Detecting Depression from Social Media Text using Manifold Learning

Danillo Barros de Souza* and Javier Del Ser†

October 20, 2025

Abstract

Despite being a real worldwide concern on mental healthcare, Mental health diseases (in particular, depression), the early signs for further subsequent professional treatments are either neglected or inaccessible. Given such struggles to detect early diagnostics of mental health diseases, there is an urgency to find alternatives alongside the rigorous diagnostics provided by professional guidance. In this aspect, transformers are often used for text vectorisation and further classification from text features. However, the current methods are used blindly, without a proper statistical analysis or a deeper understanding of the text embedding and classification machinery. We investigated 4 recent text embeddings through a manifold learning perspective, and developed a simple and efficient text analysis that can detect the early signs of depression from text sentiment. To this end, we design a proof-of-concept machine learning workflow for detecting depression from social media posts from the Reddit depression dataset. Our benchmark revealed a high accuracy score for depression detection, which outperformed recent proposed workflows on the referred dataset (97.5% against 80.1%). Such results reinforce the importance of the employment of reliable machine learning methods aiming at auxiliary methods for monitoring mental health diseases.

1 Introduction

Despite the importance of mental health care, mental health diseases (MHD) are still neglected and have raised concern worldwide across decades [1–4]. In particular, depression is a considerable portion of MHD that silently represents a risk to human health. Indeed, a recent work reports the worsening of health conditions from depression, ranging from brain infection diseases as dementia, Alzheimer's disease, cardiovascular complications, obesity in women and severity of AIDS/HIV, to severe depressive disorders [5]. Depression is also statistically associated to alongside mental health disorders, e.g., schizophrenia [6], post-traumatic stress disorder (PTSD [7]), obsessive-compulsive disorder (OCD [8]), anxiety disorder [9] and bipolar disorder [10]. Furthermore, the mental health conditions are not properly addressed across many work environments, often affecting work classes as academics [11], engineers [12], health-care professionals [13],

*Mathematical Computational and Experimental Neuroscience Group (MCEN), Basque Center for Applied Mathematics (BCAM), Alameda de Mazarredo 14, 48009, Bilbao, Spain

†Department of Mathematics, University of the Basque Country (UPV/EHU), 48940, Leioa, Spain

among many others [14], where low-income countries are the most affected by the lack of mental health quality and assistance [15]. Also, the lack of accessible, reliable, and professionally guided data for mental health analysis difficult scientific advances in this regard. Alongside the digital era, social media platforms became popular and widely used for expressing daily personal experiences. On the other hand, text transformers [16] are employed a crucial role in the analysis text analysis and embeddings from semantic information and are widely used for providing insights into text analysis and classification. However, most approaches neglect important statistical features before text-to-vector transformation, which may lead to bias and inaccurate models. To tackle this, we study the embeddings provided by 4 novel text transformers in the literature by using a manifold learning dimensionality reduction method prior to the training process as an attempt to understand the data through a geometric perspective. Crucially, in this work, we analyze the social media posts and the quality of text transcription from text embeddings. Impressively, our proposed workflow supersedes the accuracy of recent works in the literature [17] (80.1% of accuracy, against 97.2% of our proposed workflow). To show our findings, this work is developed as follows: In Section 2, we present the different available methods adopted in our workflow that are efficient for text vectorization, clustering and statistical analysis. Section 3, we benchmark the proposed methods. We finally provide the discussion of our recent findings in the 4.

2 Materials and Methods

In this section, we describe the workflow adopted in our approach. In subsection 2.1, we describe the dataset used for investigating depression from text publications. Subsequently, in subsection 2.2 we introduce the text transformers tested for text vectorisation and a comparison between the features of each model is presented. In subsection 2.3, we provide the most effective dimensionality reduction methods for data visualization, and we benchmark the results obtained. We use clustering methods to identify patterns in the text data in section 2.4. Subsequently, we discuss the choices from a variety of classification methods in section 2.5. We finally test the predictability of the data by performing different classification methods in section 2.6.

2.1 The Dataset

As a starting point for analysing text from social media, we used the data collection of a supervised depression dataset from Reddit posts. The original data was provided by the Reddit Self-reported Depression Diagnosis (RSDD [18]), consisting of tabular data with approximately 9000 self-diagnosis posts (i.e., the users claimed to have been diagnosed with depression by a mental health professional), labelled as non-depressed and depressed posts (0 or 1, respectively). The cleaned data is free from emoji symbols and misspellings, resulting in a dataset of 7731 posts, which is freely available at [19].

2.2 The choice of text embeddings for mental health contextualization

Text content carries relevant information that goes beyond coherence and text semantics. Indeed, the sentiment analysis is the key point for detecting patterns that are crucial for text classification. In this aspect, the choice of a proper transformer plays an important role in providing text embeddings that contextualize the mental health scenario. In order to investigate the performance of the text vectorization, we used Sentence-BERT transformers [20] (from the python package implementation sentence_transformers [21, 22]) due to its simplicity of use and low computational power requirements. In our approach, we used the clean data described in section 2.1 as an input for the transformers. As outputs, a high-dimensional vector is provided from each text post, and each representation may vary according to the model approached. Explicitly, we adopted the models from sentence transformers provided from Hugging Face repositories, namely, **all-MiniLM-L6-v2**, **paraphrase-multilingual-MiniLM-L12-v2**, **distiluse-base-multilingual-cased-v2** and **all-distilroberta-v1**. For convenience, we will name them Model 1 to 4, respectively.

2.3 The need of Dimensionality reduction methods

Naturally, text transformers are known for providing text embeddings that encode rich semantic, syntactic, and contextual information. Consequently, there is an inevitable need for high-dimensional spaces to compact such information into a vector representation. In order to provide insights into data interpretability and statistical visualization, the use of efficient data dimensionality reduction methods is mandatory. In the context of mental health disease detection, the analysis from such methods benefits from revealing semantic clusterings that are relevant for discriminating features that hallmark depression patterns. However, this task can be challenging for Big Data, where the large number of features is a limiting factor for complex computations. Taking this fact into account, the choice of an efficient and robust dimensionality reduction method is crucial for studying intrinsic statistical features from text embeddings. In this aspect, there is a variety of methods that could be applied. Some commonly used methods are the Principal Component Analysis (PCA [23]), t-distributed Stochastic Neighbour Embedding (t-SNE [24]) and Isometric Mapping (Isomap [25]). In particular, the Uniform Manifold Approximation and Projection (UMAP [26]) had emerged due to its fast performance, low noise sensitivity and data geometry preservation. We used UMAP in our analysis for both providing statistical insights and as a pre-processing step prior to classification methods.

2.4 Clustering and outlier detection

Clustering methods help to identify underlying structures in data and to reveal hidden patterns without requiring a supervised viewpoint. Furthermore, it enables simplification of complex datasets by grouping similar data, allowing for summarisation and visualization, as well as outlier detection. Among the most famous, we have the k-means [27, 28], the Hierarchical Clustering [29,30], Spectral clustering [31,32], and Density-Based Spatial Clustering of Applications with Noise (DBSCAN [33, 34]). As we analyse the cluster quality provided by previous dimensionality reduction methods, we decided to adopt DBSCAN in our approach as a powerful and simple tool for mapping the clusters in the data.

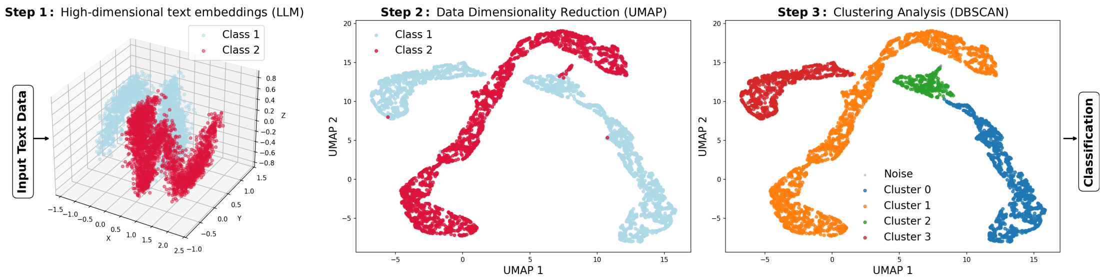


Figure 1: Workflow description, from the text input data to the classification process.

2.5 Classification methods

Classification is a core task in supervised machine learning where the goal is to assign input data to one of several predefined categories or classes based on labelled training examples. Such methods play a crucial role in data-driven decision-making to provide data interpretability. In our work, we tested 4 different methods for benchark the quality of the predictability of depression cases, namely, k-nearest neighbours (KNN) [35], neural networks (NN) [36], decision tree (DC) [37], and support vector machines (SVM) [38] classifiers.

2.6 The Proposed Modelling Workflow

To benchmark our hypothesis on improved accuracy via geometric insights from dimensionality reduction, we benchmarked the traditional text vectorization from the text embeddings as a baseline to compare the results with the pre-geometrized information from the manifold learning, generated from UMAP embeddings. We further tested the classification quality of those methods by utilising k-Nearest Neighbours, a simple neural network model, decision tree and support vector machine. To validate each method, the data was split into train, test and validation data, and the results were benchmarked with a confusion matrix [39]. Explicitly, our proposed geometric supervised workflow is described as follows:

1. The text information is provided as input for a text embedding (models 1 to 4);
2. The resulting text embedding undergoes a dimensionality reduction to provide data visualization (in this case, with UMAP);
3. The clustering quality and anomalies are investigated from clustering and outlier detection methods (in our case, with DBSCAN) for further classification;
4. A benchmark is performed from varius classification methods (KNN, NN, DT or SVM) from a parameter grid in order to search the optimal parameters to infer the depression from embedded text.

The Figure 1 elucidates our workflow.

3 Results

Here, we provide the preliminary statistical analysis performed, followed by the clustering and classification benchmark across the proposed workflow.

3.1 Statistical Analysis

Prior to the text embeddings, we measured the text length of each speech as a starting point for statistical information. As observed in Figure 2, the text lengths range from short sentences to extremely large texts. Interestingly, it was noticed that text lengths above 42 do not have depressive labels, which may suggest a probabilistic indicator of depressive post for large texts. In contrast, the detection of depression in short texts is challenging from a probabilistic viewpoint.

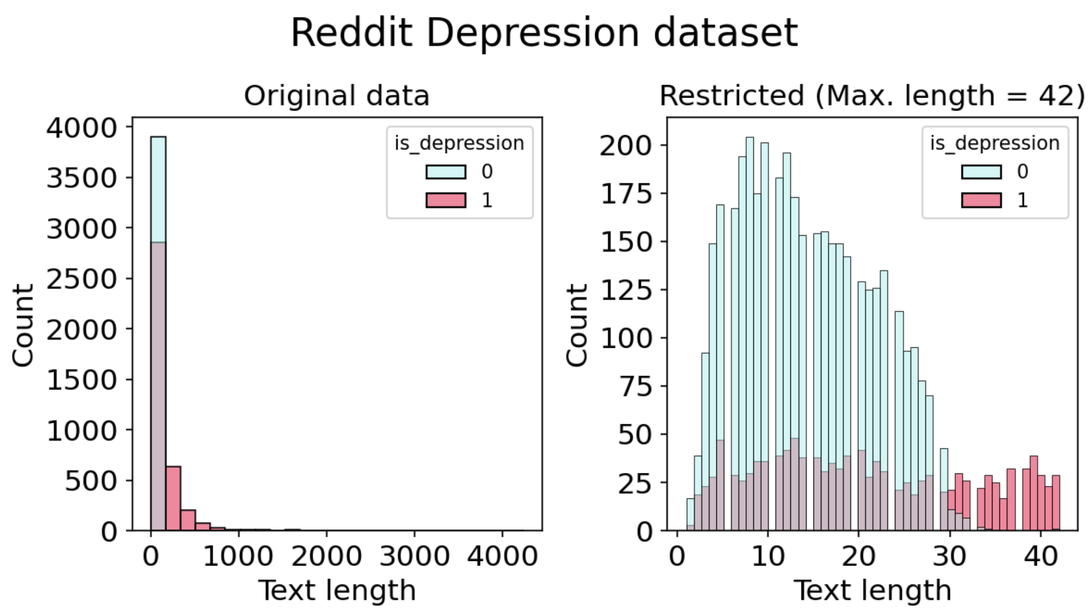


Figure 2: Reddit dataset and the statistical information by text length.

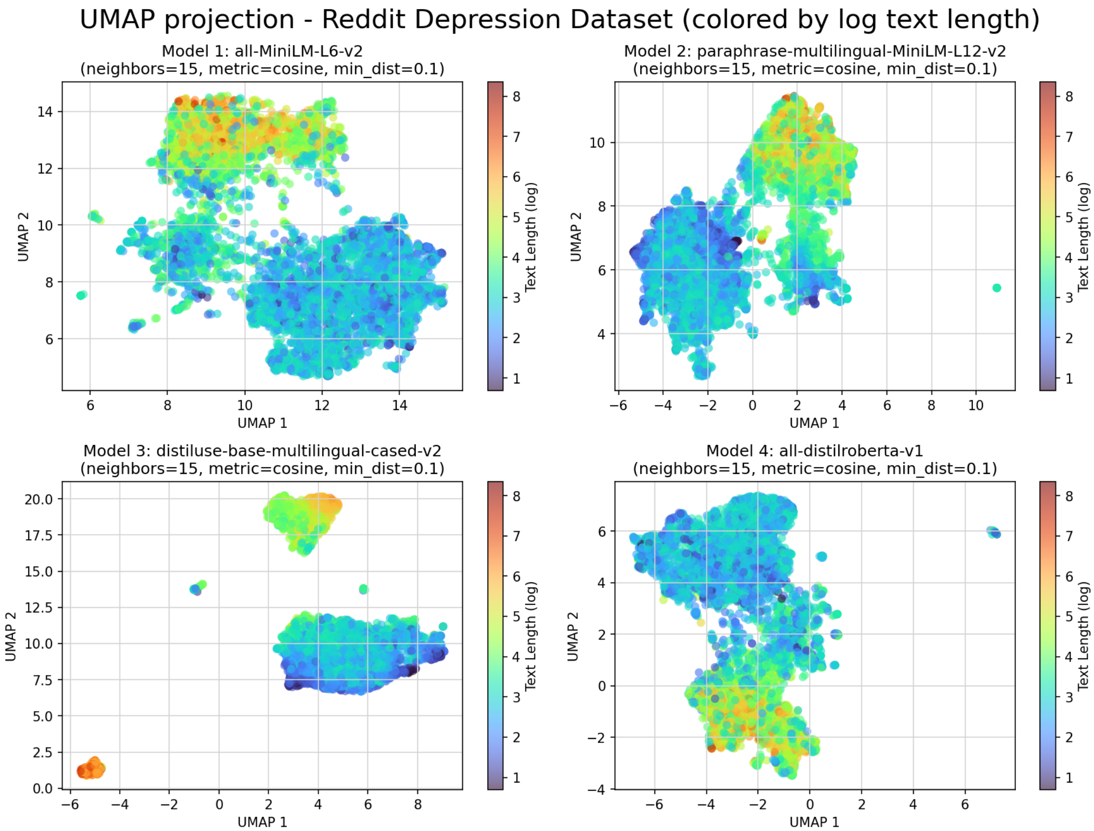


Figure 3: Comparing the text length sensitivity of the text embeddings from their respective 2D-UMAP projection.

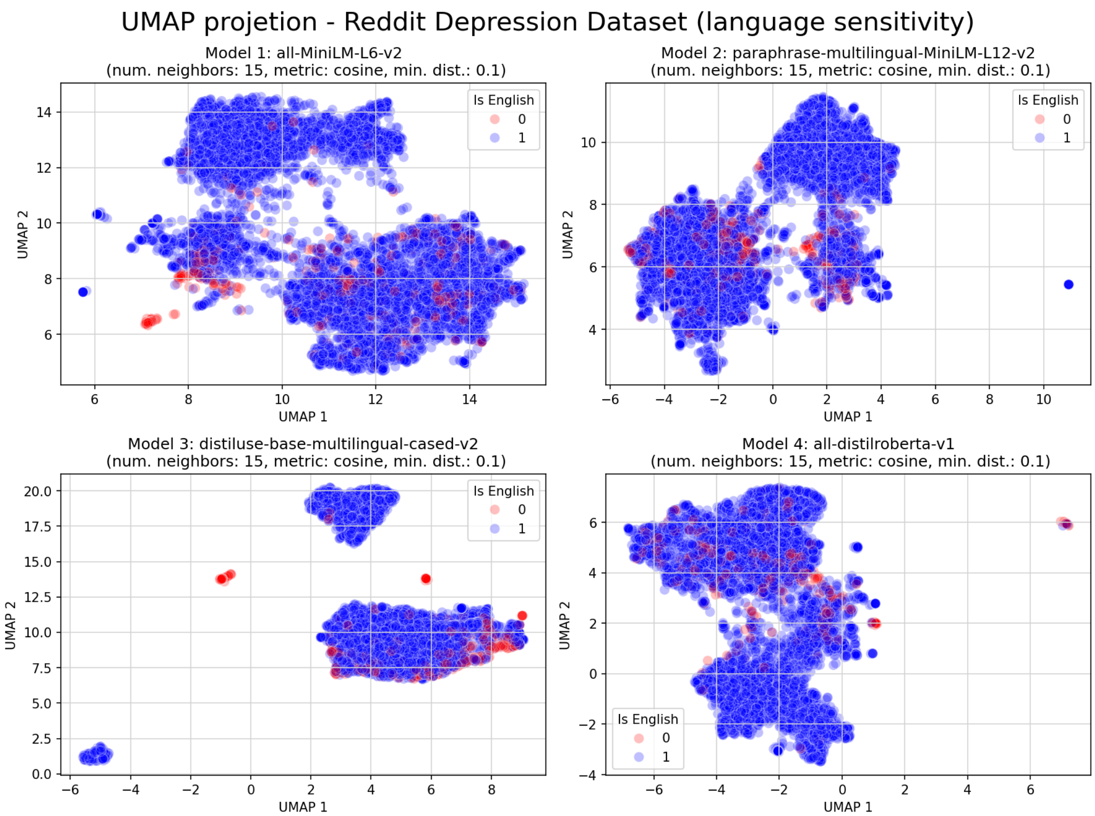


Figure 4: Comparing the language sensitivity of the text embeddings from their respective 2D-UMAP projection.

We investigated the clustering quality of the 4 transformers utilised in our analysis. To this end, we first performed a data dimensionality reduction with UMAP, with a parametric space 15 nearest neighbours, a minimum point distance of 0.1. We used the cosine metric for measuring distances due to its advantage regarding other metrics for dealing with vectorized information [40, 41]. In Figures 3 and 4 we display the model's sensitivity to text length (in log scale) and languages (English and non-English text), respectively. It was noticed that the model 4 provided text embeddings that are hypersensitive to both text length and language, which compromised its reliability for depression detection. The opposite performance is verified in the models 1, 2 and 4. Subsequently, we test the sensitivity to depression text in Figure 5. We noticed that models 1, 2 and 4 presented a clear separability between the depressive and non-depressive group, despite the sensitivity to outliers. In contrast, model 3 provided solid clustering information; however, the distinction between the depressed and non-depressed groups was compromised. In Figure 6, we provide the clustering information for each cluster detected by DBSCAN. Once more, models 1, 2, and 4 provided the best clustering quality and fair clustering per classes, despite the sensitivity to noise. In contrast, the clustering of model 3 showed poor sensitivity to text from depressed and non-depressed classes, resulting in a low reliability on the separation between those groups. We also tested the model's sensitivity to the language input (English and non-English text). The results can be accessed from Figures SM 1–5 in the Supplementary Material. In this section, we restrict the results to the approaches and parameter spaces that provided the highest accuracy scores. The remaining results (including the pre-clusterized approach) are provided in the Figures SM 7–9 in the Supplementary Material.

3.2 Benchmarks

To test the accuracy of detecting depression of the proposed workflow, we split the data into 60% and 40% for training and test data, respectively, for KNN, DT and SVM classifiers. Similarly, to train the neural network model, we split the data into 60%, 20% and 20% for train, test and validation data, respectively. In all models, we used a batch training of size 5. In all approaches, we provided the tests for the pre-geometrized data from UMAP embeddings, and the vectorized text as inputs for classification for baseline comparison. The details of the benchmark (including time processing and accuracy scores) can be reproduced from the codes in our GitHub repository [42].

`k-Nearest Neighbors classifier:` We performed a grid search on the k-nearest neighbours classifier, for $k \in \{10, 15, 25, 30, 35, 40\}$ and different distance metrics (Euclidean, Squared-Euclidean, Manhattan and Cosine), for both vectorised text classification and with UMAP dimensionality reduction as a preliminary clustering. For reproducibility, the codes are publicly available at [42]. The best scoring is shared by both approaches (UMAP embedding and text embedding). The best scores for the kNN approach were obtained from model 2, with accuracy scores accuracy score 96.6% and 96.7%, respectively. It is worth noting that the dimensionality reduction provides a significant improvement in the time processing of the models, which should be considered in the construction of models from large datasets. The details can be found in Figure 7.

`Neural Network model:` To test across methods, we also trained a simple neural network for both vectorised data and pre-geometrized data. Table 1 shows the details of the proposed neural network. For the compilation setup, we used ADAM as an optimized and the binary Cross-Entropy as a loss function. We trained the model utilising the accuracy as a metric. In Figure 8 we provide the results for the highest accuracy approach, which was obtained from the vectorised input data (model 2, f1-score 97.1%, against 96.8% from the UMAP embedding approach). We conclude that there was no accuracy and significant time improvement in training the model with UMAP embeddings, which may indicate that the UMAP projection is ineffective for the NN approach.

`Decision Tree:` For the search grid parameters, we used Gini, Entropy and Log loss for the criterion. In the Maximum depth, we used values in $\{\emptyset, 5, 10, 15\}$. For the Minimum samples split, we adopt parameters in the set $\{2, 5, 10\}$. For the minimum sample leaf parameter, we used the values in $\{1, 2, 4\}$. As a result, we obtained 96.3% as an accuracy score from the geometrized approach (i.e., the UMAP embeddings from LLMs as inputs), against 91.7 from the text embedding approach. The details of the best parameter model can be found in the Figure 9.

`Support Vector Machine:` Finally, we adopted the kernels RBF, sigmoid and polynomial for the SVM classifier. Additionally, we used C in $\{0.1, 1, 10, 100\}$, gamma in $\{0.001, 0.01, 0.1, 1\}$ in the parametric space for the grid search. Once more, the best accuracy score result (97.5%) was obtained from the non-geometric approach (against 96.5% from the UMAP embeddings). Additionally, the substantial computational time increment from the UMAP embedding approach should be considered. The details of the SVM benchmark are elucidated in Figure 10.

We concluded that the best-scoring results are provided by the classification obtained from SVM, NN, kNN and DT, in this sequence. We also summarise the benchmark results obtained from our proposed workflow in the Table 2.

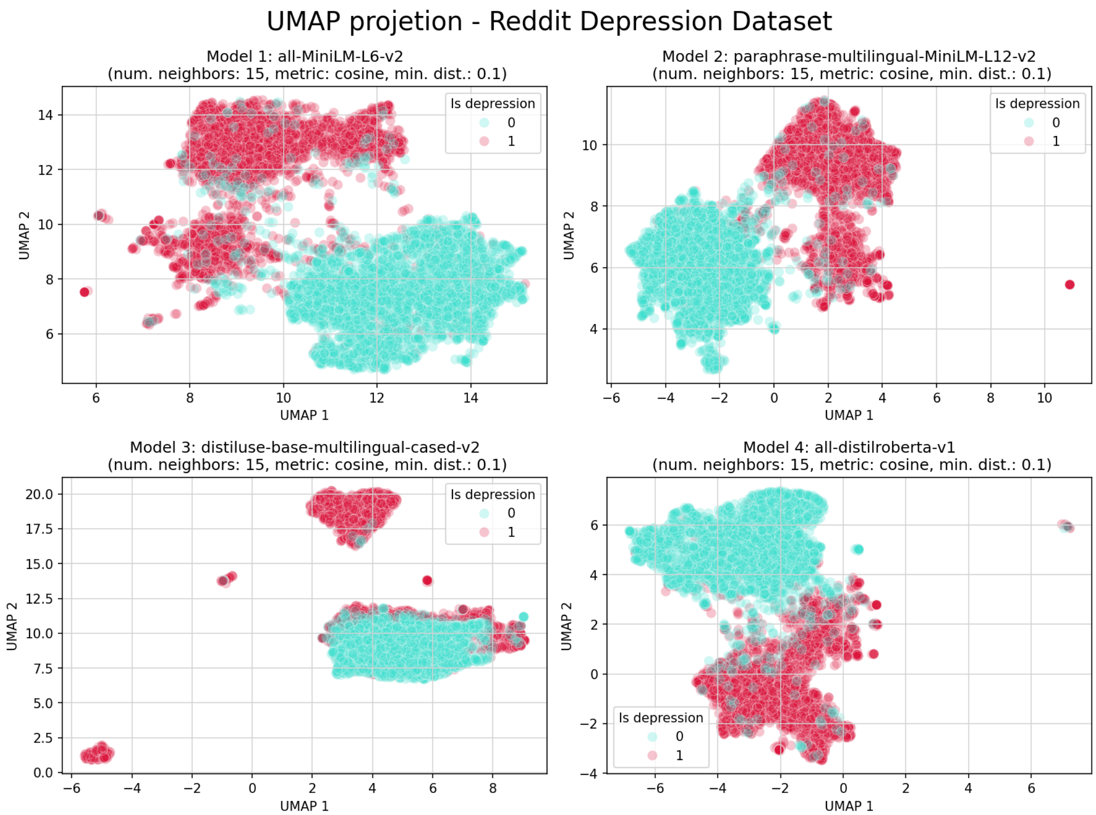


Figure 5: Comparison of the text embeddings from the 2D-UMAP projection.

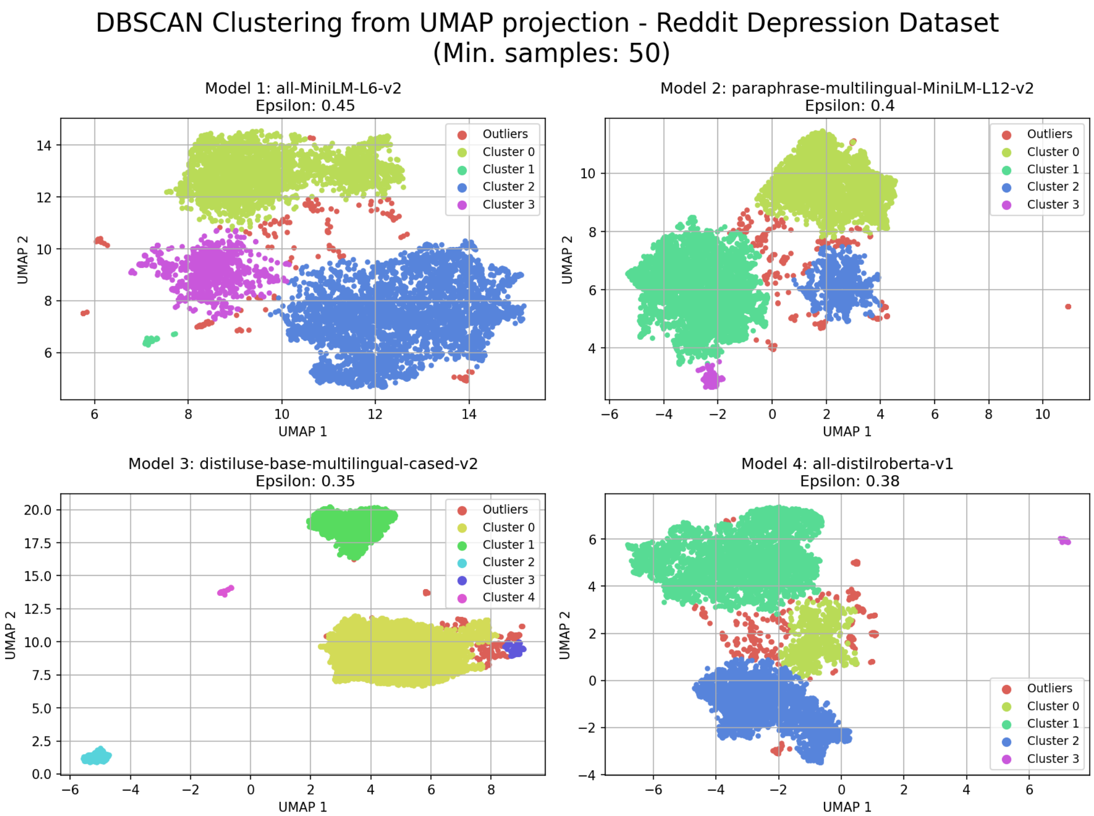


Figure 6: Clustering comparison across large language models.

Layer	Activation Function
Input	–
Dense (32 neurons)	ReLU
Dropout (0.2)	–
Dense (1 neuron)	Sigmoid

Table 1: Neural network architecture for the Reddit depression dataset.

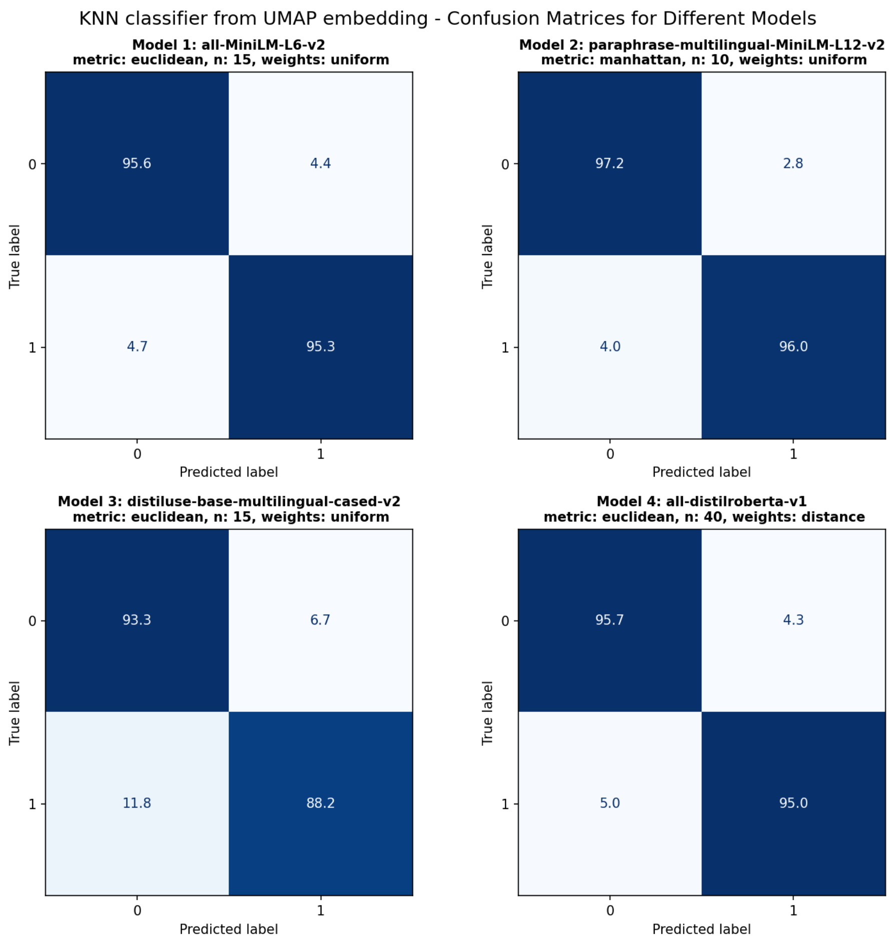
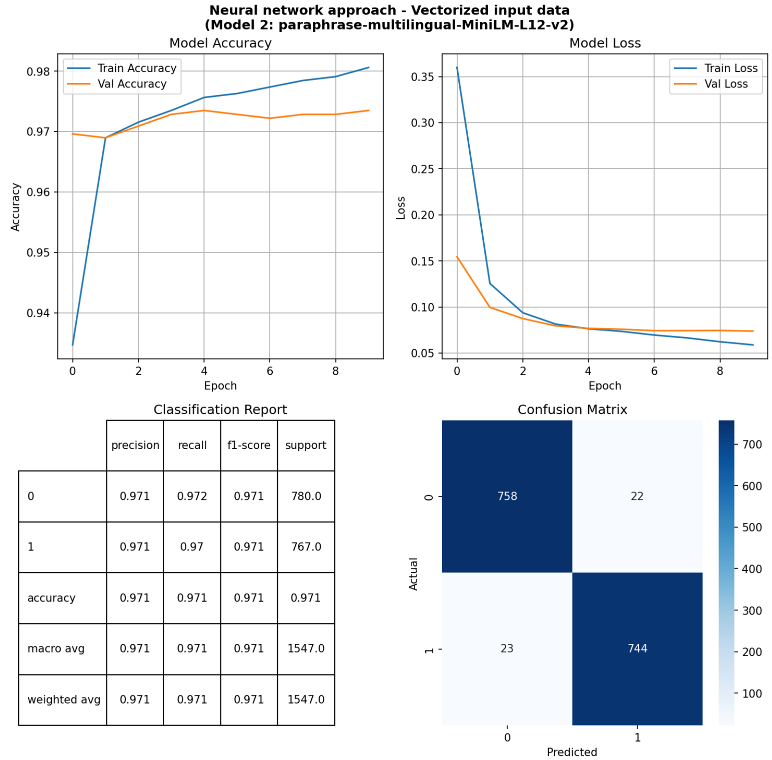


Figure 7: Confusion matrices comparison between different text embeddings. Here, we computed the 2D UMAP dimensionality reduction as a preliminary step of k-nearest neighbours classification.



	precision	recall	f1-score	support
0	0.971	0.972	0.971	780.0
1	0.971	0.97	0.971	767.0
accuracy	0.971	0.971	0.971	0.971
macro avg	0.971	0.971	0.971	1547.0
weighted avg	0.971	0.971	0.971	1547.0

Figure 8: Neural network training on vectorised text by model 2.

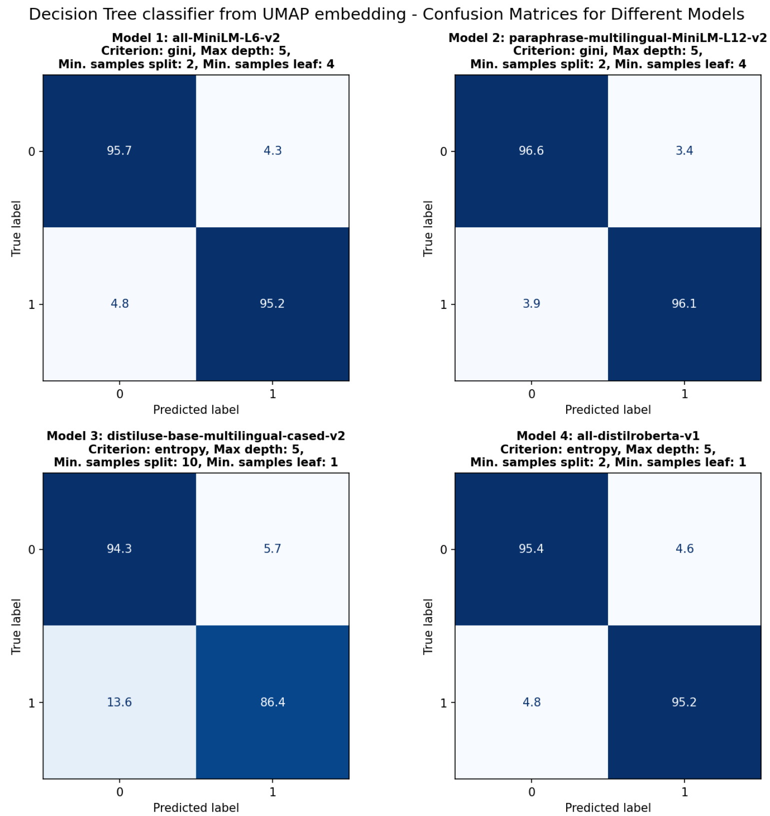


Figure 9: Confusion matrices comparison between different text embeddings using UMAP dimensionality reduction and Decision Tree classifier.

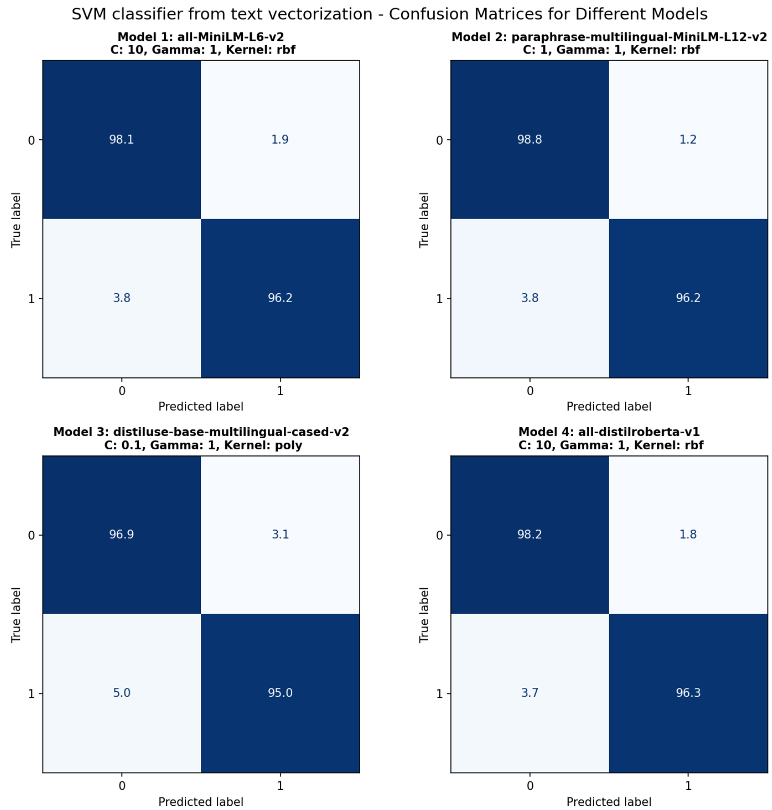


Figure 10: Confusion matrices comparison between different text embeddings and SVM classifier.

Classifier	Best LLM	UMAP	Computational Complexity	Accuracy Score (%)
SVM	2	No	High	97.5
NN	2	No	Low	97.1
KNN	2	Yes	Low	96.6
DT	2	Yes	Middle	96.3

Table 2: Benchmark summary of the adopted transformers models, methods, and classifiers.

4 Conclusion

We proposed a manifold learning workflow to investigate the relevant statistical features of depression diseases from social media text embeddings generated from 4 novel transformers. To this end, we investigated a range of classification models and parameter spaces to benchmark and provide optimal choices for reliable models. Our findings demonstrate the efficiency of the proposed methods, achieving an accuracy score of 97.5%, compared to 80.1% reported in recent work [17]. This result underscores the dual benefit of manifold learning: it reduces data complexity (thereby lowering computational costs in some cases) while simultaneously providing a geometric perspective that enhances classification accuracy. Although the incorporation of geometric information does not always yield accuracy gains and reduced computational costs, manifold learning is a key factor for investigating hallmarks of depression from text embeddings, as well as their sensitivity to text language and length. Overall, this work highlights the value of integrating statistical analysis with geometric insights into data structure, suggesting manifold learning as a promising complement to current machine learning methodologies for the detection of depressive behaviour in social media text.

Acknowledgement

This research is supported by the grant
PID2023-146683OB-100 funded by MICIU/AEI /10.13039/501100011033 and by ERDF, EU. Additionally, it is supported by Ikerbasque Foundation and the Basque Government through the BERC 2022-2025 program and by the Ministry of Science and Innovation: BCAM Severo Ochoa accreditation CEX2021-001142-S / MICIU / AEI / 10.13039/501100011033. Moreover, the authors acknowledge the financial support received from BCAM-IKUR, funded by the Basque Government by the IKUR Strategy and by the European Union NextGenerationEU/PRTR. We also acknowledge the support of ONBODY no. KK-2023/00070 funded by the Basque Government through ELKARTEK Programme.

References

[1] Vikram Patel. Why mental health matters to global health. *Transcultural psychiatry*, 51(6):777–789, 2014.

[2] Patricio V Marquez and Shekhar Saxena. Making mental health a global priority. In *Cerebrum: the Dana forum on brain science*, volume 2016, pages cer–10, 2016.

[3] China Mills. From 'invisible problem'to global priority: The inclusion of mental health in the sustainable development goals. *Development and Change*, 49(3):843–866, 2018.

[4] Modhurima Moitra, Shanise Owens, Maji Hailemariam, Katherine S Wilson, Augustina Mensa-Kwao, Gloria Gonese, Christine K Kamamia, Belinda White, Dorraine M Young, and Pamela Y Collins. Global mental health: Where we are and where we are going. *Current psychiatry reports*, 25(7):301–311, 2023.

[5] Alix M Arnaud, Teri S Brister, Ken Duckworth, Phyllis Foxworth, Tonya Fulwider, Ellison D Suthoff, Brian Werneburg, Izabela Aleksanderek, and Marcia L Reinhart. Impact of major depressive disorder on comorbidities: a systematic literature review. *The Journal of clinical psychiatry*, 83(6):43390, 2022.

[6] Wen Li, Yuan Yang, Feng-Rong An, Ling Zhang, Gabor S Ungvari, Todd Jackson, Zhen Yuan, and Yu-Tao Xiang. Prevalence of comorbid depression in schizophrenia: a meta-analysis of observational studies. *Journal of Affective Disorders*, 273:524–531, 2020.

[7] Nina K Rytwinski, Michael D Scur, Norah C Feeny, and Eric A Youngstrom. The co-occurrence of major depressive disorder among individuals with posttraumatic stress disorder: A meta-analysis. *Journal of traumatic stress*, 26(3):299–309, 2013.

[8] Anna Merrill, Beth Gershuny, Lee Baer, and Michael A Jenike. Depression in comorbid obsessive-compulsive disorder and posttraumatic stress disorder. *Journal of clinical psychology*, 67(6):624–628, 2011.

[9] Ravi Philip Rajkumar. Comorbid depression and anxiety: Integration of insights from attachment theory and cognitive neuroscience, and their implications for research and treatment. *Frontiers in Behavioral Neuroscience*, 16:1104928, 2022.

[10] Thomas Munk Laursen, Esben Agerbo, and Carsten Bocker Pedersen. Bipolar disorder, schizoaffective disorder, and schizophrenia overlap: a new comorbidity index. *Journal of Clinical Psychiatry*, 70(10):1432, 2009.

[11] Astrid Müller. Mental health disorders: prevalent but widely ignored in academia? *Journal of Physiology*, 598(7), 2020.

[12] Cristina Martinez Montes, Birgit Penzenstadler, and Robert Feldt. The factors influencing well-being in software engineers: A cross-country mixed-method study. *arXiv preprint arXiv:2504.01787*, 2025.

[13] Xie Zhang, KE Zhao, Guohua Zhang, Ruihua Feng, Jianjun Chen, Dongwu Xu, Xiaodong Liu, Arlette J Ngoubene-Atioky, Hong Huang, Yanlong Liu, et al. Occupational stress and mental health: a comparison between frontline medical staff and non-frontline medical staff during the 2019 novel coronavirus disease outbreak. *Frontiers in psychiatry*, 11:555703, 2020.

[14] Birgit Aust, Caleb Leduc, Johanna Cresswell-Smith, Clíodhna O'Brien, Reiner Rugulies, Mallorie Leduc, Doireann Ni Dhalaigh, Arilda Dushaj, Naim Fanaj, Daniel Guinart, et al. The effects of different types of organisational workplace mental health interventions on mental health and wellbeing in healthcare workers: a systematic review. *International archives of occupational and environmental health*, 97(5):485–522, 2024.

[15] Julia Lohmann, Olzhas Shulenbayev, Danielle Wilhelm, Adamson S Muula, and Manuela De Allegri. Psychological wellbeing in a resource-limited work environment: examining levels and determinants among health workers in rural malawi. *Human resources for health*, 17(1):85, 2019.

[16] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[17] Manuel Kanahuati-Ceballos and Leonardo J Valdivia. Detection of depressive comments on social media using rnn, lstm, and random forest: comparison and optimization. *Social Network Analysis and Mining*, 14(1):44, 2024.

[18] Andrew Yates, Arman Cohan, and Nazli Goharian. Depression and self-harm risk assessment in online forums. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2958–2968. Association for Computational Linguistics, 2017.

[19] Infamouscoder. Depression reddit cleaned. `https://www.kaggle.com/datasets/infamouscoder/depression-reddit-cleaned`, 2023. Accessed: 2025-07-21.

[20] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*, 2019.

[21] Nils Reimers. Sentence-transformers documentation. `https://www.sbert.net/`. Accessed: 2025-07-22.

[22] Nils Reimers and Iryna Gurevych. sentence-transformers: Multilingual sentence embeddings using bert and other transformers. `https://github.com/UKPLab/sentence-transformers`, 2020. Accessed: 2025-07-22.

[23] Hervé Abdi and Lynne J Williams. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4):433–459, 2010.

[24] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(Nov):2579–2605, 2008.

[25] Joshua B Tenenbaum, Vin de Silva, and John C Langford. A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323, 2000.

[26] Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

[27] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B (methodological)*, 39(1):1–22, 1977.

[28] KMMN Krishna and M Narasimha Murty. Genetic k-means algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 29(3):433–439, 1999.

[29] Fionn Murtagh and Pedro Contreras. Algorithms for hierarchical clustering: an overview. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 2(1):86–97, 2012.

[30] Frank Nielsen. Hierarchical clustering. In *Introduction to HPC with MPI for Data Science*, pages 195–211. Springer, 2016.

[31] Andrew Ng, Michael Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. *Advances in neural information processing systems*, 14, 2001.

[32] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416, 2007.

[33] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231, 1996.

[34] Erich Schubert, Jörg Sander, Martin Ester, Hans Peter Kriegel, and Xiaowei Xu. Dbscan revisited, revisited: why and how you should (still) use dbscan. *ACM Transactions on Database Systems (TODS)*, 42(3):1–21, 2017.

[35] Leif E Peterson. K-nearest neighbor. *Scholarpedia*, 4(2):1883, 2009.

[36] Yu-chen Wu and Jun-wen Feng. Development and application of artificial neural network. *Wireless Personal Communications*, 102(2):1645–1656, 2018.

[37] Shan Suthaharan. Decision tree learning. In *Machine learning models and algorithms for big data classification: thinking with examples for effective learning*, pages 237–269. Springer, 2016.

[38] Shan Suthaharan. Support vector machine. In *Machine learning models and algorithms for big data classification: thinking with examples for effective learning*, pages 207–235. Springer, 2016.

[39] Robert Susmaga. Confusion matrix visualization. In *Intelligent information processing and web mining: proceedings of the international IIS: IIPWM '04 conference held in zakopane, Poland, may 17–20, 2004*, pages 107–116. Springer, 2004.

[40] Tan Thongtan and Tanasanee Phienthrakul. Sentiment classification using document embeddings trained with cosine similarity. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*, pages 407–414, 2019.

[41] Wei Zhao, Michael Strube, and Steffen Eger. Discoscore: Evaluating text generation with bert and discourse coherence. *arXiv preprint arXiv:2201.11176*, 2022.

[42] Danillo Barros de Souza. Using manifold learning workflow for detecting depression from reddit posts. `https://github.com/danillodbs16/AI_UPV_2025_depression_reddit/`, 2025. GitHub repository, accessed 2025-09-12.

Supplementary Material -Detecting Depression from Social Media Text using Manifold Learning

Herein, we provide the complementary content from our statistical analysis and model constructions performed in our work.

SM0.1 Additional Figures

In Figure SM1 we display the difference between clustering information from UMAP manifold learning provided from DBSCAN in each of the 4 text transformers approached. In Figures SM2 to SM5. In Figure SM6, the classification via kNN from vectorized text (no UMAP embedding applied) is benchmarked among the 4 approached text transformers. The training process using neural networks for the same vectorized input data can be found in Figures SM7 to SM9. In Figures SM10 to SM13 the benchmark between the models are performed from UMAP embedding input data.

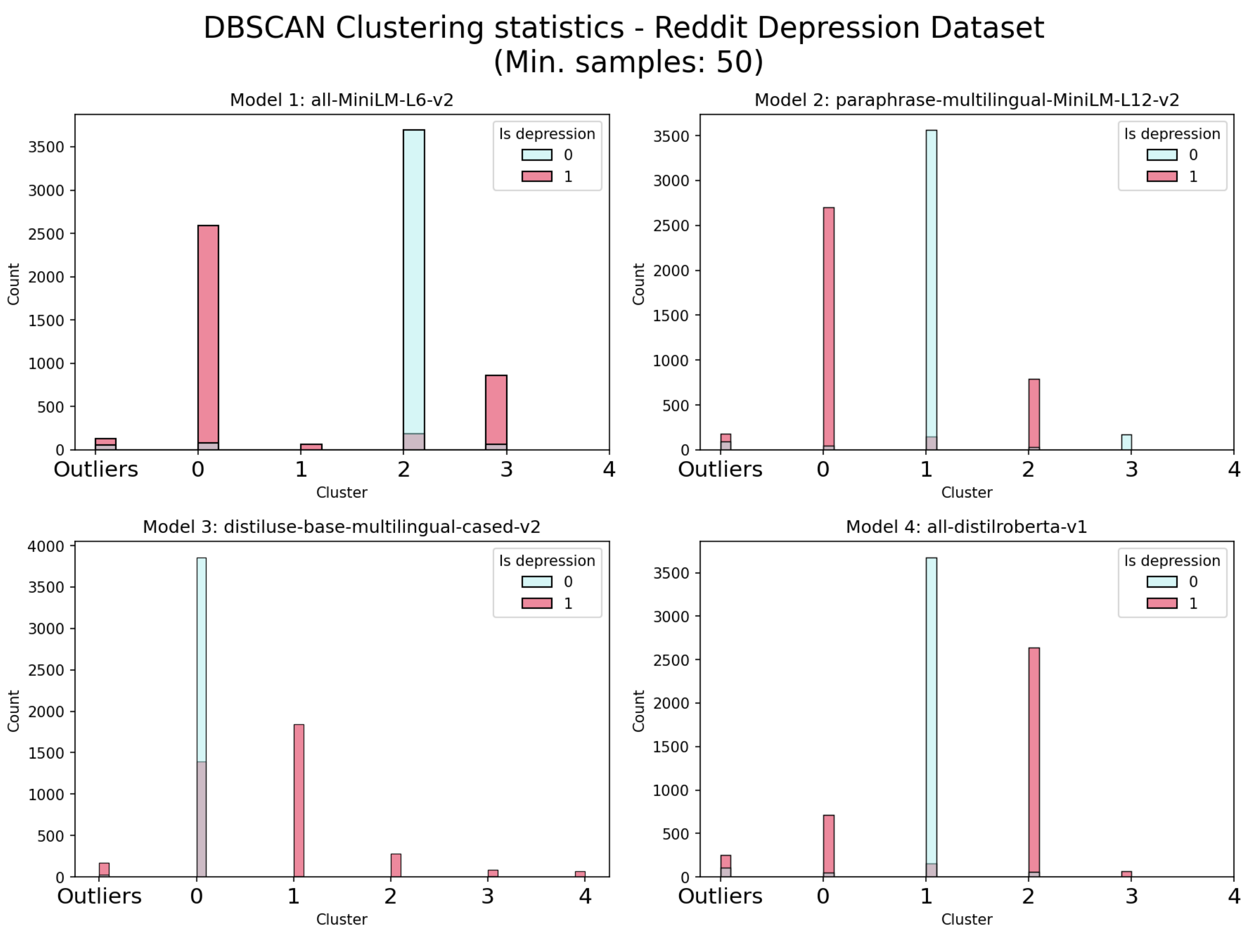


Figure SM1: Statistical information per cluster from large language models.

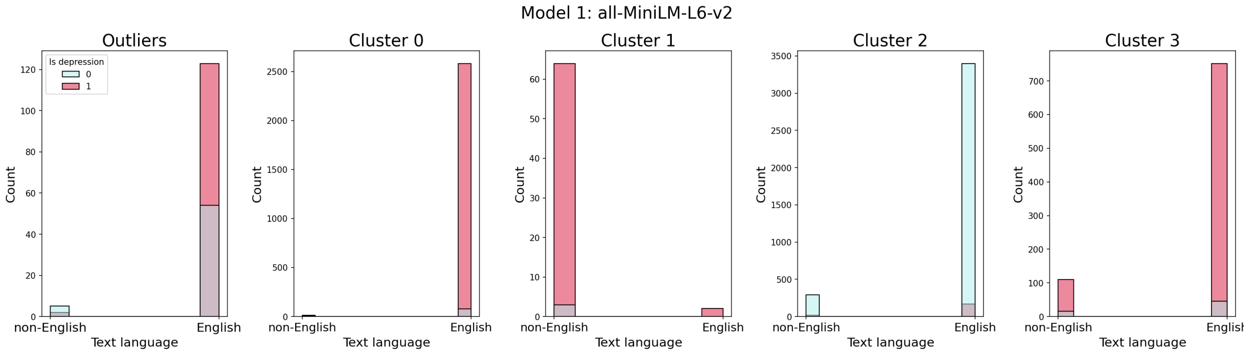


Figure SM2: Model 1 and its sensitivity to non-English text.

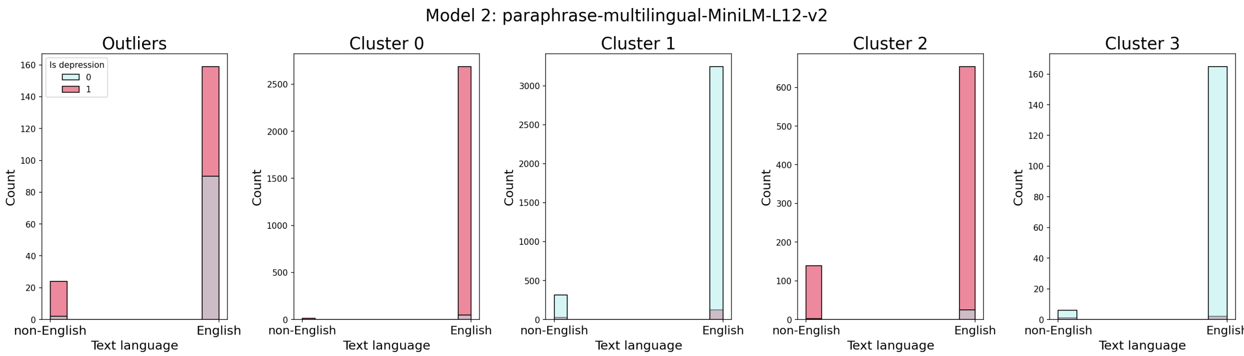


Figure SM3: Model 2 and its sensitivity to non-English text.

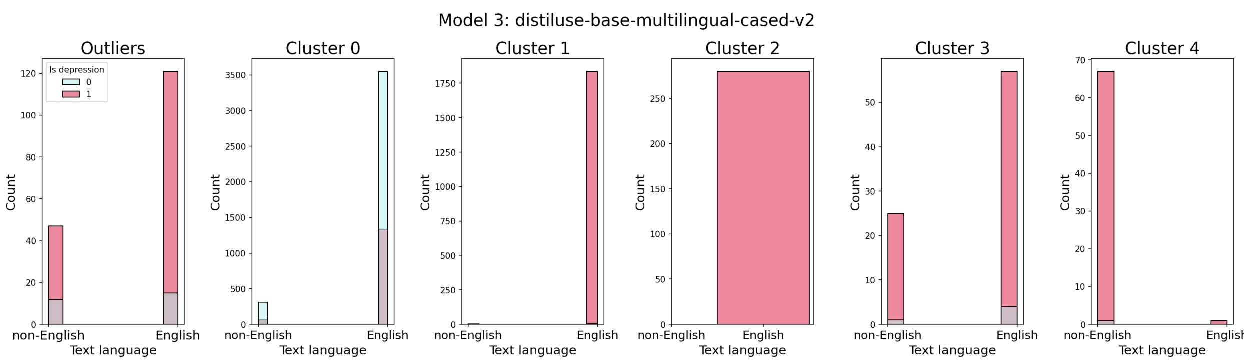


Figure SM4: Model 3 and its sensitivity to non-English text.

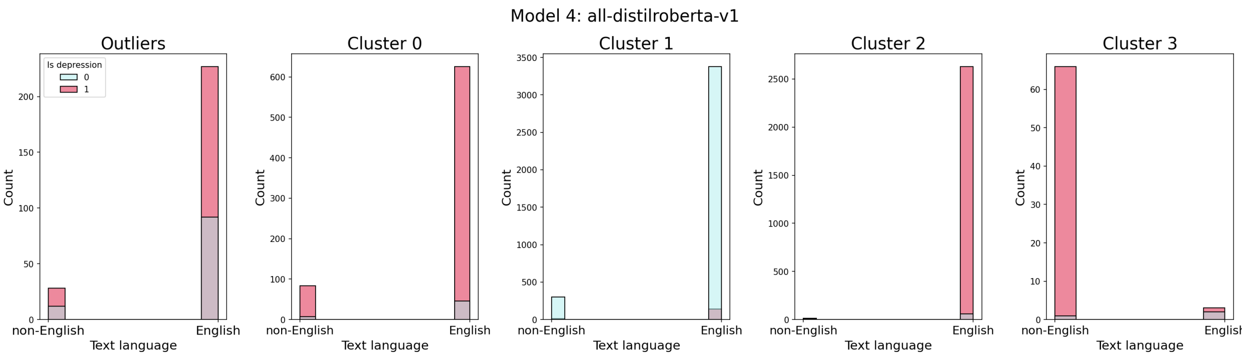


Figure SM5: Model 4 and its sensitivity to non-English text.

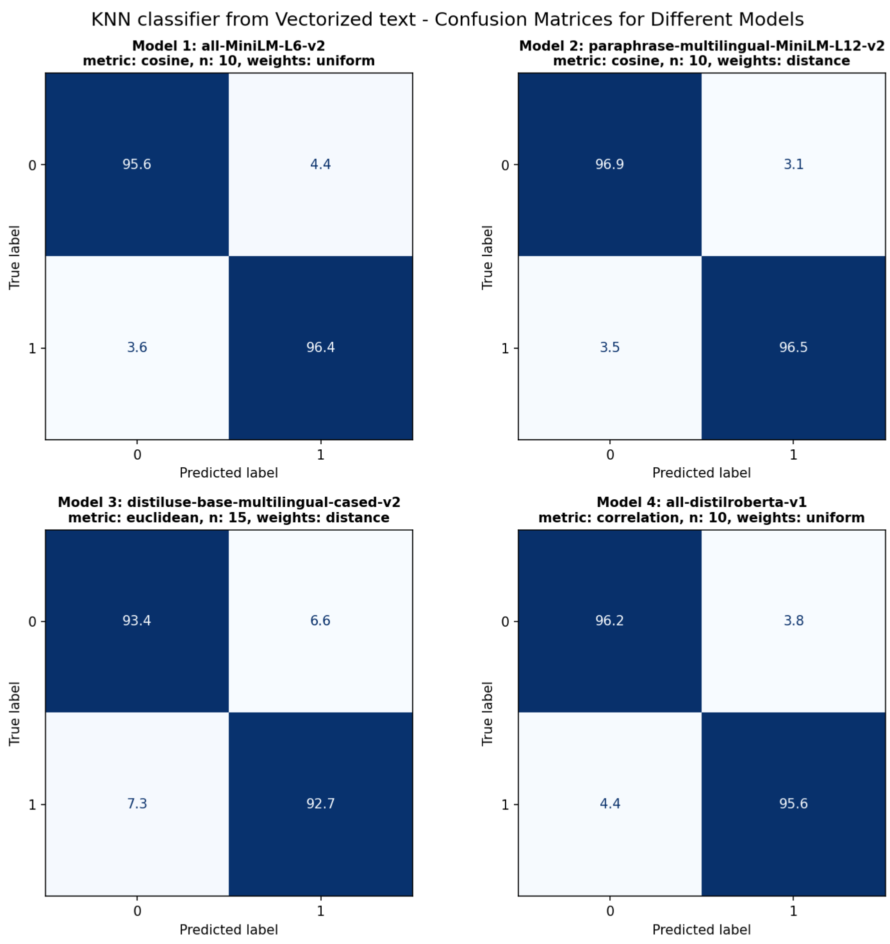


Figure SM6: Confusion matrices comparison between different embeddings.

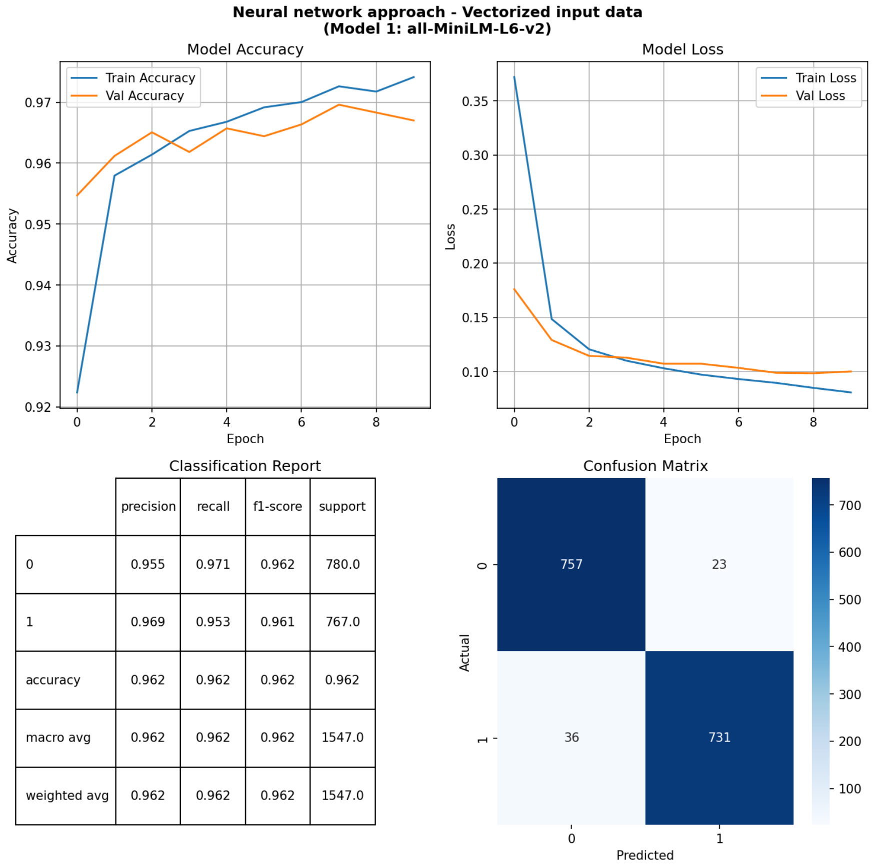
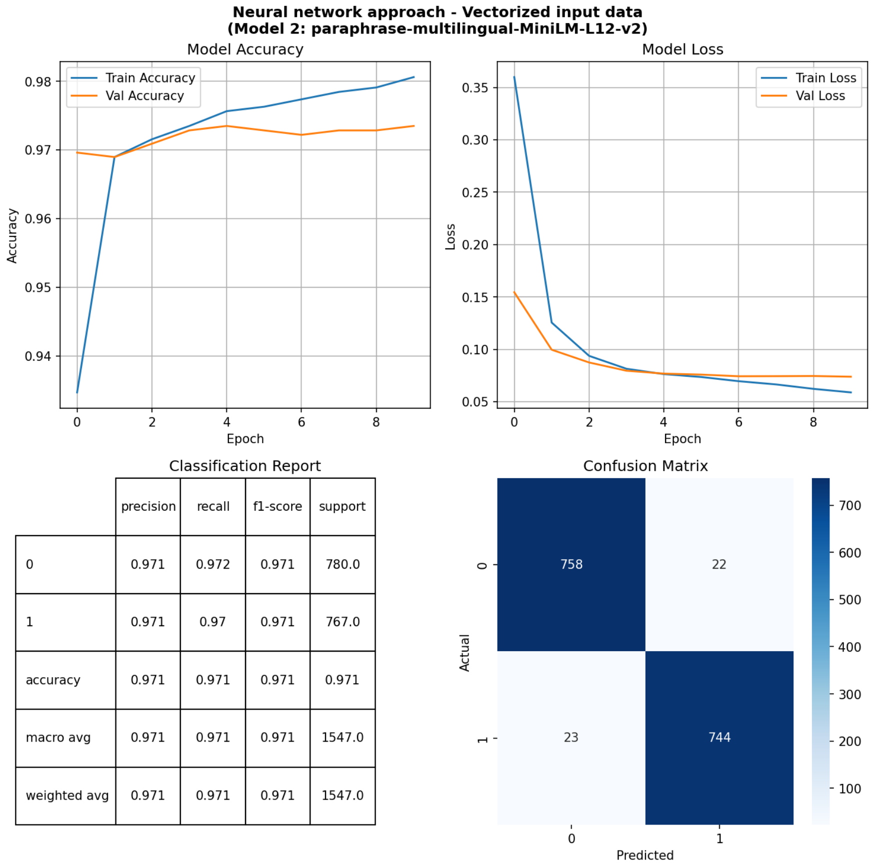
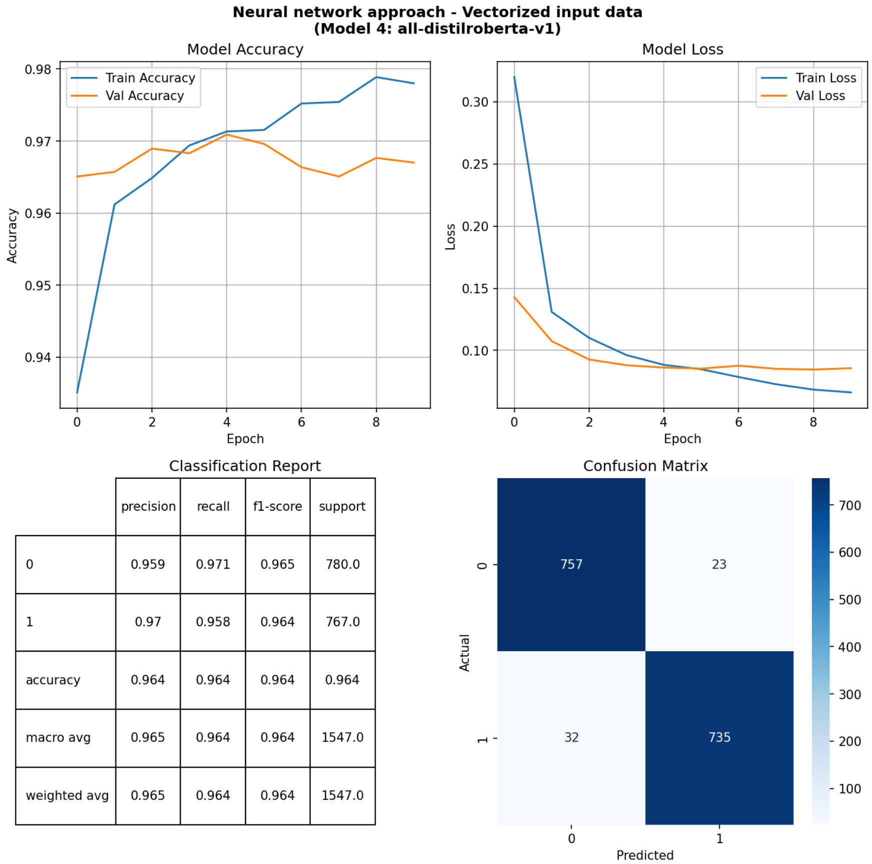


Figure SM7: Neural network training on vectorised text by model 1.



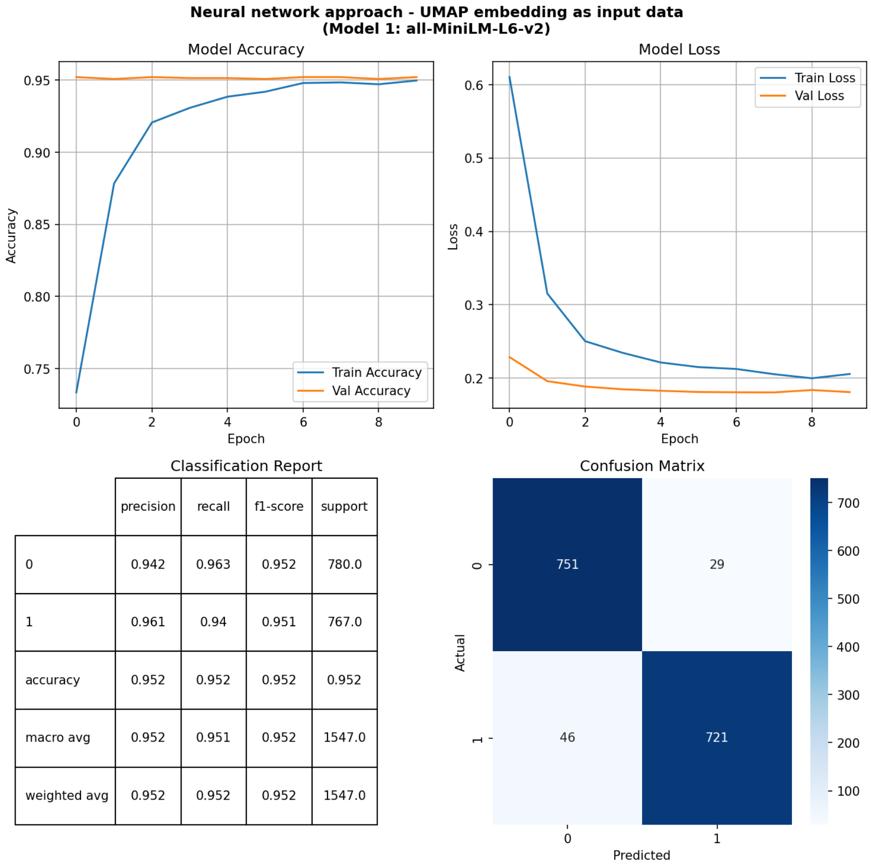
	precision	recall	f1-score	support
0	0.971	0.972	0.971	780.0
1	0.971	0.97	0.971	767.0
accuracy	0.971	0.971	0.971	0.971
macro avg	0.971	0.971	0.971	1547.0
weighted avg	0.971	0.971	0.971	1547.0

Figure SM8: Neural network training on vectorised text by model 2.



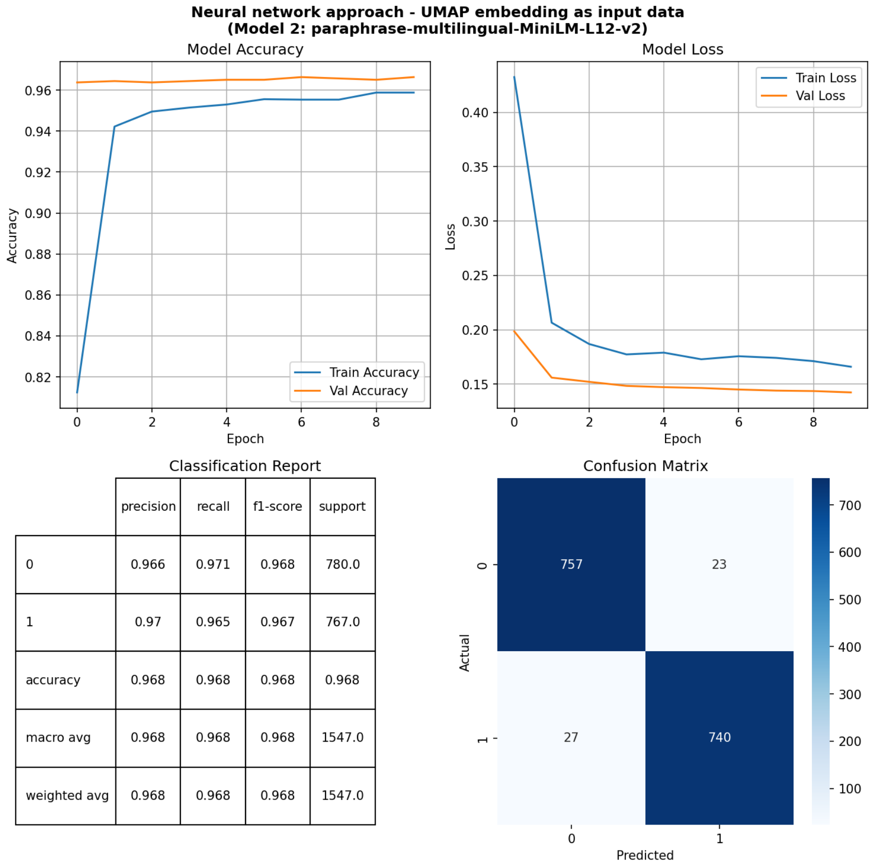
	precision	recall	f1-score	support
0	0.959	0.971	0.965	780.0
1	0.97	0.958	0.964	767.0
accuracy	0.964	0.964	0.964	0.964
macro avg	0.965	0.964	0.964	1547.0
weighted avg	0.965	0.964	0.964	1547.0

Figure SM9: Neural network training on vectorised text by model 4.



	precision	recall	f1-score	support
0	0.942	0.963	0.952	780.0
1	0.961	0.94	0.951	767.0
accuracy	0.952	0.952	0.952	0.952
macro avg	0.952	0.951	0.952	1547.0
weighted avg	0.952	0.952	0.952	1547.0

Figure SM10: Neural network training on UMAP embedding from vectorized text by model 1.



	precision	recall	f1-score	support
0	0.966	0.971	0.968	780.0
1	0.97	0.965	0.967	767.0
accuracy	0.968	0.968	0.968	0.968
macro avg	0.968	0.968	0.968	1547.0
weighted avg	0.968	0.968	0.968	1547.0

Figure SM11: Neural network training on UMAP embedding from vectorized text by model 2.

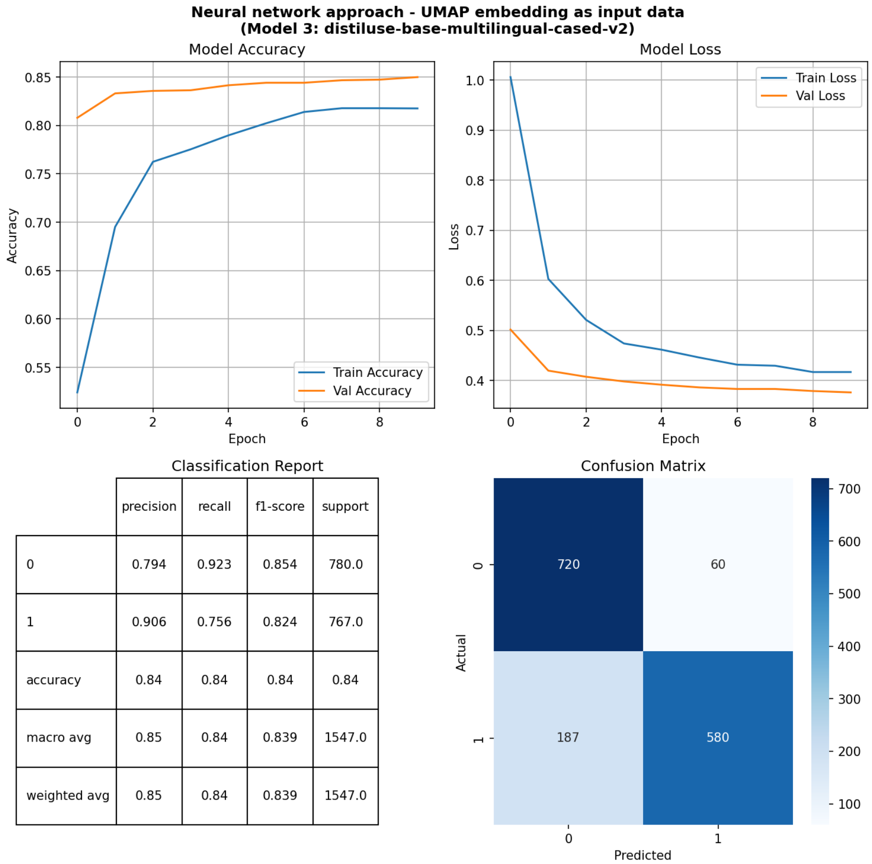
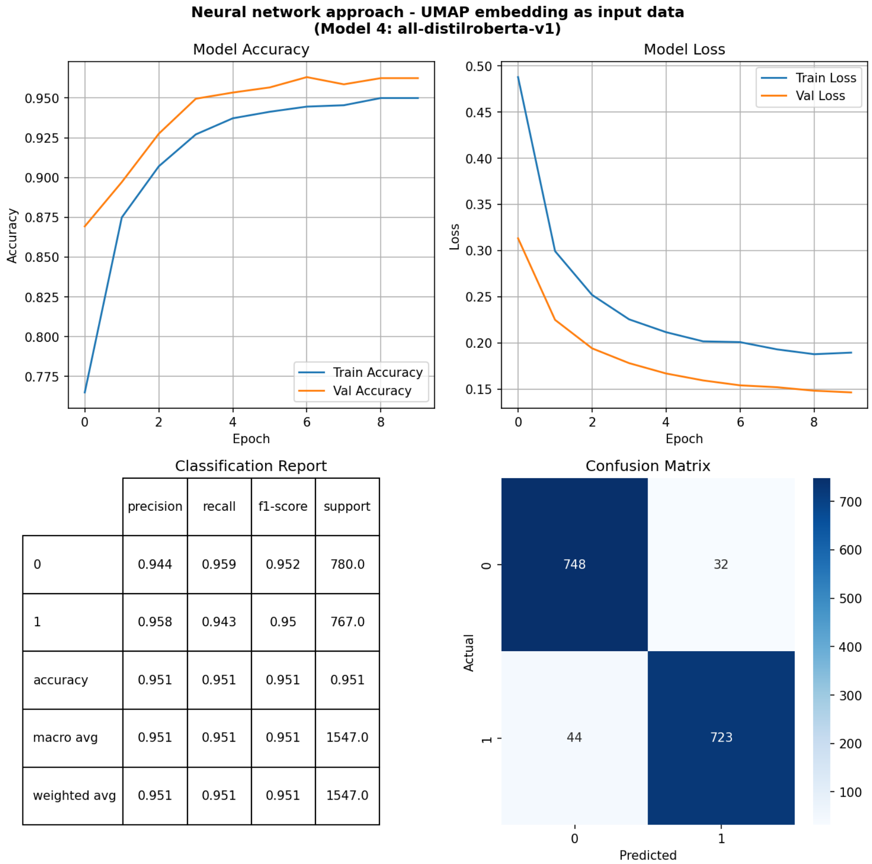


Figure SM12: Neural network training on UMAP embedding from vectorized text by model 3.



	precision	recall	f1-score	support
0	0.944	0.959	0.952	780.0
1	0.958	0.943	0.95	767.0
accuracy	0.951	0.951	0.951	0.951
macro avg	0.951	0.951	0.951	1547.0
weighted avg	0.951	0.951	0.951	1547.0

Figure SM13: Neural network training on UMAP embedding from vectorized text by model 4.