MANTECH Publications

Journal of Artificial Intelligence, Machine Learning and Soft Computing

Volume 3, Issue 1, January-June 2018

Diagnosis Children's with Dyslexia Using Machine Learning Technique

Dr. K. Vikram[1], Geeta Bhimrao Atkar[2]

Professor[1], Assistant Professor[2]

Department of Computer Science & Engineering,

G.H. Raisoni College of Engineering & Manaegemet, Wagholi, Pune

Email ID: *vikramkarimella@gmail.com*

Abstract

Worldwide, around 10% of the population has dyslexia, a specific learning disorder. Most of previous eye tracking experiments with people with and without dyslexia have found differences between populations suggesting that eye movements reflect the difficulties of individuals with dyslexia. In this paper, we present the first statistical model to predict readers with and without dyslexia using eye tracking measures. The model is trained and evaluated in a 10-fold cross experiment with a dataset composed of 1,135 readings of people with and without dyslexia that were recorded with an eye tracker. Our model, based on a Support Vector Ma- chine binary classifier, reaches 80.18% accuracy using the most informative features. To the best of our knowledge, this is the first time that eye tracking measures are used to predict automatically readers with dyslexia using machine learning.

Categories and Subject Descriptors K.4.2 [Computers and Society]: Social Issues—Assistive technologies for persons with disabilities; I.2.1 [Artificial Intelligence]: Applications and Expert Systems—Medicine and science.

Keywords: *Dyslexia, eye tracking, eye movements, diagnosis, detection, prediction, machine learning, support vector machine.*

INTRODUCTION

Dyslexia is the most common neurological learning dis- ability [24]. It affects from 10 to 17.5% of the population in the U.S.A. [23] and from 8.6 to 11% in Spain [6, 25] and has a considerable presence in web text [1]. Competitive reading and writing is required in our education system; therefore school failure is associated with dyslexia, even if dyslexia is not related to overall intelligence [24].Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, Dyslexia detection is crucial. When diagnosed, dyslexia can be treated avoiding its consequences such as high rates of academic failure. At the same time, diagnosing dyslexia is not a trivial task; it is expensive and it normally requires an expert. Also, dyslexia manifestations vary depending on the language. In fact, dyslexia is called a hidden disability due to the difficulty of its diagnosis in languages with shallow orthographies [72], as Spanish. Previous eye tracking studies with people with dyslexia from psychology research have concluded that the eye movements of people with dyslexia are not the cause but the reflection of the difficulties they have while reading [21, 45, 49]. Although there are number of studies that presents how eye tracking measures show individual differences [12, 37, 40], most of the studies agree in finding significant differences among readers with and without dyslexia. Human-computer interaction studies that use eye tracking with people with dyslexia have normally focused in finding the most accessible text presentations [55, 57]. Again, differences between people with and without dyslexia were found.

Even if eye tracking measures have already been used to predict where people tend to look [26] or to improve the interface design of search engines [17], among others; we believe this is the first effort to automatically detect dyslexia using eye tracking measures.In this paper we present a statistical model to classify readers with and without dyslexia using a Support Vec- tor Machine binary classifier. We trained the model with a dataset composed of 1,135 readings of Spanish speakers with and without dyslexia from 11 to 54 years old recorded with an eye tracker.

The contribution of this paper is:
– a statistical model to classify Spanish readers with and without dyslexia that achieves 80.18% accuracy in a 10-fold cross validation experiment.1

Next section focuses on dyslexia and its detection. Section 3 explains related work that have shown differences between populations using eye tracking. We present the dataset used in Section 4 and show the experiments with our statistical model in Section 5. We discussed the results in Section 6. Finally, we draw conclusions and future work in Section 7.

DYSLEXIA DETECTION

Dyslexia is defined as a specific learning disability with neurological origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result as a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities. Sec- ondary consequences may include problems in reading com- prehension and reduced reading experience that can impede growth of vocabulary and background knowledge [24, 30].

Dyslexia detection is crucial since academic failure is as- sociated with dyslexia when it is not detected and treated accordingly [14]. Diagnosing dyslexia and early detection of risk of dyslexia have been addressed from different fields, especially in psychology and neuroimaging.

In psychology, traditional paper based diagnosed of dyslexia such as TALE [70] for Spanish or Diagnostischer Rechtschreibtest [18] for German, analyze both reading and writing skills. Diagnoses of dyslexia are confirmed when the reading and the spelling performance of the child is signif- icantly under the level expected due to her or his age and general intelligence.2
More recently, Lyytinen et al. [32] created the computer game Literate, later called GraphoGame [31],3 which was de- veloped to identify children at risk of having dyslexia before school age in Finland. Its exercises are aimed towards the connection of graphemes (letters) and phonemes (sounds). They conducted two user studies with 12 and 41 children between 6 and 7 years old with very promising results. Al- though these

exercises were conceived as preventive train- ing, children who used Literate improved their accuracy in grapheme-phoneme connections, reading words, and naming phonemes after playing for less than 4 hours.

On the other hand, neuroimaging with children with dyslexia has revealed relationships between brain responses at infancy and later reading performance. Molfese [36] re- ported that there are brain responses (event-related brain potentials) to speech sounds within 36 hours of birth that can be used to discriminate children who would become readers with dyslexia with 8 years old. The accuracy of this prediction is 81%.

EYE MOVEMENTS AND DYSLEXIA

Related work on diagnose and early detection of risk of dyslexia is very extensive and comes from different fields such as cognitive neuroscience, psychology or biology. In this related work we only focus on how eye tracking measures have been studied in relationship with dyslexia. These studies come from psychology and human-computer interaction research.

Psychology Research

Rayner [50] presents a review of the studies from the mid 70's to the 90's that have used eye movements to investigate cognitive processes. He demonstrates that eye movement measures can be used to infer moment-to-moment cognitive processes in reading. For instance, shorter fixations are associated with better readability while longer fixations can indicate that processing loads are greater. As a matter of fact, non impaired readers present longer fixations at low-frequency words than at high-frequency words [22, 27, 46, 51, 54, 59].

The eye movements of readers with dyslexia are different from regular readers. People with dyslexia as well as beginner readers, make longer fixations, more fixations, shorter saccades4 and more regressions than readers without dyslexia [11, 35, 50].

During the 80's-90's it was discussed to which extent eye movements are the cause of reading problems. If eye movements were a causative factor, then dyslexia could easily be diagnosed with a simple eye movement test. There have been some studies concerning whether eye movements are the cause of dyslexia: erratic eye movements [41, 42, 43], in-stability during fixation [11], and selective attentional deficit [13, 15, 44]. However, the results presented in those experiments could not be later replicated by others (see [3, 5, 38, 39, 65, 66] for erratic eye movements; Raymond et al. [47] for instability during fixation; and [16, 28, 52, 53, 64] for selective attentional deficit). At the same time, Tinker [67, 68] and Rayner [48, 50] studies show how eye movements are generally not a cause of the reading disability but were a reflection of other underlying problems. More concretely, there are three studies that are consistent with the conclusion that eye movements reflect the difficulties that individuals with dyslexia have while reading and that are not the cause of the reading problem.

First, Hy¨ona and Olson [21] found that readers with dyslexia show the typical word frequency effect in which low-frequency words are fixated longer (fixation duration, number of fixations, and regressions) than high-frequency words.

Second, Pirozzolo and Rayner [45] and Olson et al. [39] found that when people with dyslexia were given a text appropriate for their reading level, their eye movements (fix- ations, saccades, and regressions) were much like those of regular readers at that particular age level.

Third, Rayner [49] showed that regular children’s eye movements (fixation durations, saccade lengths, and the size of the perceptual span) shared the characteristics of read- ers with dyslexia when they were given a text that was too difficult for them.

Taking into account all these studies the evidence sug- gests that the vast majority of people with dyslexia have a language processing deficit and, that their eye movements simply reflect their difficulty processing language [50].

Later literature on eye tracking and dyslexia have focused on the effect of other factors, such as regular orthography [20], orthographic neighbors [33], or letter length and phoneme length [34].

Human-Computer Interaction

In her PhD thesis, Rello [55], performed twelve eye track- ing experiments –and eleven sub experiments– to explore the impact of text presentation and text content on the readabil- ity and comprehension of people with dyslexia.

The text presentation parameters tested were typeface, font size, colors, character spacing, line spacing, paragraph spacing, different grey scales for text, and column width [61]. representations [60], verbal paraphrases, graphical schemes, keywords, and different strategies of lexical simplification [58]. Each experiment had between 23 to 48 participants with dyslexia plus a control group. In all the experiments, except from one –where the participants had to read texts with orthographic errors [56]– there were found significant differences between participants with and without dyslexia regarding eye tracking measures.

What is missing?

Previous literature on eye movements and dyslexia have found pattens, correlations, individual differences, and sig- nificant differences between populations. However, to the best of our knowledge, none of them have applied machine learning to classify people with and without dyslexia, that is, to detect readers with dyslexia.

What is missing is the connection of two points that were not previously joined: eye tracking measures and machine learning techniques to predict readers with dyslexia automatically.

DATASET

We used a dataset derived from an eye tracking experiment with 97 subjects with normal or corrected-to-normal vision; 48 of then with diagnosed dyslexia. The participants with dyslexia (22 female, 26 male) presented a confirmed diagnosis of dyslexia. Their ages ranged from 11 to 50 ($\bar{x}$ = 20.96, s = 9.98). Except from 3 participants, all of the participants were attending school or high school (26 partic- ipants), or they were studying or had already finished uni- versity degrees (19 participants). The group of participants without dyslexia was composed of 49 people (28 female, 21 male). Their ages ranged from 11 to 54 ($\bar{x}$ = 29.30, s = 9.03). Except from 5 participants, the rest were either attending or had finished school or high school (17 participants) or university (27 participants).

This dataset was derived from a within-subject design ex- periment. Each participant read 12 different texts with 12 different typefaces. The texts and the fonts were counter- balanced to avoid sequence effects. Therefore, the data with respect to text-font combinations was evenly distributed.

The twelve fonts were: Arial, Arial Italic, Times and Times Italic – the most common fonts used on screen and printed texts, respectively [8]–; OpenDyslexic and

OpenDyslexic Italic – designed specifically for people with dyslexia–5; Verdana, recommended by the British Dyslexia Association [4]; Courier – the most common example of monospaced font [8]–; Helvetica and Myriad – broadly used in graphic design and typeface of choice of Microsoft and Apple, respectively –; Garamond – for its strong legibility for printed materials [8]– and CMU –widely used in scien- tific publishing, as is the default of the typesetting program TeX, as well as a free typeface supporting many languages [29].

The readings of each text were recorded using eye track- ing, the user preferences towards the fonts were gathered using questionnaires with five-point Likert scales. Compre- hension questions were presented at the end of each text as a control variable.

The text used in the experiments met comparability re- quirements. They were extracted from the same book, Im- postores (‘Impostors’),6 by Lucas S´anchez [62]. They all had the same genre and same style; the same number of words (60 words); similar word length, with an average length ranging from 4.92 to 5.87 letters; absence of numerical ex- pressions, acronyms, and foreign words, because people with dyslexia especially encounter problems with such words [9]. An example of a text read by the participants is given in the Appendix.

The text presentation was also controlled, except from the typeface. All the texts were left-justified, using a 14 points font size, and the column width did not exceeded 70 characters/column, as recommended by the British Dyslexia Association [4]. The color used was the most frequently used in the Web for text: black text on white background.

The equipment used was the eye tracker Tobii 1750 [69], which has a 17-inch TFT monitor with a resolution of 1,024×768 pixels. The time measurements of the eye tracker have a precision of 0.02 seconds. Hence, all time values are presented in the dataset with an accuracy of two decimals.

The eye tracker was calibrated individually for each partic- ipant and the light focus was always in the same position. The distance between the participant and the eye tracker was constant (approximately 60 cm. or 24 in.) and con- trolled by using a fixed chair.

For more details about the experimental design on how these readings were collected, please refer to [55].

Therefore, our dataset is composed of readings marked as D if the participant has dyslexia and N if not, there are 12 readings per participant, that is 1,164 readings; 29 of these readings were not properly recorded with not a number values. Hence, we removed those readings from the dataset having a final dataset containing 1,135 readings. From the dataset we extracted the following features:

- Age of the participant: ranging from 11 to 54 years old.

- Typeface: One of the 12 typefaces used for the text.

- Italic: This is a binary feature with two values, italic when the text had an italic type and roman when the text had a roman type.

- Serif: This is a binary feature with two values, sans serif when the font of the text had an typeface with- out serif -Arial, Helvetica, Myriad, and Verdana–, and serif when the text had typefaces with serif –CMU, ***Garamond, and Times***

- Typeface designed for dyslexia: A binary feature that shows when the font in the text had a typeface specifically designed for people with dyslexia.

- Typeface preference of the participant: Value given to a typeface by the participant using a five-point Likert scale.

- Number of visits: Total number of visits to the area of interest.

- Mean of visit: Duration of each individual visit within the area of interest (the text).
- Sum of visits (in the following, reading time): Sum of all the visits. This is equivalent to the reading time of the whole text.

- Mean of fixation: When reading a text, the eye does not move contiguously over text, but alternates sac- cades and visual fixations, that is, jumps in short steps and rests over pieces of text. It denotes how long the eye rests still on a single spot of the text.

- Number of fixations: Total number of fixations while reading a text per visit.

- Sum of fixations: Sum of all fixations.

Some of the features have numeric (real or integer) values, so we established some ranges for each of them to discretize the data. For instance, the age of the participants is divided in 3 different groups: (1) younger than 14 years old, (2) from 14 to 19 years old, and (3) from 20 to 54 years old.

EXPERIMENTS

In order to find out whether it is feasible to detect read- ings of users with dyslexia, we set up a machine learning ex- periment. Machine learning is the scientific discipline that studies algorithms that can learn from data and make pre- dictions. The output of a machine learning algorithm is called a model which is capable of making predictions given unseen data (normally for evaluation).

We carried out an experiment with a binary classifier of LIBSVM [7] in the polynomial Support Vector Machine (SVM) set-up. An SVM is a method for supervised learning that analyzes data and recognize patterns for classification. Given a set of training examples, each marked as belonging to a category, an SVM training algorithm builds a model that assigns new examples into the categories. It represents the examples as points in space and classifies them accord- ing to hyperplanes. When there is an input for the classifier it tries to assign a hyperplane to the input and then this is the classification output. Our SVM is trained on datasets as the one described in Section 4, and it is able to perform predictions on new readings.

We performed a 10-fold cross validation experiment by di- viding the data in 10 different roughly equal subsets (10% of the data in each subset). Then we trained a statistical model on the rest of the data (90%) and tested on the cor- responding fold by iterating 10 times, at the end we had all the data tested independently. We randomized the data and we did stratified sampling to ensure a similar distribution of data in all folds. We also kept all readings by the same user in the same fold, meaning that we had, in each fold, a similar number of readings marked as participants with and without dyslexia, and that a user does not serve for training a model that will predict readings of the same user. The idea is to see how the results of the statistical analysis will generalize on an independent dataset, in our case: new readings. We used 10-fold cross validation because it is normally recommended for smaller datasets when a single train-development test split might not be informative enough [2].

Table 1 shows the accuracy of the SVM models for each of the folds. This result suggests that the model is able to predict readings of users with dyslexia quite accurately with a final result of 80.18% by using the most informative features (reading time, mean of fixations and age of the par- ticipants),7 meaning that the statistical models are able to make a correct prediction in 910 of the 1,135 readings.

Table 1: Accuracy of the classifiers in the 10-fold cross validation experiment

Dataset	Accuracy
Fold-1	83.62% (97/116)
Fold-2	96.26% (103/107)
Fold-3	69.90% (72/103)
Fold-4	89.74% (105/117)
Fold-5	86.48% (96/111)
Fold-6	73.15% (79/108)
Fold-7	61.21% (71/116)
Fold-8	82.41% (89/108)
Fold-9	85.47% (100/117)
Fold-10	74.24% (98/132)
All	80.18% (910/1,135)

The features that we found useful for classification were: (1) reading time, (2) mean of fixations, and (3) age of the participant. Some features that were useful standing alone, such as (1) number of visits or (2) number of fixations were not useful when they are used jointly with the features listed above, due to redundancy as they express very similar infor- mation. Other features, such as typeface, italic or serif do not affect at all in the predictions, which is also expected, and good news, because they are independent from the par- ticipant. This means that we are able of detecting users with dyslexia without taking into account the typeface.

DISCUSSION

First, it is worth noting that the age of the participants range from 11 to 54 years old, and the users with dyslexia tend to improve their reading skills with age. In order to test whether this was affecting the final result we run the same experiment (with the same folds), as in Section 5, by removing the age of the participant as a feature. Table 2 shows the results of the SVM models without considering the age of the participants. The final result is 76.38 of final accuracy (losing 3.8 points). This indicates that the age of the users shows clearer differences in their reading perfor- mance. Nonetheless, in the dataset, the age average of the participants with dyslexia is 20.96, with a standard deviation of 9.98 while the age average of the participants without dyslexia is 29.20 with a standard deviation of 9.03 [55]. If the ages of both groups were perfectly matched, we could expect to have more homogenous results between folds. Nonethe- less, we had more participants without dyslexia without higher education (5) than participants with dyslexia (3) [55].

Fleshing out a bit more the results, we also observe that some folds achieved higher results than others, being 96.26% the highest and 61.21% the lowest, even though we per- formed stratified sampling we encountered this outcome. This means that some of the readings are difficult to pre- dict. For instance, a participant with dyslexia who is 50 years old and might have already overcome most of its reading issues, would be ideally classified as a participant with dyslexia, however our model fails and classifies it as a participant without dyslexia. We believe that the main problem is the size of the dataset, since having more input data would lead our model to generalize better in those cases. There- fore, in order to improve our results, we plan to increase the dataset by carrying out more eye tracking experiments. Moreover, the age of the participants is also playing a role in this issue, since we have participants that range from 11 to 54 years old, we also believe that having different datasets depending on the age of the participants will also lead to a proper classification of the readings, since our model is gen- eral and a first attempt trying to predict dyslexia based on eye tracking measures.

Table 2: Accuracy of the classifiers in the 10-fold cross validation experiment without considering the age of the participant as a feature

Dataset	Accuracy
Fold-1	83.62% (97/116)
Fold-2	85.98% (92/107)
Fold-3	65.05% (67/103)
Fold-4	84.62% (99/117)
Fold-5	74.77% (83/111)
Fold-6	72.22% (78/108)
Fold-7	56.03% (65/116)
Fold-8	82.41% (89/108)
Fold-9	85.47% (100/117)
Fold-10	73.48% (97/132)
All	76.39% (867/1,135)

As we mention above, the features related with the type- face were not useful for classification. Even that previous work has shown that typeface have a significant impact on the readability of people with and without dyslexia, this fact does not have an impact to differentiate their readings. The fonts that improve the readability of people without dyslexia are also beneficial for the readability of people with dyslexia [55]. This can explain why typeface is not an informative feature for classification, similarly of the font related features, such as, serif, italic, and fonts designed for people with dyslexia.

Even if the results of this experiment are only valid for Spanish language, previous literature on eye tracking has found significant differences on eye movements for languages with deeper orthographies such as English [21], German [20] or Bulgarian [19] as well as it was reported for Spanish [35]. Therefore we believe that dyslexia prediction in other languages using eye tracking measures is feasible, especially in languages with shallow orthographies where reading speed is a strong indicator for diagnosing dyslexia, such as Spanish [63] or Italian [71].

CONCLUSIONS AND FUTURE WORK

The eye movements of readers with dyslexia are different from regular readers. People with dyslexia have longer reading times, make longer fixations, and make more fixations than readers without dyslexia. In this paper we have shown how these characteristics can be used to train a machine learning model. We have presented a method that predicts readings of people with dyslexia with 80.18% of accuracy. To the best of our knowledge, this is the first attempt to build a statistical model to predict automatically readers with dyslexia using eye tracking measures. The model uses common eye tracking features such as reading time, mean of fixation, as well as the age of the participant. However, note that these are only seminal results on this topic and that the dataset was intended for other purposes.

Dyslexia is called a hidden disability because it is hard to diagnose. Diagnosing dyslexia is crucial to address this condition but estimations of dyslexia are much higher than the actual diagnosed population. Our model is just the first attempt on predicting dyslexia by using a machine learning approach with eye tracking measures. However, we believe it has a great potential impact.

This study suggests that eye tracking measures have the potential to be used to diagnose dyslexia in the future. Eye trackers are becoming more and more affordable and reading a text in silence is less intrusive than being exposed to the tests needed in current diagnoses. Moreover, there are other human-computer interaction measures that were found to be related to eye tracking measures, such as mouse tracking, that are yet to be explored. Their potential applications go from diagnosing to user modelling via their interaction with the computer.

For future work we plan to enlarge the dataset by carrying out more eye tracking experiments in different languages. Then, we also plan to try with other kind of classifiers, such as perceptron learning, recursive neural networks, or conditional random fields. Currently, we are exploring other human-computer interaction measures to detect dyslexia in collaboration with schools.

ACKNOWLEDGMENTS

This paper was partially developed under a grant from the Department of Education, NIDRR grant number H133A130057.

REFERENCES

1. R. Baeza-Yates and L. Rello. Estimating dyslexia in the Web. In Proc. W4A ’11), Hyderabad, India, 2011. ACM Press.

2. M. Ballesteros and J. Nivre. Maltoptimizer: A system for maltparser optimization. In Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC’12). European Language Resources Association (ELRA), may 2012.

3. J. L. Black, D. W. K. Collins, J. N. De Roach, and

4. S. Zubrick. A detailed study of sequential saccadic eye movements for normal-and poor-reading children. Perceptual and motor skills, 59(2):423–434, 1984.

5. British Dyslexia Association. Dyslexia style guide, January 2012. http://www.bdadyslexia.org.uk/.

6. B. Brown, G. Haegerstrom-Portnoy, C. Yingling,

7. J. Herron, D. Galin, and M. Marcus. Tracking eye movements are normal in dyslexic children. American journal of optometry and physiological optics, 60(5):376, 1983.

Authors' Profile

Dr. K. Vikram working is a professor in G.H. Raisoni College of Engineering & Management. He published many research papers and published two text books like Digital Image processing and Computer organization and editorial board member of many international journals.

Ms Geeta Bhimrao Atkar Asst.professor,G..H.Raisoni College of Engineering and Management,Wagholi,department of Computer Engineering Total work experience is 7.8 years. Pursuing PhD in Computer Science and Engineering from Vellore Institute Technology,Chennai Completed Mtech in Computer Engineering from PCCOE,Pune Total paper published 8 in international journals.