A multi-source graph database to showcase a recommender system for dyslexic students

Karim El Hage, Adel Remadi, Yasmina Hobeika, Ruining Ma, Victor Hong, Francesca Bugiotti

▶ **To cite this version:**

Karim El Hage, Adel Remadi, Yasmina Hobeika, Ruining Ma, Victor Hong, et al.. A multi-source graph database to showcase a recommender system for dyslexic students. IEEE International Workshop on Data Engineering and Modeling for AI (DEMAI) - IEEE Big Data 2023, Dec 2023, Sorrento, Italy. 10.1109/BigData59044.2023.10386535 . hal-04415709

HAL Id: hal-04415709

https://hal.science/hal-04415709v1

Submitted on 24 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A multi-source graph database to showcase a recommender system for dyslexic students

1st Karim El Hage
CentraleSupélec
Gif-Sur-Yvette, France
karim.elhage@student-cs.fr

2nd Adel Remadi
CentraleSupélec
Gif-Sur-Yvette, France
adel.remadi@student-cs.fr

3rd Yasmina Hobeika
CentraleSupélec
Gif-Sur-Yvette, France
yasmina.hobeika@student-cs.fr

4th Ruining Ma
CentraleSupélec
Gif-Sur-Yvette, France
ruining.ma@student-cs.fr

5th Victor Hong
CentraleSupélec
Gif-Sur-Yvette, France
victor.hong@student-cs.fr

6th Francesca Bugiotti
LISN, CNRS, *CentraleSupélec*
Gif-Sur-Yvette, France
francesca.bugiotti@centralesupelec.fr

***Abstract*—This paper addresses the need to better support dyslexic students in higher education using data-driven methods. Our approach lies in modeling relationships between dyslexic students, problems inherent to their condition, and potential solutions. The proposed graph database integrates multiple data sources, in different formats and languages, and captures complex relationships between entities that are not identifiable when considering each data source independently. This paper's main contribution is a hybrid recommender system that first filters potential solutions through the navigation of the modeled graph, then utilizes a Neural Network to solve an ordinal classification problem: effectively ranking the filtered recommendations based on their predicted usefulness. Several documented approaches to solving the ranking algorithm's prediction task were implemented and compared. The models that achieved the highest ranking accuracy, approximately 74%, were 3-Layer Neural Networks trained with Ordinal Log-Loss and self-guided EMD² loss. In summary, our work not only facilitates the identification of patterns essential for crafting personalized recommendations to address the most severe difficulties of dyslexic students, but also establishes a structured foundation for the scalable integration of additional data sources. These results strongly support future research and application development related to dyslexia in higher education.**

***Index Terms*—Dyslexia, Graph Database Modeling, Recommendation System, Ordinal Classification**

I. Introduction

Dyslexia, a genetic variation impacting cognitive tasks such as problem-solving, communication skills, spelling, reading, and learning facts, is characterized by difficulty in acquiring accurate and fluent reading skills [1]. It affects 5-10 percent of the population, with a higher prevalence among children (up to 17.5%) [2] [3]. Consequences include reduced academic self-concept and challenges in mental health [4]. These issues persist beyond the classroom in daily life and adulthood [4].

Treating dyslexia is particularly challenging in higher education settings, as it often goes unnoticed by educators [5] [6]. Dyslexic students face academic challenges, including note-taking, writing assignments, and organization [5]. Moreover, lack of access to proper information hinders their preparedness for higher education [7].

Existing literature highlights the importance of evidence-based interventions for dyslexic students [8]. However, the considerable time and effort required to identify appropriate measures has caused educators not to realize the potential and benefits of evidence-based interventions [8]. Given the mentioned need for evidence-based learning, Vrailexia, an EU-funded project run by a consortium of universities, is collecting multi-lingual data using three sources: questionnaires, transcripts of interviews with experts, and VR (Virtual Reality) simulations. [9]. The Vrailexia project aims to develop tools to address the students' most severe difficulties [9]. It is important to note that the collected data is not medical but is instead a collection of subjective self-assessments from students and expert opinions. This, however, allows for the extraction of data-driven insights that could benefit dyslexic students. At the initial stage of this study, these data sources existed independently, meaning that no integrated pipeline allowed to run multi-source analyses.

For the sake of space, this paper shall not focus on how the data has been extensively pre-processed, extracted, modeled, and consequently made into an interconnected graph. The main contribution of this paper shall instead be on showcasing a powerful use case of such data modeling that is directly relevant to the mission of the Vrailexia project: integrating data from various multi-lingual sources in a centralized graph database and proposing a recommendation system connecting dyslexic students with solutions that tailor to their most severe difficulties. The paper's structure is as follows: Section II introduces related work supporting our scientific methodology. Section III summarizes characteristics of the modeled graph databases and presents the developed hybrid recommendation system. Section IV outlines and discusses the results. Finally, Section V concludes and proposes future avenues for exploration.

This research was co-funded by European Union Committee within the Erasmus+ Program 2014-2020-Key Action 2: Strategic Partnership Project (Agreement n. 2020-1-IT02-K203-080006).

II. Related Work

The related work below illustrates documented methods to use Neo4j as a recommendation system by content filtering. This is complemented by bringing forth previous publications on ordinal classification, the core task of the neural network responsible for ranking identified recommendations.

A. Methods to use Neo4j as a Recommendation System

Prior to proposing a methodology for the recommender system, literature has been studied and compared to understand how Neo4j can be leveraged for such use. In [10], we find a framework for a hybrid medical recommendation system that clusters similar drugs together based on attributes and then uses a neural network to rank the drug within each cluster. A neural network is proposed for the ranking portion to deal with complex and inherent non-linear relationships between drugs and symptoms [10].

In paper [11], the authors discuss using Neo4j as a scalable tool to create relationships between data used by a recommendation system. This is done through a collaborative filtering approach of grouping patients with similar attributes and aggregating their preferences to create a list of highest-rated items [11]. In [12], we find a graph modeled on Neo4j to build a social recommendation system where users receive tailored recommendations only from users they already have some form of connection with (social trust). This allows a very exclusive form of recommendation filtering. Finally, multiple data sources within Neo4j are integrated in [13], where the authors import and use textual similarities to link the various data sources. This integrated graph utilizes Neo4j's graph data science library to create node embeddings that are consequently used to support their movie recommendation system [13]. Taking these related works as inspiration, the proposed recommendation system described in Section III-B also extensively harnesses Neo4j and its data modeling capabilities.

B. Ordinal Classification

The ranking algorithm developed as part of our proposed recommendation system for dyslexic learners involves predicting the usefulness of potential solutions on a scale of 1 to 5, treating it as an ordinal prediction task. This approach, also known as ordinal regression, considers the inherent order between different values and has been shown to outperform classical regression approaches in machine learning [14] [15].

Different techniques exist for ordinal classification. Some convert the task into independent binary classification subproblems [16], while others propose reduction frameworks introducing a final ranking rule [17]. These inspired the development of end-to-end models using CNNs [18], on which other methods such as CORAL [19] and CORN [20] built upon to address classification inconsistencies and improve performance results. Alternatively, another family of techniques has mainly relied on threshold-based methods, such as the proportional odds model (POM) by [21], its adaptations into neural networks by [22], the perceptron ranking (PRank) algorithm proposed by [23] and the GAT surrogate by [15].

Our study has focused on a last family of techniques, called the loss-based approaches, to address the ordinal classification problem discussed in Section IV. These techniques offer the convenience of tackling ordinality without altering models' architecture or setting thresholds [24] [25]. These methods include Weighted Kappa loss (WKL) [26], soft-labels (SOFT) [27], CO2 loss [24], squared Earth Moving Distance (EMD^2) [28], and Ordinal Log-Loss (OLL) [25] [29]. Study [25] experimented with OLL on Computer Vision tasks and showed that it outperformed the CORN method [20] and the CO2 loss [24]. OLL was also used by [29] on Natural Language Processing tasks and reportedly obtained overall better performances than the losses introduced by [19], [26], [27] and [28].

Evaluation metrics for ordinal classification are expected to capture the nuances between adjacent and distant classes. Since the standard accuracy metric falls short on such aspect [24] [29] [30], our study instead considered the following evaluation metrics in Section IV: Mean Absolute Error (MAE), Mean Squared Error (MSE), Maximum Mean Absolute Error (MMAE) [31], and Accuracy within n (ACC_n) [30].

III. Methodology

As previously discussed, we choose Neo4j to store the information from the various data sources. Hence, the conception of the graph schema is conducted in a manner that follows the conventions of graph design and conception. The main focus of this section is to highlight the final graph structure and describe the methodological approach behind the recommendation system, emphasizing how the graph representation is leveraged to filter suggestions, which are then ranked using a neural network algorithm.

A. Multi-Source Integration into a Graph Data Model

Figure 1 outlines the final schema interconnecting the entities from multiple data sources. The entities related to the respondents are sourced from the questionnaire and VR tests. Specifically, the questionnaire was a subjective evaluation that asked respondents to rank the severity experienced by different dyslexia-associated problems and the usefulness of learning tools & strategies in their everyday lives. This study used questionnaire data collected in France and Spain. The ranks, collected on a scale of 1-5, are stored as an edge property of the link between the respondent and the addressed entities. As an example, to find the respondents who consider specific problems to be the most severe, one could use a navigation scheme making use of the `Strength` edge attribute (the property capturing rank values) as follows:

```
(: Respondent)-[: HAS {strength: 5}]->(:Problem)
```

The VR test alternatively gauged respondents both on their confidence level and their reading comprehension skills. Finally, the experts are introduced via interview transcripts that document their opinions and recommendations relating to dyslexia, especially within the context of education. This data source provides the ability to create causality links between problems that respondents were asked to evaluate and the tools and strategies they rated to be beneficial to their everyday lives. However, it is important to consider that semantic differences such as the problem "Reading Difficulties" in the questionnaire and "Difficulty reading and understanding sentences" in the interview would not be linked in the graph database by default. Hence, similarity links were introduced to create connections between nodes of the same entity but originating from different data sources: the usefulness of which shall be demonstrated in the next section.

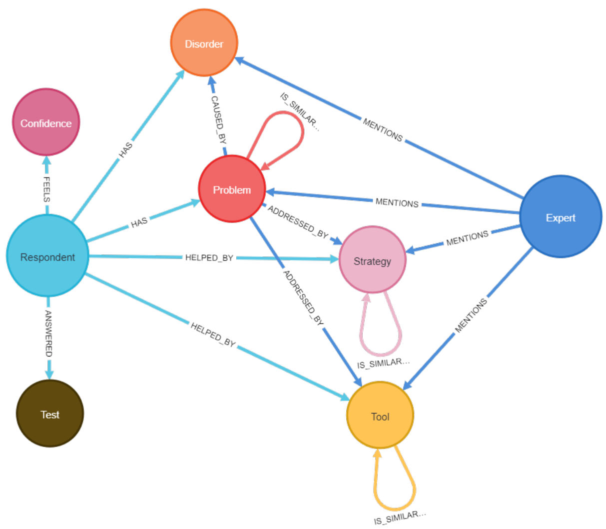

Fig. 1. Final graph representation of the modeled schema

B. Recommendation System

The unique structure of the knowledge graph offers ample opportunities for exploration. A recommendation system that can propose tools and strategies to dyslexic students based on their problems is a relevant use case that can demonstrate this. As part of this implementation, it was decided to focus only on the data originating from the interviews and questionnaires since no expert opinions relating to the VR test insights had yet been collected. The proposed recommendation system consists of two primary components respectively in charge of:

1) Filtering candidate suggestions based on graph navigation by using causal links provided by experts.
2) Ranking suggestions by solving an ordinal classification task with a neural network.

The first building block in charge of filtering suggestions by graph navigation can be thought of as solving a link prediction problem. The link prediction task relies on the graph representation to extract valuable insights. By leveraging inputs from respondents and causal links provided by experts, subsets of relevant strategies and tools can be identified for specific problems. This process takes advantage of the interconnectedness of the data in the graph structure, allowing for the extraction of meaningful relationships between various entities. This component of the recommendation system focuses on `Problem` nodes that a respondent has stated to be very severe (`Strength` edge attribute ≥ 4) to ensure the offering of targeted suggestions. Figure 2 illustrates an example of a respondent that has severe "Reading Difficulties" and for whom an expert recommendation is to propose, via a similarity link, "Audio recording of lessons" as a means to address this severe problem.

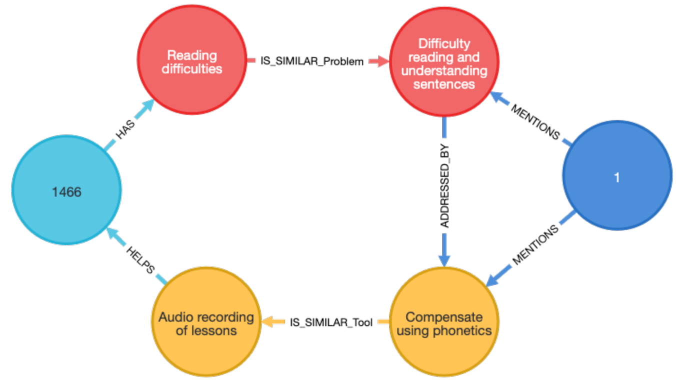

Fig. 2. Graph navigation to retrieve causality link between problems and tools

Using this approach makes it possible to filter lists of tailored suggestions for any respondent in the database, while capitalizing on the insights and connections made by experts in the field of dyslexia. This component also has the advantage of being extremely fast to execute, as it consists of a series of Cypher queries that can be run from a pipeline in Python or any other application with an access to the database.

The second building block, the ranking algorithm, employs a neural network model to predict the usefulness that any respondent would assign to different strategies and tools on a scale of 1 to 5. The core objective of the neural network is therefore to solve an ordinal classification task. As inputs, the model uses various features extracted from the graph database. Once the ordinal classification task is performed, the predictions rank the sets of tools and strategies according to their inferred usefulness. Applying this ranking on the filtered suggestions provided by the first building block produces the final output of the proposed recommendation system - a tailored list of recommendations.

Both the model selection and the final set of input variables to be extracted from the graph database were determined based on performance on the test set. While more than 30 features were considered, it was finally concluded that only the respondents' disorders and problems would be selected, representing 17 input features. Considering the substantial number of target variables (about 40 tools and strategies), a neural network was employed for its ability to handle non-linear multi-output problems.

In terms of modeling, several architectures and loss functions were considered and compared. The model architecture was tuned to start from a 1-Layer network (with no activation function) up to a 4-Layer setup (with ReLU activations). It was decided to test and compare several appropriate loss functions cited in Section II-B. The Ordinal Log Loss (OLL) [25] [29], the self-guided EMD² loss (EMD²) [28], the CO2 loss [24] and the Soft-labels loss (SOFT) [27] have been implemented to address the problem of ordinality. The Mean Squared Error (MSE) and the Cross-Entropy (CE) were also employed for comparison, addressing the problem as regression and classification tasks, respectively. The model's predictive performance was evaluated using the ACC_1 (Accuracy within 1) [30], MAE, RMSE, and MMAE [31] metrics.

IV. Results and Discussion

In terms of evaluation, there was no quantitative way to benchmark the performance of the hybrid recommendations (graph navigation filtering + ranking algorithm) since they relied on the opinions of experts and would require feedback from dyslexic respondents. However, it was still possible to demonstrate that our data modeling has led to the successful ability to leverage the recommendations from experts using the graph navigation filtering. Furthermore, it is possible to quantitatively evaluate the ordinal classification algorithm used to rank these recommendations.

A. Evaluation of the Ordinal Classification Algorithm

The final ordinal classification model was determined after comparing the performance of various architectures and losses. As part of the model selection process, hyperparameters such as the size of hidden layers, the learning rate of the Adam optimizer, and the number of epochs were tuned. The implemented losses also involved hyperparameters that were considered while tuning the models. Table I summarizes the best performance obtained per implemented loss function.

A critical detail is that the only evaluated outputs were those for which the usefulness was filled by a respondent since a blank answer would indicate they had not previously used the concerned tool or strategy. This minimizes the possibility of bias in the evaluation. During training, blank values of usefulness were replaced by the training set's sample means. This led to better performance results than when replacing blanks with zero, median, or mode.

TABLE I
BEST MODEL PER IMPLEMENTED LOSS FUNCTION

Loss	Best Architecture	ACC_1	MAE	RMSE	MMAE
EMD²	3-Layer NN	**74.46 %**	**1.04**	1.34	1.81
OLL	3-Layer NN	74.13 %	**1.04**	**1.33**	1.80
MSE	2-Layer NN	72.69 %	1.07	1.37	1.83
CO2	2-Layer NN	68.90 %	1.14	1.54	2.70
SOFT	4-Layer NN	66.01 %	1.19	1.61	**1.75**
CE	3-Layer NN	65.53 %	1.18	1.58	1.76

Among the various explored models, the OLL and the self-guided EMD² reached the highest ACC_1 of 74.13% and 74.46%, both with 3-Layer Neural Networks. As commented by the authors of [29], a value of 1.5 proved to be a good tradeoff for the OLL's hyperparameter α. The hyperparameters selected for the self-guided EMD² loss were $\lambda = 10^9$, $\omega = 1.5$ and $\mu = 0$.

Interestingly, setting $\lambda = 10^9$ amounts to discard the Cross-Entropy term. In [28], the EMD² term was introduced only as a regularizer to the Cross-Entropy loss because it faced convergence issues. These issues were also encountered experimentally by [29]. This observation was, however not met in our study, possibly due to the different nature of the data and the lower complexity in model architectures.

Both the self-guided EMD² and OLL losses showed significantly better performance than the other implemented losses on all metrics except MMAE. The best models on that metric were the Cross-Entropy (CE) and the Soft-labels (SOFT) losses. However, these models obtained significantly worse performances on all other metrics. While Cross-Entropy is known not to be the most appropriate choice for ordinal classification [24], the soft-labels relying on a label embedding designed explicitly for ordinal prediction tasks only had a slightly better performance. In light of these results, it was eventually decided to select the model using the self-guided EMD² loss as the ranking algorithm of the recommendation system.

B. Demonstration of the Recommendation System Use case

In this section, a randomly selected dyslexic student is taken as an example to illustrate the results and corresponding discussion of using the hybrid recommendation system.

TABLE II
EXAMPLE OF RECOMMENDATIONS

Most Severe Problems	Recommended tools
Reading Difficulties	1) Use a special font for easy reading 2) Use Audio Books 3) Numerical tutor (e.g., Siri) to which it is possible to query verbal explanations on challenging concepts 4) Words written in different colors
Difficulties to focus during online courses	1) A clearer presentation of the study material
Difficulties to understand complex or rare words	1) Register courses 2) Underline text with different colors 3) Conceptual sketches made by oneself 4) Repeat the studied contents 5) Summaries prepared by oneself

The first step would be to identify the respondent's most significant problems and utilize the graph navigation component of the recommendation system to use the experts' opinions and filter a list of relevant strategies and tools. In parallel, the trained ordinal classification model would take the respondent's declared disorders and problem strengths as inputs to infer the usefulness of all the strategies and tools. Combining both components by ranking the filtered suggestions in decreasing order of usefulness provides a tailored list of recommendations. Table II illustrates, as an example, the random respondent's three most severe problems and the corresponding top 5 tools to address each of them.

The results show that by navigating the graph, it is possible to use the knowledge of experts to recommend tools and strategies to address the most severe difficulties of dyslexic students who have answered the questionnaire. The technique is limited by the amount of expert interview data that exists. For example, the second problem displayed in Table II only has one recommendation because the current experts' opinions stored in the database only refer to this one tool as a method to address it. Hence, scaling the database to include more data will make such recommendations more refined.

V. Conclusion and Future Work

In summary, this study aimed to provide tailored recommendations and an integrated knowledge base for challenges faced by dyslexic students in higher education, demonstrating the effective use of Neo4j. The developed graph database schema addressed key issues, suggesting personalized tools and strategies based on respondents' profiles and difficulties. The ordinal prediction task was found to be best addressed by the self-guided EMD² [28] and OLL [25] [29] losses achieving ACC_1 performances of 74.46 % and 74.13 % respectively.

Moving forward, it is possible to improve on several aspects of the work to produce a scalable and credible tool to be used by the stakeholders of the Vrailexia project. One limitation of the work is that the recommendations have yet to be qualitatively tested over a period of time. Through the introduction of Students/Experts-in-the loop, it would not only be possible to validate the obtained results, but also improve the recommendations given by this hybrid model. Considering the ethical concerns of providing recommendations in such a domain, it is important to introduce such checks into the pipeline. Other current limitations of our study are the manual extraction of entities from the interview transcripts and the manual creation of similarity links between phrases from various data sources. Named Entity Recognition models or Large Language Models (LLM) could be employed to automatically process textual inputs, determine similarity links and insert data into the database. Such a contribution would allow the Vrailexia project to significantly scale up the knowledge graph on expert interview transcripts and even on any other unstructured textual sources. This would in turn increase the potential of the proposed recommendation system and offer countless other opportunities for research aiming at helping dyslexic students.

References

[1] M. Snowling, C. Hulme, and K. Narion, "Defining and understanding dyslexia: past, present and future," Oxford Review of Education, vol. 46, pp. 501–513, 2020.

[2] J. Roitsch and S. Watson, "An overview of dyslexia: definition, characteristics, assessment, identification, and intervention," Science Journal of Education, vol. 7, pp. 81–86, 2019.

[3] S. Shaywitz and B. Shaywitz, "Dyslexia (specific reading disability)," Biol Psychiatry, vol. 57, pp. 1301–1309, 2005.

[4] A. Wilmot, P. Hasking, S. Leitão, E. Hill, and M. Boyes, "Understanding Mental Health in Developmental Dyslexia: A Scoping Review," International Journal of Environmental Research and Public Health, vol. 20, p. 1653, 2023.

[5] M. Pino and L. Mortari, "The Inclusion of Students with Dyslexia in Higher Education: A Systematic Review Using Narrative Synthesis," Dyslexia, vol. 20, pp. 346–369, 2014.

[6] S. Riddel and E. Weedon, "What counts as a reasonable adjustment? Dyslexic students and the concept of fair assessment," International Studies in Sociology of Education, vol. 16, pp. 57–73, 2006.

[7] M. Madriga, "Enduring disablism: students with dyslexia and their pathways into UK higher education and beyond," Disability & Society, vol. 22, pp. 399–412, 2007.

[8] T. Santangelo, A. Ruhaak, A. Kama, and B. Cook , "Constructing effective instructional toolkits: a selective review of evidence-based practices for students with learning disabilities," Emerald Group Publishing Limited, vol. 26, pp. 221–249, 2013.

[9] Vrailexia "Home Page." https://vrailexia.eu/ (accessed Aug. 03, 2023).

[10] Q. Zhang, G. Zhang, J. Lu and D. Wu, "A Framework of Hybrid Recommender System for Personalized Clinical Prescription," in 10th Int. Conf. on Int. Sys. and Know. Eng., 2015, pp. 189–195.

[11] B. Stark, C. Knahl, M. Aydin, M. Samarah and K. O. Elish, "BetterChoice: A migraine drug recommendation system based on Neo4J," in 2nd IEEE Int. Conf. on Comp. Intel. and App., China, 2017, pp. 382–386.

[12] A. Virk and R. Rani, "Efficient Approach for Social Recommendations Using Graphs on Neo4j," in 2018 Int. Conf. on Inv. Research in Comp. App., 2018, pp. 133–138.

[13] A. Origlia, M. Di Bratoo, M. Di Maro, and S. Menella, "A multi-source graph representation of the movie domain for recommendation dialogues analysis," in 13th Lang. Res. and Eval. Conf., 2022, pp. 1297–1306.

[14] S. Zhang, L. Yang, M. Mi, X. Zheng, and A. N. Yao, "Improving Deep Regression with Ordinal Entropy," 2023, arXiv:2301.08915.

[15] F. Pedregosa, F. Bach, A. Gramfort, "On the consistency of ordinal regression methods," Journal of Machine Learning Research, vol. 18, pp. 1–35, 2017.

[16] E. Frank, M. Hall, "A simple approach to ordinal classification," in 12th Euro. Conf. on Mach. Lear., 2001, pp. 145–156.

[17] L. Li and H. Lin, "Ordinal regression by extended binary classification," in Adv. Neural Inform. Proces. Syst. 19, 2006, pp. 865–872.

[18] Z. Niu, M. Zhou, L. Wang, X. Gao and G. Hua, "Ordinal Regression with Multiple Output CNN for Age Estimation," in 2016 IEEE Conf. on Comp. Vis. and Pat. Recog., 2016, pp. 4920–4928.

[19] W. Cao, V. Mirjalili, S. Raschka, "Rank consistent ordinal regression for neural networks with application to age estimation," Pattern Recognition Letters, vol. 140, pp. 325–331, 2020.

[20] X. Shi, W. Cao, S. Raschka, "Deep neural networks for rank-consistent ordinal regression based on conditional probabilities," Pattern Analysis & Applications, vol. 26, pp. 941–955, 2023.

[21] P. McCullagh, "Regression models for ordinal data", Journal of the Royal Statistical Society, vol. 42, pp. 109–142, 1980.

[22] J. Verwaeren, W. Waegeman, B. DeBaets, "Learning partial ordinal class memberships with kernel-based proportional odds models," Computational Statistics and Data Analysis, vol. 56, pp. 928–942, 2012.

[23] K. Crammer, Y. Singer, "Pranking with Ranking," in Adv. in Neur. Inf. Proc. Sys. 14, 2001, pp. 641–647.

[24] T. Albuquerque, R. Cruz, JS. Cardoso, "Ordinal losses for classification of cervical cancer risk," PeerJ Computer Science 7, 2021.

[25] G. Polat, I. Ergenc, H. T. Kani, Y. O. Alahdab, O. Atug, and A. Temizel, "Class distance weighted cross-entropy loss for ulcerative colitis severity estimation," in Med. Ima. Under. and Anal., 2022, pp. 157–171.

[26] J. de la Torre, D. Puig, A. Valls, "Weighted kappa loss function for multiclass classification of ordinal data in deep learning," Pattern Recognition Letters, Vol. 105, pp. 144–154, 2018.

[27] L. Bertinetto, R. Mueller, K. Tertikas, S. Samangooei and N. A. Lord, "Making Better Mistakes: Leveraging Class Hierarchies With Deep Networks," in 2020 IEEE/CVF Conf. on Comp. Vis. and Pat. Recog., 2020, pp. 12503-12512.

[28] L. Hou, C.P. Yu, and D. Samaras, "Squared earth mover's distance-based loss for training deep neural networks," 2016, arXiv:1611.05916.

[29] F. Castagnos, M. Mihelich, C. Dognin, "A simple log-based loss function for ordinal text classification," in 29th Int. Conf. on Comp. Ling., 2022, pp. 4604–4609.

[30] L. Gaudette, N. Japkowicz, "Evaluation Methods for Ordinal Classification," in 22nd Conf. Adv. in Art. Intel., 2009, pp. 207–210.

[31] M. Cruz-Ramírez, C. Hervás-Martínez, J. Sánchez-Monedero, P.A. Gutiérrez, "Metrics to guide a multi-objective evolutionary algorithm for ordinal classification," Neurocomputing, Vol. 135, pp. 21–31, 2014.