

# NeuroLit:

## A Visualization & Prediction Tool for Reading Research

[Authors]:

Maggie Clarke, Speech & Hearing Sciences

Patrick Donnelly, Speech & Hearing Sciences

Sritam Kethireddy, Applied Mathematics

# 80%

of children with learning disabilities have a  
reading disability

Reading research aims to:

1. Find struggling readers early
2. Individualize treatment
3. Learn how we can tailor a faulty educational system to work for all students





# How can NeuroLit help?

Incorporating the power of Data Science

- Building models
- Exploring data in methodical ways
- Looking for hidden relationships





# The Data

Two datasets:

1. Behavioral measures of reading skill, standardized
2. Survey responses, mostly binary + 1 Likert scale rating of reading ability


\*\* All data used is de-identified and available by permission of the research participants in accordance with the UW IRB \*\*



# The Interface

## NeuroLit Github Repository



 This repository Search

Pull requestsIssuesMarketplaceExplore

Unwatch 7Star 0Fork 0

CodeIssues 1Pull requests 0Projects 0WikiInsightsSettings

uwseeds-group-neurolit created by GitHub Classroom

Add topics

92 commits1 branch0 releases5 contributorsMIT

Branch: masterNew pull request

Create new fileUpload filesFind fileClone or download

Patrick Donnelly added examples folder with some sample iterations of the project note... Latest commit 8e0ad56 2 hours ago

Docs	Update data.md	15 hours ago
Images	added cover image	a month ago
examples	added examples folder with some sample iterations of the project note...	2 hours ago
neurolit	updated scikt-learn module to model-selection and changed data pull t...	3 hours ago
LICENSE	Initial commit	2 months ago
README.md	Update README.md	a day ago
dash_visualize.ipynb	updated scikt-learn module to model-selection and changed data pull t...	3 hours ago
predict.ipynb	updated scikt-learn module to model-selection and changed data pull t...	3 hours ago
setup.py	new setup.py file	9 days ago

README.md

NeuroLit

A Visualization and Prediction Tool for Reading Research

Our group aims to implement machine learning to examine the relationship between qualitative survey responses given by parents/guardians and the quantitative reading skill of their children. Data from the [UW Reading & Dyslexia Research Program](#) (UW RDRP) will be used.

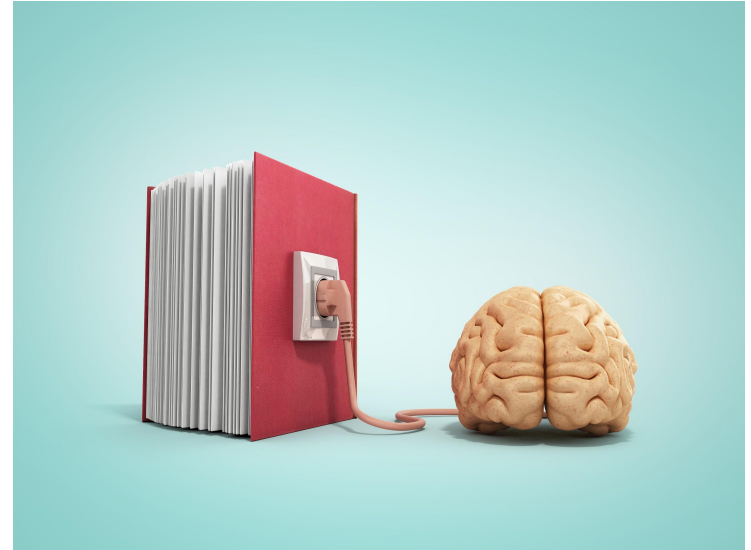
# Use Cases

## Data Visualization:

1. Can we create a tool that will produce fast, easy, and presentation-worthy visualizations of relationships in our data?
2. Using our survey data, do parent perceptions of reading difficulty correlate with assessed reading skill?

## Model-building:

3. Is our reading assessment able to predict the diagnosis of reading disability?



# NeuroLit Visualization Tool

UW Reading and Dyslexia Research Program.

Perceived Reading Skill

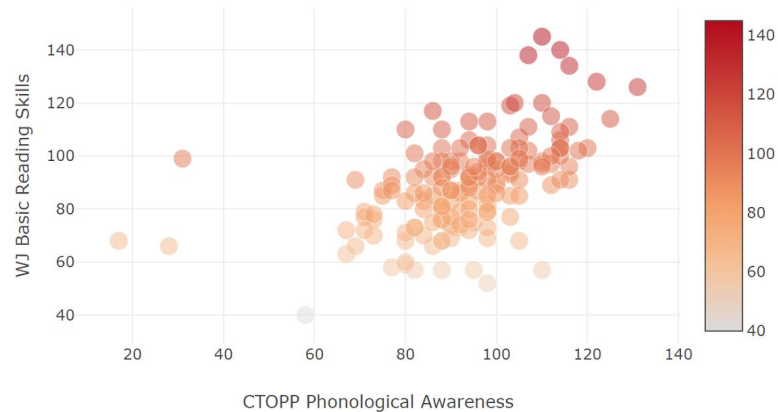
× ▼

CTOPP Phonological Awareness

× ▼

WJ Basic Reading Skills

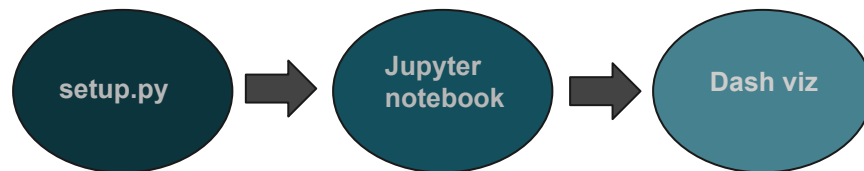
× ▼



# NeuroLit Model Building

User: UW researchers/students interested in using reading data to predict dyslexia diagnosis

- Pull RedCap data
- Select variables of interest
- Visualize relationships
- Use machine learning to predict outcomes
- Visualize outcomes



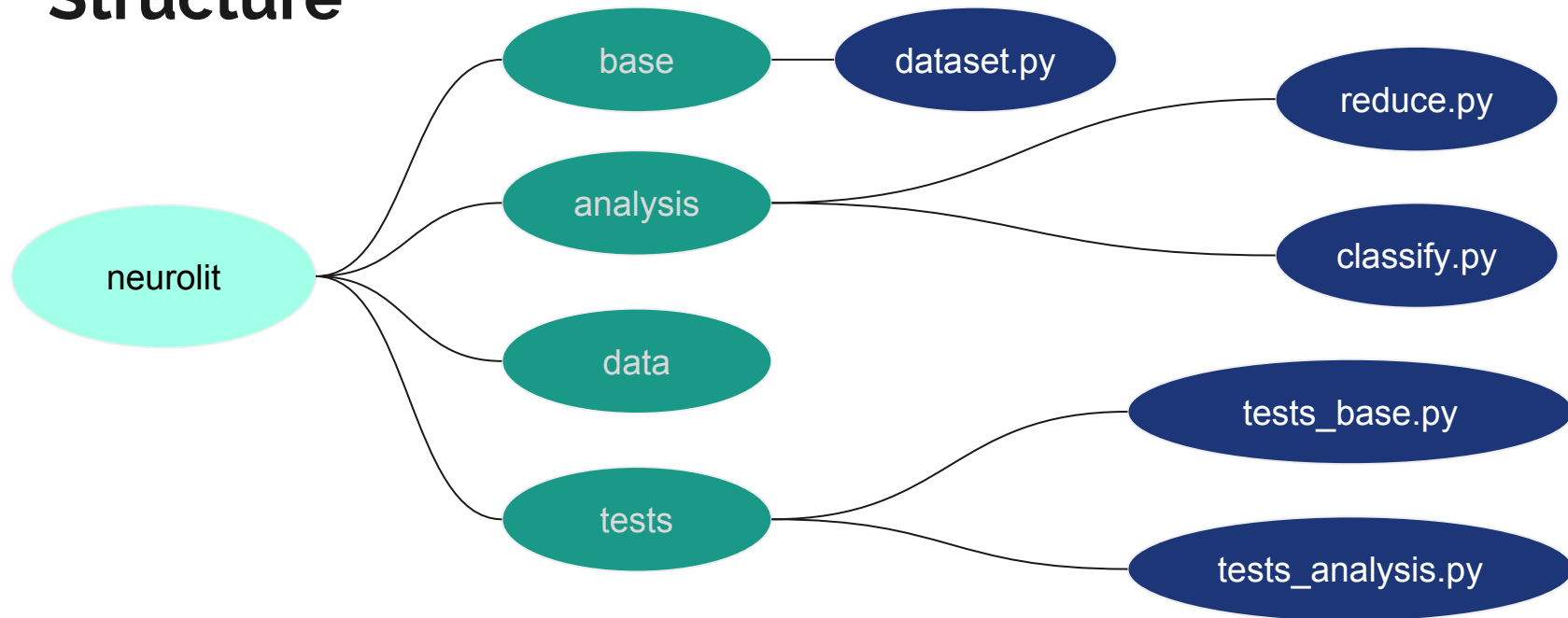
## Layout

- Setup.py
- Ipython notebook
  - Imports scripts to call functions
  - Visualization script





# Structure





# Design

## Curating Data


- Data retrieval
- Dataset object
- Variable selection
- Missing data imputation
- Data normalization

## Data Analysis

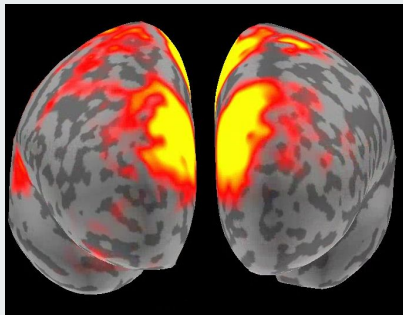
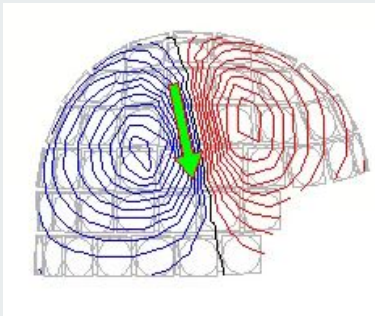
- Dimensionality reduction
- Classification
- Regression
- Clustering

## Visualization

- Variable relationships
- Model parameters
- Model performance
- Raw data distributions



# Lessons Learned/Future Directions

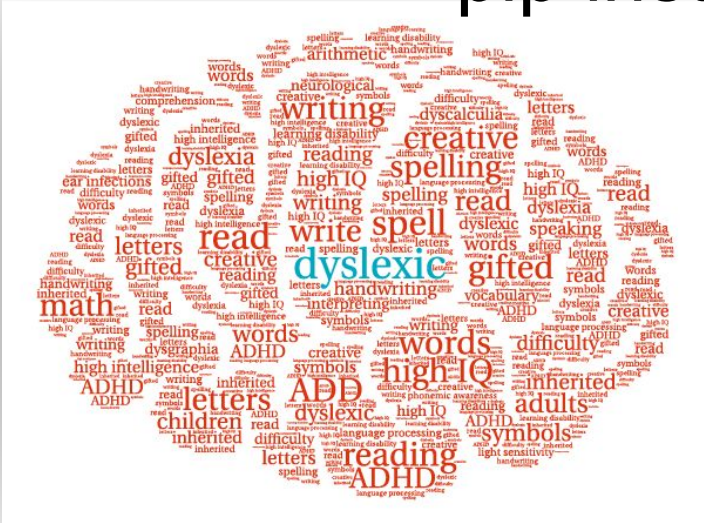


Lessons:

- Time is limited
- Start simple and build on it

Looking to future:

- Visualize machine learning outcomes
- Incorporate functional brain-imaging data
- Generalize to other kinds of categorical data

[illegible][illegible]