 3rd SEPLN Workshop on Evaluation of Human Language Technologies for Iberian Languages (IBEREVAL), Seville, Spain, September, 2018.

[10] J. A. García-Díaz, S. M. Jiménez-Zafra, M. T. Martín-Valdivia, F. García-Sánchez, L. A. Ureña-López, R. Valencia-García, Overview of PoliticES 2022: Spanish Author Profiling for Political Ideology, Procesamiento del Lenguaje Natural 69 (2022).

[11] E. S. Tellez, S. Miranda-Jiménez, M. Graff, D. Moctezuma, Gender and Language-Variety Identification with MicroTC, in: Conference and Labs of the Evaluation Forum, 2017.

[12] S. S. Carrasco, R. C. Rosillo, LosCalis at PoliticEs 2022: Political Author Profiling using BETO and MarIA, in: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2022). CEUR Workshop Proceedings, CEUR-WS, A Coruna, Spain, 2022.

[13] E. Villa-Cueva, I. González-Franco, F. Sanchez-Vega, A. P. López-Monroy, NLP-CIMAT at PoliticEs 2022: PolitiBETO, a Domain-Adapted Transformer for Multi-class Political Author Profiling, in: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2022), CEUR Workshop Proceedings, CEUR-WS, 2022.

[14] H. Cabrera, E. S. Téllez, S. Miranda, INFOTEC-LaBD at PoliticES 2022: Low-Dimensional Stacking Model for Political Ideology Profiling, in: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2022). CEUR Workshop Proceedings, CEUR-WS, A Coruna,

Spain, volume 3202 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2022.

[15] L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, 2018. URL: https://arxiv.org/abs/1802.03426. doi:10.48550/ARXIV.1802.03426.

[16] A. Ng, M. Jordan, Y. Weiss, On Spectral Clustering: Analysis and an Algorithm, Advances in neural information processing systems 14 (2001) 849–856.

[17] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research 12 (2011) 2825–2830.

[18] J. A. García-Díaz, S. M. Jiménez-Zafra, M. T. Martín-Valdivia, F. García-Sánchez, L. A. Ureña-López, R. Valencia-García, Overview of PoliticES at IberLEF 2023: Political ideology detection in Spanish texts, Procesamiento del Lenguaje Natural 71 (2023).

[19] B. J. Frey, D. Dueck, Clustering by Passing Messages Between Data Points, science 315 (2007) 972–976.

[20] P. J. Rousseeuw, Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis, Journal of Computational and Applied Mathematics 20 (1987) 53–65. URL: https://www.sciencedirect.com/science/article/pii/0377042787901257. doi:https://doi.org/10.1016/0377-0427(87)90125-7.