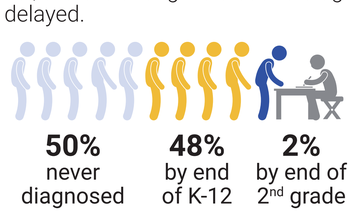
Can We (and *Should We*) Use AI to Detect Dyslexia in Children's Handwriting?

Katie Spoon[1,2], David Crandall[1], Katie Siek[1], Marlyssa Fillmore[3]
[1]Indiana University Bloomington, [2]IBM Research, [3]The Academic Achievement Center

Introduction

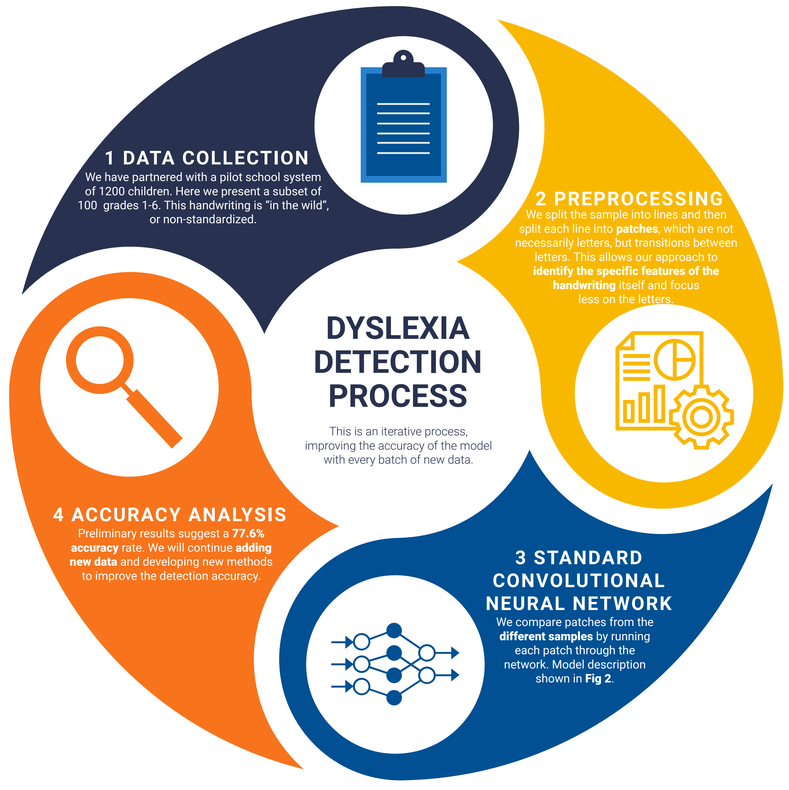
10.2 million
kids in the US with dyslexia will not be diagnosed by the recommended age[2]

- **Literacy** is the best predictor of success later in life. **Dyslexia** is a **learning disability (LD)** characterized by a difficult to read or interpret words, letters, and symbols according to their shapes[1].
- Dyslexia is **not tied to IQ**, which is a common myth. Students with dyslexia perform just as well with **accomodations**. To get these accomodations, students meet with a school psychologist. However, there is a long waitlist and diagnosis is often delayed.



- If students are placed in the waitlist by the end of 2nd grade, they improve their chances of graduating from high school[2]. However, **teachers usually don't have the training** to detect LDs[3] and often detection comes too late.

Our Approach



Goal: Create a deep learning-based early screening tool to **identify characteristics of dyslexic handwriting** and place students into the diagnostic queue by the end of 2nd grade

We Can; Should We?

Quantifying "Messiness"

- One of the authors hand-labeled patches of handwriting—5100 of them—in various categories of messiness: (1) illegible (cannot make out the letter or transition), (2) partially illegible (can make out a letter but not sure what it is), and (3) legible (can clearly tell the letter or transition).
- We re-validated the network and found that 84% of the patches marked illegible and 60% of the patches marked partially illegible were from samples of children with dyslexia. This seems to confirm that **students with dyslexia have objectively "messier" handwriting than their peers**, and the network is using "messiness".

Dangers of this Technology

- **False negatives:** While a false positive would refer a student for testing who may not need it, a false negative has more dire consequences.
- **Use outside of intended use**: What is to stop someone from doing the same thing to try to distinguish varying levels of intelligence in children? It is difficult to know the scope of use that will follow, and even though this project intends to help, there is always future potential for someone else to do harm.
- **Bias**: Some of the handwriting we receive labeled as not characteristic of dyslexia could be from a child with an *undiagnosed* reading or writing problem. Additionally, we need to investigate intersectional breakdowns of accuracy by gender, race, SES, etc.

Reading and Writing are Linked

Students who struggle to read typically also struggle to write. Students can be good readers with poor writing (dysgraphia), but typically not vice versa.

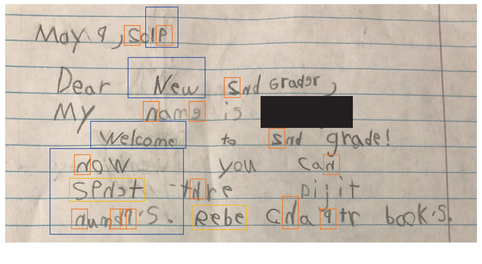


Fig 1. **Standard dyslexia evaluations depend partially on writng.** School pyschologists use many evaluation metrics (perhaps the most widely used method is the Barton method), including **letter confusion**, **misspellings**, and **self-correction**, among others (captilization, punctuation, content, pencil grip, etc.).

Neural Network Architecture & Potential Feature Identification

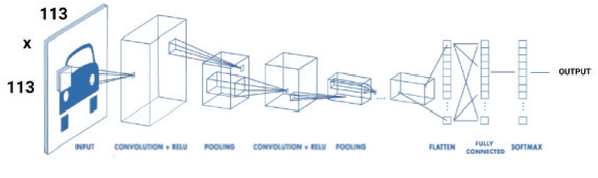


Fig 2. **Our model.** This model[4] was used for a different problem. It takes in **a patch of handwriting**. It has 5 convolutional layers, 3 max-pooling layers, and 2 fully-connected layers. Additional variants studied in our previous work include siamese networks (two different streams take patches from students with and without dyslexia) and the addition of a bayesian classifier to the output. We omit these variants for this work as accuracy improvements were not the focus here.

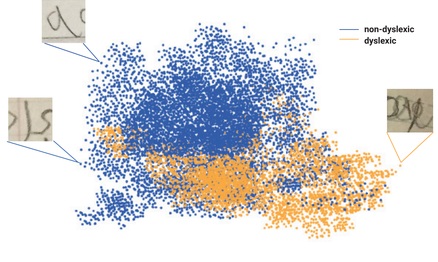


Fig 3. **Dyslexic handwriting tends to be messier.** TSNE is an ML **visualization technique** that maps a high-dimensional feature vector (in our case, from the last layer of a neural network) into 2D space. **Points closer together are more similar**. After plotting a sample of patches from 2nd grade, it appears that while there is a lot of overlap, the dyslexic handwriting tends to be messier or harder to identify.

Future Work

- We will improve the model for detecting dyslexia after we **utilize additional data**, build a **front-end application** for teachers and parents, and employ data visualization techniques to better understand which features the network uses to make decisions.
- We would also like to eventually **expand the system** to detect different types of LDs and to monitor English Language Learners, as the tool is not language-specific.

Contact
Katie Spoon: kspoon@iu.edu

1. Learning Disabilities Association of America (LDA)
2. National Center for Learning Disabilities (NCLD)
3. National Council on Teacher Quality (NCTQ)
4. Priya Dwivedi, English Deep Writer, (2018), Github repository, https://github.com/priya-dwivedi/Deep-Learning/tree/master/handwriting_recognition
5. Hernandez, D. J. (2012). Double jeopardy: How third-grade reading skills and poverty influence high school graduation. Baltimore, MD: The Annie E. Casey Foundation.
6. British Dyslexia Association