Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2024) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

28th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2024)

Towards an Intelligent Model for Dysgraphia Evolution Tracking

Redouane Bouhamoum[a,*], Maroua Masmoud[a], Youssef Lyousfi[a], Hajer Baazaoui[a], Deepti Mehrotra[b]

[a]*ETIS UMR 8051, CY University, ENSEA, CNRS, Cergy, France*
[b]*Amity University Uttar Pradesh, Noida, India*

Abstract

Learning disabilities present significant barriers in the lives of individuals, particularly children and students, as they can impede their learning process and skill development. Dysgraphia, a form of learning disability, can adversely affect an individual's writing ability. While various approaches have been proposed to detect learning disorders, there is a lack of methods for tracking the progression of these disorders.

In this work, we propose an intelligent model for tracking the evolution of dysgraphia. Our approach utilizes a probabilistic machine learning algorithm to compute a Dysgraphic class score for each individual. By computing this score at different intervals, we can monitor the individual's progress over time. To achieve this, we trained various probabilistic classifiers and fuzzy clustering algorithms on a labeled dataset to select the model with the best performance for tracking. Our experimental evaluation demonstrates that our model successfully tracks the evolution of individuals with dysgraphia.

Keywords: Dysgraphia, Evolution Tracking, Artificial Intelligence, Probabilistic Machine Learning ;

1. Introduction

Learning disabilities represent a significant obstacle in the lives of individuals, particularly children and students, as they can hinder their learning process and the acquisition of skills. Dysgraphia is a disorder characterized by difficulties in the acquisition of writing skills [1]. Early diagnosis of dysgraphia allows for assisting children in improving their handwriting, and adapting teaching methods accordingly. However, the diagnosis is a time-consuming and costly process requiring professional intervention.

Recent advances in artificial intelligence could make the diagnosis of dysgraphia easier and less costly, greatly helping those affected. Several works have proposed intelligent approaches to detect individuals with a particular type of learning disability [2]. Other approaches have been proposed to detect various types of learning disability, etc. [3]. Some studies also suggest that people with dysgraphia can improve their writing and reading skills [4].

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000.
E-mail address: redouane.bouhamoum@ensea.fr

However, the aspect of evolution tracking of Dysgraphic people has not been addressed. To the best of our knowledge, there is no existing approach for tracking the progress of individuals with dysgraphia. Indeed, effective evolution tracking is crucial as it enables the adaptation of learning scenarios according to the progress or regression of the disease. We point out that the study presented in this paper is part of a research line with a wider scope. In a previous work, a novel system for helping people suffering from dysgraphia was proposed, tackling spelling mistakes, grammatical mistakes and poor handwriting quality [5]. In this work, we propose an approach for automatically tracking the evolution of dygraphic individuals over time. The proposed approach, based on probabilistic classifiers, enables us to compute the probability of belonging to each class (Dysgraphic / Non-Dygraphic). We will use this probability as a score or degree to determine whether a Dysgraphic person shows possible improvement over time or not. The present paper details our proposal to perform the evolution tracking task.

The rest of the paper is organized as follows. Section 2 discusses the state of the art related to learning disabilities detection and treatment. The general idea of our approach is introduced in section 3. Section 4 details the phases of the construction of our evolution tracking model. Experimental results are presented in section 5. Finally, a conclusion is provided in section 6.

2. Related Work and Motivations

Several works have utilized machine learning (ML), deep learning techniques (DL) and fuzzy logic to address the problems of detecting Dysgraphic persons and providing assistance through corrections to their handwriting, along with proposing exercises as part of their treatment.

The detection of dysgraphia is crucial, as early diagnosis enables children to perform well at home and school allowing parents to adapt to the environment according to the child's needs [6]. Consequently, several approaches have been proposed to automatically detect dysgraphia using ML algorithms. For example, the the Adaptive Boosting (AdaBoost) algorithm achieves an accuracy of 80% [7] while the Random Forest algorithm demonstrates an accuracy up to 99% [8, 9]. Additionally, other approaches have employed DL algorithms such as the Deep Transfer Learning which achieves an accuracy of 99.75% for dysgraphia detection [10], and the Convolutional Neural Network with an 87% accuracy [11].

While the previous approaches addressed the problem of dysgraphia detection, other works have proposed to assist Dysgraphic persons in improving their writing skills. Quenneville [12] mentioned various tools that aid those with learning disabilities, such as word predictors and auditory feedback. The approach presented by Scheidl et al. [13] used TensorFlow, a popular ML tool in python, to achieve accurate handwriting recognition. Toutanova et al. [14] described a spelling correction algorithm based on pronunciation. In our previous work, we presented an approach that combines multiple solutions for efficient handwriting correction, including handwritten text recognition using a CNN-RNN-CTC model, a spelling correction model based on the SymSpell and Phoneme models, and grammar correction using the GECToR model.

In this work, we tackle the challenge of tracking the evolution of dysgraphia using probabilistic ML. Researchers in [15] utilized unsupervised learning with the K-Means clustering method to categorize children into three clusters. Cluster C1 exhibited mild dysgraphia, while C2 and C3 exhibited severe dysgraphia. However, the classification lacks the provision of evolution information if the child does not transition through the different clusters. To the best of our knowledge, there is no existing proposal addressing the tracking of the evolution of dysgraphia with a score reflecting the individual's situation at each evaluation.

3. General Approach

Our tracking approach for Dysgraphic individuals aims to automatically evaluate the evolution of the disease over time. Our proposal is based on a fuzzy classification algorithm that assigns a score to each individual indicating their likelihood of belonging to the Dysgraphic class. The intuition is that this score reflects the stage of the disease, allowing us to monitor the evolution of dysgraphia.

To achieve that, we compare different ML methods to build a decision model capable of assigning a score to an individual indicating their likelihood of belonging to the Dysgraphic class and thus providing a degree of evolution. Figure 1 provides an overview of our approach, illustrating the different steps in the training phase and the principles of tracking phase.

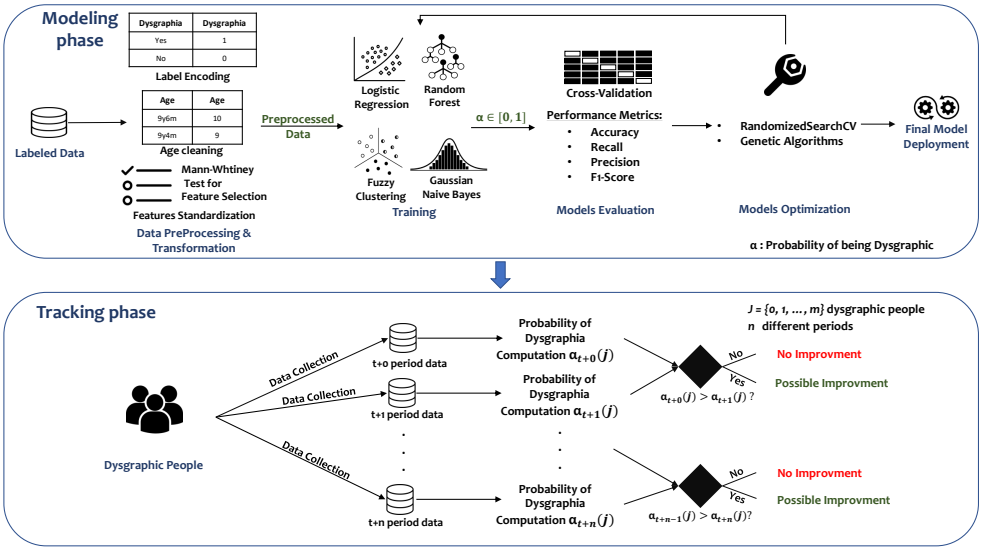


Fig. 1: Overview of the Dysgraphia Tracking Approach

As shown in this figure, the first phase consists in building the tracking model. It starts by applying necessary pre-processing techniques to the labeled dataset and selecting important features in order to reduce learning time and avoid overfitting. Subsequently, well-known fuzzy ML algorithms are applied on the pre-processed and labeled dataset. Then, based on classification metrics (F1-Score, Recall, Precision and Accuracy), we select the best fuzzy algorithm for classifying Dysgraphic persons. Each model undergoes an optimization using various techniques to determine the combination of parameters that improve the model's quality. The algorithm achieving the highest test scores is selected and deployed for the next phase. Finally, the constructed model is used in the second phase to monitor the evolution of Dysgraphic persons. At each data collection point, individuals are classified, and based on the computed score, we determine whether they have experienced positive or negative evolution.

4. Proposed Dysgraphia Evolution Tracking Model

The object of our proposed model is to assess the evolution of Dysgraphic persons over time. We detail in this section the driven dataset used to conduct the construction of our model, the features selection stage, and the selection of the ML algorithm that provides the best result regarding the tracking of dysgraphia.

4.1. Description of the dataset

The dataset we used in this work to build our model is collected by [16]. They recruited 280 children, including 62 Dysgraphic children. The participants underwent the BHK test, which involves reproducing text to identify various attributes (such as letter shape, size, alignment, spacing, etc.) used to calculate a final handwriting score.

Then, 12 different digital features were selected from the work of Asselborn [9], based on their importance on the simple binary classification between children with or without dysgraphia. These features can be categorised as follows:

- **Static features**: Representing the geometric characteristics of the written text. The static features in the dataset include Space Between Words, SD of Handwriting Density, Median of Power Spectral of Tremor Freq.
- **Kinematic features**: Features describing the dynamic of the handwriting process. kinematic features consist of Median of Power Spectral of Speed Freq, Distance to mean of Speed Freq, and In Air Time Ratio.
- **Pressure features**: Features using the notion of pressure measured between the pen tip and the tablet surface. Pressure features include Mean Pressure, Mean Speed of Pressure Change, and SD of speed of Pressure Change.
- **Tilt features**: Grouping features using the notion of tilt between the pen and the surface of the tablet. Tilt features comprise Distance to Mean of Tilt-x Freq, Bandwidth of Power Spectral of Tilt-x Freq, and Median of Power Spectral of Tilt-y Freq.

The dataset includes additional features that capture general information about the children involved in this study. Three categorical variables represent additional information (Gender, Class, Laterality), one numerical feature which is the Age, and there is an ID to identify each individual. Finally, there is a grouping variable indicating whether each child belongs to the non-Dysgraphic or Dysgraphia group. This results in a dataset with a total of 20 columns, including the ID of each child.

4.2. Feature Selection

The data pre-processing and feature selection phases are essential concepts in a ML project, aiming to prepare the data for modeling and optimize results and performances. In this work, we perform various data transformations. First, we engage in feature reduction to decrease the number of features, thereby reducing learning time and preventing overfitting. Subsequently, we generate a polynomial combination of features, allowing the creation of more complex models that better capture the nonlinear relationships between the features and the target variable.

First, we select significant features using Mann-Whitney U test, a statistical method suitable for comparing the central tendency of two groups without assuming the normal distribution of data. The test relies on the MW U statistic, which measures the difference between the ranks of the observations in each group [17].The significance value p_value was set to $p < 0.05$. Table 1 lists the selected features with their test scores and the p_values.

Table 1: Mann-Whitney U test Results

Features	Score	p_value
BHK_raw_quality_score	164.5	0
median_Freq_speed	384	0
mean_d_P	802	0
dist_Freq_tilt_x	837	0
dist_Freq_speed	988	0
median_Freq_tilt_y	1002	0
Space_Between_Words	1135	0
bandwidth_tilt_x	1218	0
std_d_P	1218	0
mean_Pressure	1260	0
in_Air	1268	0
BHK_raw_speed_score	1275	0
std_Density	1448	0.01
median_Freq_tremolo	1500	0.02
Age	1596	0.05

Regarding the categorical variables, we applied chi-squared test to determine if there is an association between the features and the Dysgraphia column. However, none of the categorical features were selected.

After selecting the features, we applied polynomial features to the chosen ones to generate polynomial combinations. This allows the creation of more complex models that can better capture nonlinear relationships between the features and the target variable. Then, we applied Sequential Features Selection to select the final combinations that will be used to train the model.

4.3. Selection of ML Algorithm

In this section, we apply a number of well-known classification and clustering algorithms to the prepared dataset to select the ML algorithm that yields the best results. Then, we evaluate the results provided by each algorithm based on the classification metrics. The algorithms are applied to the dataset several times, and the parameters of the algorithms are optimized after each iteration until we obtain the best results.

4.3.1. ML Algorithms

Fuzzy C-Means (FCM): It is a clustering algorithm that assigns each data point to a cluster based on its degree of membership. In FCM, each data point can belong to multiple clusters, with degree of belonging computed using the matrix U [18].

Random forest: This Classifier is a bagging technique that combines multiple decision trees to predict final value. For a given number of trees N, Random Forest Classifier builds and trains N Decision Trees in order to get the prediction of each one and computes the Mode as final result [19].

Gaussian Naive Bayes (GNB): Naïve Bayes (NB) is a supervised learning algorithm that uses Bayes' theorem to calculate the probability of a given instance belonging to a certain class. Gaussian Naive Bayes is a variant of Naive Bayes that assumes the likelihood of the features to be Gaussian. The GNB model uses MAP estimation to calculate the mean and variance of the feature values for each class. These parameters are then used to calculate the conditional probability of the features given the class, which is used in the classification [19]. The Gaussian Naive Bayes assumes the normality of the data to provide accurate results; therefore, the normality of our data needs to be evaluated [20].

Normality Test: Normal distribution, also known as Gaussian distribution, is one of the most important distributions in statistics and ML, playing a crucial role in various algorithms and statistical tests. One statistical test employed to assess the normality of data is the Kolmogrov-Smirnov test. This test evaluates the goodness-of-fit between observed data and the normal distribution, offering a measure of the agreement between empirical data and the excepted characteristics of a Gaussian distribution. The null hypothesis of the Kolmogrov-Smirnov test states that data are drawn from a normally distributed population, with a p_value greater than 0.05 indicating acceptance of this hypothesis [21].

Table 2: Kolmogrov-Smirnov test Results

Features	Score	p_value
BHK_raw_quality_score	1	0
median_Freq_speed	0.5	2.289×10^{-65}
mean_d_P	0.49	9.52×10^{-64}
dist_Freq_tilt_x	0.5	2.91×10^{-65}
dist_Freq_speed	0.5	2.91×10^{-65}
median_Freq_tilt_y	0.501	1.31×10^{-65}
Space_Between_Words	1	0
bandwidth_tilt_x	0.501	1.43×10^{-65}
std_d_P	0.82	5×10^{-211}
mean_Pressure	1	0
in_Air	0.62	7.81×10^{-106}
BHK_raw_speed_score	1	0
std_Density	0.99	0
median_Freq_tremolo	0.5	1.51×10^{-65}
Age	1	0

Logistic Regression: It is a statistical method used to model the relationship between a binary outcome variable and one or more predictor variables. The goal of logistic regression is to estimate the probability of the binary outcome (e.g., 0 or 1) as a function of the predictor variables. The logistic regression model uses a logistic function to map the linear combination of the predictor variables to the probability of the outcome [19].

4.3.2. Selection Process

In order to select the best ML algorithm for our model, we evaluated and compared different well-known algorithms based on the Cross-Validation method. This method is one of the most widely used protocols for model evaluation, employing a statistical approach to assess and compare learning algorithms by dividing data into two segments: one for learning or training a model and the other used for validating it. In typical cross-validation, the training and validation sets must cross-over in successive rounds, ensuring that each data point has a chance of being validated against. For this project, we used 5K Cross-Validation.

To evaluate the quality of the results provided by the different algorithms, we considered accuracy, precision, recall and F1 metrics.

4.3.3. Hyper-parameters Optimization:

Hyperparameters are configuration settings that are established before the training process begins and cannot be learned from the data. These settings wield substantial influence over the model's performance and the quality of its

outcomes. The primary aim of hyperparameter optimization is to identify the most effective values for these hyperparameters, with the ultimate goal of enhancing prediction accuracy.

To achieve this, hyperparameter optimization typically involves assessing the model's performance on a validation dataset while experimenting with different combinations of hyperparameters. The combination that yields the highest performance is then selected. Various techniques can be employed for hyperparameter optimization, including grid search, random search, Bayesian optimization and genetic algorithms. In this work, we have used 2 techniques:

- Genetic algorithm: an evolutionary algorithm that begins with a randomly generated initial set of parameters (i.e chromosomes) of a potential solution (6 chromosomes in our case) (e.g.: {N of trees: 10, Criterion: log loss, Max Depth: 10, bootstrap: False, Max Features: None} is a chromosome). Then, a score is calculated based on these chromosomes using the cross-validation score. The 2 highests individuals are selected to pass their genes to the next generation, in which they randomly exchange 2 of their parameters values to create 2 new chromosomes (i.e. off-springs). These off-springs are subject to a mutation, where we randomly select for each one a parameter and change its value, to create 2 new chromosomes (i.e. Mutant chromosomes). This whole process is applied on these 6 (2 best chromosomes, 2 off-springs and 2 mutants) and goes on until satisfying a certain criterion (either achieving a pre-defined score or a duration).
- Random search: an optimization technique that explores different combinations of parameters values. The model is trained with these combinations and evaluates the performance of each one, repeating the process for a number of iterations. The algorithm returns the combination with the maximum score.

Table 3: Optimal Hyperparameters.

Model	Parameter	Optimal Value
Random Forest	N of trees	148
	Criterion	log loss
	Max depth	58
	Bootstrap	True
	Max Features	None
Logistic Regression	Solver	Saga
	Penalty	None
	C	1
Gaussian Naive Bayes	Priors	[0.9, 0.1]

We compute the hyperparameters for each algorithm using both Genetic algorithm and Random search separately, then we evaluate the results using the cross validation technique. Finally, we keep the hyperparameters that provide the best results with respect to the identification of Dysgraphic persons. The hyperparameters selected for each algorithm are presented in table 3.

4.3.4. Results:

In this work, our main goal is to detect Dysgraphic persons and to compute for each individual the degree of belonging to this class. In other terms, we need to maximize the true positives and minimize the false negatives. The metric which reflects this aspect is the recall. Thus, our model should maximize the recall metric to detect these individuals well.

Table 4: Cross-Validation results.

	Accuracy	Recall	Precision	F1 Score
Random Forest	0.9644	0.9000	0.9418	0.9177
Logistic Regression	0.9745	0.9027	0.9800	0.9350
Fuzzy C-Means	0.5210	0.6110	0.2580	0.3630
GaussianNB	0.9592	0.9500	0.8930	0.9140

The results presented in table 4 shows that Gaussian Naive Bayes outperforms the other models in terms of Recall. Furthermore, Gaussian Naive Bayes obtained good results in the other metric. As a consequence, we decided to use the Gaussian Naive Bayes model in the tracking phase.

Applied on our driven dataset, Gaussian Naive Bayes selected the following features in order to train the model: **BHK_raw_quality_score**, **BHK_raw_speed_score**, **median_Freq_speed**, **Space_Between_Words** and **Age**. Indeed, it includes in its process a feature selection step. The following figure shows scatter plots of these features for both groups (non-Dysgraphic & Dysgraphic).

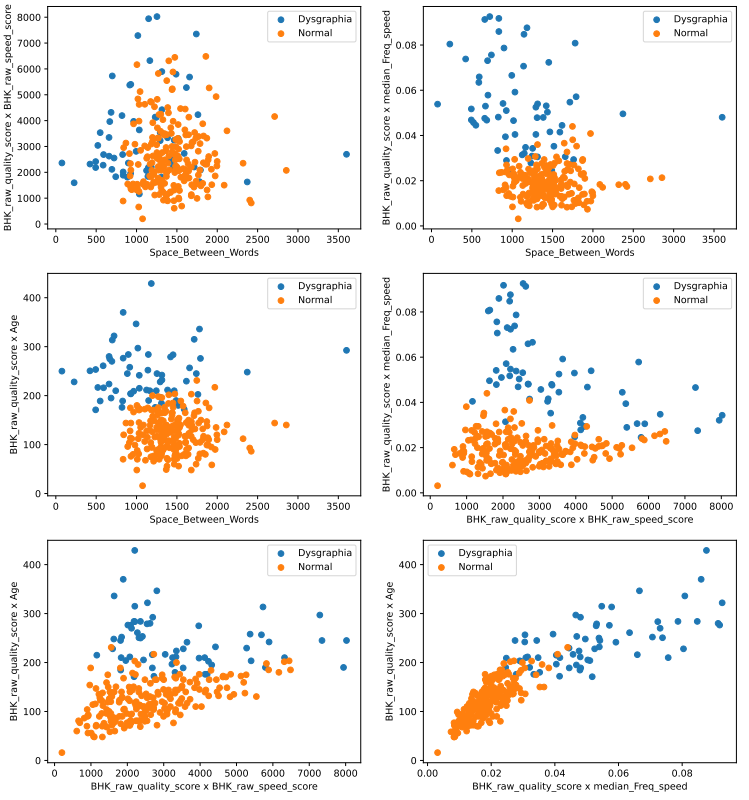


Fig. 2: Final selected features

4.4. Tracking Process

The object of this phase is to build the model used for the tracking of Dysgraphic persons, based on the results achieved during the first phase. Thus, the tracking is achieved using the Gaussian Naive Bayes algorithm applied on the features combination selected in section 4.2.

In order to track the evolution of Dysgraphic persons, our approach is applied on the data collected at each instant t to compute for each individual i the score $s(t, i)$ of belonging to the Dysgraphic class. Then the score $s(t, i)$ obtained at instant t is compared with the score $s(t-1, i)$ obtained at instant $t-1$. The evolution is defined based on this comparisons:

- positive evolution if $s(t, i) < s(t-1, i)$
- negative evolution if $s(t, i) > s(t-1, i)$
- no evolution if $s(t, i) = s(t-1, i)$

At each data collection, the process is repeated on the new data and the scores for each individual is compared with previous data which allow to draw the evolution curve per person.

5. Experimental Evaluation and Interpretation

This section presents our experiments to show the effectiveness of our approach for tracking the evolution of Dysgraphic persons. We have first evaluated the quality of the results achieved by our model. Then we have evaluated the capacity of our approach to track both positive and negative evolution of the individuals. The source code of the project is publicly available on GitHub [1].

5.1. Datasets

To evaluate the robustness and the capability of our model to generalize better on unseen data, we have created a synthetic dataset using Generative Adversarial Network for tabular datasets **TABGAN** proposed by [22] based on the driven dataset used in this work [16]. This approach ensures that the generated distributions match the distributions of the discrete variables in the training data.

First of all, we have trained a multi-layer perceptron on the original dataset, which achieved an AUC (Area under the ROC curve) of 0.802 on the test set. AUC metric is a widely used measure for assessing the performance of classifiers, particularly for imbalanced data [23]. Then, TABGAN model is used to generate a synthetic dataset from the original. The resulting generated dataset comprises 157 samples, 130 non-Dysgraphic and 27 Dysgraphic samples. To assess the quality of this newly generated dataset, we employed it as a test set for the previously implemented multi-layer perceptron, resulting in an AUC of 0.78.

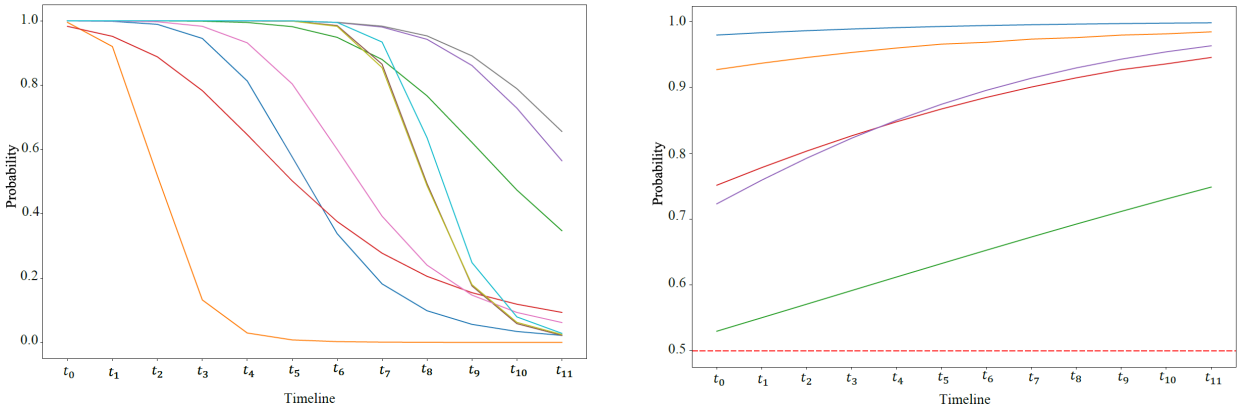
To the best of our knowledge, there are no datasets online containing information of Dysgraphic persons collected at different times. Thus, to evaluate the capacity of our approach to track Dysgraphic persons, we have simulated the evolution of the characteristics of Dysgraphic people. First, we generate a series of data showing positive evolution on which we apply our algorithm. Positive evolution implies that the person's health improves. Second, we evaluate the capability of our algorithm to track negative changes by generating a series of data reflecting such trends. The object is to show that our algorithm is able to understand, through the series of data, that an individual is making positive or negative evolution.

To do that, we have implemented the algorithm 1 to generate for each person, series of data from instant t_1 to instant t_n, knowing that instant t_0 is the original data of each Dysgraphic person. To generate series of data with positive evolution (negative evolution), we determine the closest non-Dysgraphic (Dysgraphic) person for each Dysgraphic (non-Dysgraphic) person **p** by computing the Euclidean distance between **p** and the each non-Dysgraphic (Dysgraphic) person **n** in **n_list** (line 4 – 5). Then, for each feature **f**, we compute the difference between **f** for the Dysgraphic person **p** and the closest person **closest_n** (line 6 – 9). Finally, the data for the next instant is calculated by the formula in line 8, which is inspired from weights update formula using the descent gradient, where α is the learning rate that can be either positive or negative (positive for improving the writing abilities, negative for the opposite).

Algorithm 1 Dysgraphic cases data simulation

Require: features, non-Dysgraphic_people_data, t_0 instant data
Ensure: data ▷ from t_1 to t_n instant data
1: $\alpha \leftarrow 0.1$ ▷ Evolution rate
2: $n\_list \leftarrow non - Dysgraphic\_people\_data$
3: **for** i in range(t_1, t_n) **do**
4: $distances \leftarrow \|(p, n)\|_2$ ▷ for each non-Dysgraphic person n in n_list
5: $closest\_n \leftarrow \text{argmin}_{n \in \mathbb{K}\_\ll \sqsupset \sim \approx}(distances)$
6: **for** f in features **do**
7: $diffs[f] \leftarrow closest\_n[f] - p[f]$
8: $data[t_i][f] \leftarrow data[t_{i-1}][f] + \alpha * diffs[f]$
9: **end for**
10: **end for**
11: **return** data

[1] https://github.com/youcef-ely/Dysgraphia_Detection/tree/main



(a) Positive Evolution of 10 Random Dysgraphic Children.

(b) Negative Evolution of 5 Random Dysgraphic Children

Fig. 3: Tracking the Evolution of Dysgraphic Children.

5.2. Robustness Evaluation

We have evaluated the robustness of our model to identify dysgrqphic persons by applying the Gaussian Naive Bayes algorithm on the dataset generated in the previous section. A person is identified as Dysgraphic if the probability to belong to this class is over 50%. The obtained results show that our model is able to identify Dysgraphic persons with good metric scores. Indeed, the accuracy, recall, precision and F1 score was respectively equal to 0.94, 0.85, 0.85 and 0.85.

5.3. Dysgraphic Person Tracking

We have first applied our model on the data showing positive evolution to compute the probability of being Dysgraphic for each person based on the collected data. Figure 3a shows the probability over 12 consecutive instants, of 10 persons selected randomly from the generated dataset.

We can see that our approach predicts positive evolution, as described by the generated data, where the probability of each person is decreasing from being Dysgraphic to being non-Dysgraphic. Indeed, children are showing positive evolution even though those having a probability above 0.5 at t_{11} are still Dysgraphic. The rest of persons reached the non-Dysgraphic class since they achieved a probability lower than 50% at one point.

In a second time, we have applied our model on the data showing negative evolution. Figure 3b shows that our approach is able to track negative evolution where the probability of each person to be Dysgraphic is increasing. This figure proves that our approach is able to predict negative evolution.

The results obtained from the different experiments indicate that our model is able to identify Dysgraphic persons, even when applied to new datasets, consistently achieving good scores. In addition, we have demonstrated the model's capability to assess the evolution of the disability, whether individuals experience positive or negative progress. In conclusion, our approach proves to be well-suited for tracking dysgraphia disease by analysing data collected from the patients at different points in time.

6. Conclusion

Learning disabilities such as dysgraphia present significant challenges in the lives of individuals, particularly children and students. Several studies have proposed methods for the early detection of dysgraphia. However, there is currently no approach for monitoring the progression of these individuals' conditions.

In this paper, we present an approach for the automatic tracking of a person's disease evolution. Our approach is based on a probabilistic ML algorithm that computes, for each individual, the score of belonging to Dysgraphic class. Computing this score at different intervals allows us to monitor the person's evolution. To achieve this, we compared various probabilistic and fuzzy ML algorithms and ultimately selected Gaussian Naive Bayes, which achieves the highest performances for the tracking task. Finally, we generated a series of data that simulates the evolution of

features for Dysgraphic individuals at 12 different points in time, encompassing both positive and negative evolutions. Our experimental evaluation revealed that our model accurately predicts both types of evolution and effectively tracks Dysgraphic persons.

In our future work, we aim to develop an approach that provides explanations for the results generated by our Dysgraphic persons' tracking method. Indeed, this explanatory aspect is intended to empower Dysgraphic individuals by offering advice and adaptations to enhance their treatment.

References

[1] R. I. Nicolson and A. J. Fawcett, "Dyslexia, dysgraphia, procedural learning and the cerebellum," *Cortex*, vol. 47, no. 1, pp. 117–127, 2011.

[2] L. Deschamps, L. Devillaine, C. Gaffet, R. Lambert, S. Aloui, J. Boutet, V. Brault, E. Labyt, and C. Jolly, "Development of a Pre-Diagnosis Tool Based on Machine Learning Algorithms on the BHK Test to Improve the Diagnosis of Dysgraphia," *Advances in Artificial Intelligence and Machine Learning*, vol. 1, pp. 114–135, Aug. 2021.

[3] G. Richard and M. Serrurier, "Dyslexia and dysgraphia prediction: A new machine learning approach," *arXiv preprint arXiv:2005.06401*, 2020.

[4] R. Kariyawasam, M. Nadeeshani, T. Hamid, I. Subasinghe, P. Samarasinghe, and P. Ratnayake, "Pubudu: Deep learning based screening and intervention of dyslexia, dysgraphia and dyscalculia," in *2019 14th Conference on Industrial and Information Systems (ICIIS)*, pp. 476–481, IEEE, 2019.

[5] R. Gupta, D. Mehrotra, R. Bouhamoum, M. Masmoudi, and H. B. Zghal, "Handwriting analysis ai-based system for assisting people with dysgraphia," in *Computational Science - ICCS 2023 - 23rd International Conference, Prague, Czech Republic, July 3-5, 2023, Proceedings, Part III*, vol. 14075 of *Lecture Notes in Computer Science*, pp. 185–199, Springer, 2023.

[6] L. Magalhaes, A. Cardoso, and C. Missiuna, "Activities and participation in children with developmental coordination disorder: A systematic review," *Research in developmental disabilities*, vol. 32, pp. 1309–16, 02 2011.

[7] B. Agarwal, S. Jain, P. Bansal, S. Shrivastava, and N. Mohan, "Dysgraphia detection using machine learning-based techniques: A survey," in *Emerging Trends in Expert Applications and Security* (V. S. Rathore, J. M. R. S. Tavares, V. Piuri, and B. Surendiran, eds.), (Singapore), pp. 315–328, Springer Nature Singapore, 2023.

[8] J. Mekyska, Z. Galaz, K. Safarova, V. Zvoncak, J. Mucha, Z. Smekal, A. Ondrackova, T. Urbanek, J. M. Havigerova, J. Bednarova, *et al.*, "Computerised assessment of graphomotor difficulties in a cohort of school-aged children," in *2019 11th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 1–6, IEEE, 2019.

[9] T. Asselborn, T. Gargot, Ł. Kidziński, W. Johal, D. Cohen, C. Jolly, and P. Dillenbourg, "Automated human-level diagnosis of dysgraphia using a consumer tablet," *NPJ digital medicine*, vol. 1, no. 1, pp. 1–9, 2018.

[10] A. Devi and G. Kavya, "Dysgraphia disorder forecasting and classification technique using intelligent deep learning approaches," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 120, p. 110647, 2023.

[11] M. Bublin, F. Werner, A. Kerschbaumer, G. Korak, S. Geyer, L. Rettinger, E. Schoenthaler, and M. Schmid-Kietreiber, "Automated dysgraphia detection by deep learning with sensogrip," 2023.

[12] J. Quenneville, "Tech tools for students with learning disabilities: Infusion into inclusive classrooms," *Preventing School Failure*, vol. 45, pp. 167–170, 01 2001.

[13] M. Diem, S. Fiel, A. Garz, M. Keglevic, F. Kleber, and R. Sablatnig, "ICDAR 2013 competition on handwritten digit recognition (HDRC 2013)," in *12th International Conference on Document Analysis and Recognition, ICDAR 2013, Washington, DC, USA, August 25-28, 2013*, pp. 1422–1427, IEEE Computer Society, 2013.

[14] K. Toutanova and R. C. Moore, "Pronunciation modeling for improved spelling correction," in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, July 6-12, 2002, Philadelphia, PA, USA*, pp. 144–151, ACL, 2002.

[15] T. Gargot, T. Asselborn, H. Pellerin, I. Zammouri, S. M. Anzalone, L. Casteran, W. Johal, P. Dillenbourg, D. Cohen, and C. Jolly, "Acquisition of handwriting in children with and without dysgraphia: A computational approach," *PLoS One*, vol. 15, no. 9, p. e0237575, 2020.

[16] T. Gargot, T. Asselborn, H. Pellerin, I. Zammouri, S. M. Anzalone, L. Casteran, W. Johal, P. Dillenbourg, D. Cohen, and C. Jolly, "Acquisition of handwriting in children with and without dysgraphia: A computational approach," *PLoS One*, vol. 15, no. 9, p. e0237575, 2020.

[17] K. Mabude, J.-C. Malela-Majika, P. Castagliola, and S. Shongwe, "Distribution-free mixed gwma-cusum and cusum-gwma mann–whitney charts to monitor unknown shifts in the process location," *Communications in Statistics-Simulation and Computation*, vol. 51, no. 11, pp. 6667–6690, 2022.

[18] L. M. Goyal, M. Mittal, and J. K. Sethi, "Fuzzy model generation using subtractive and fuzzy c-means clustering," *CSI transactions on ICT*, vol. 4, no. 2, pp. 129–133, 2016.

[19] M. Rybczak, N. Popowniak, and A. Lazarowska, "A survey of machine learning approaches for mobile robot control," *Robotics*, vol. 13, no. 1, 2024.

[20] S. Naiem, A. E. Khedr, A. M. Idrees, and M. I. Marie, "Enhancing the efficiency of gaussian naïve bayes machine learning classifier in the detection of ddos in cloud computing," *IEEE Access*, vol. 11, pp. 124597–124608, 2023.

[21] P. Mishra, C. M. Pandey, U. Singh, A. Gupta, C. Sahu, and A. Keshri, "Descriptive statistics and normality tests for statistical data," *Annals of cardiac anaesthesia*, vol. 22, no. 1, p. 67, 2019.

[22] I. Ashrapov, "Tabular gans for uneven distribution," *arXiv preprint arXiv:2010.00638*, 2020.

[23] T. Yang and Y. Ying, "Auc maximization in the era of big data and ai: A survey," *ACM Computing Surveys*, vol. 55, no. 8, pp. 1–37, 2022.