Our problem statement is to address a real-world problem under the hood of special education, the main idea is to identify learning disabilities in children using informal educational assessment tools. A child is said to have learning disabilities or difficulties when he or she is unable to perform on par with their peers and faces some sort of difficulty in displaying certain skills like reading, writing, mathematics, spellings and in processing certain information like sequencing, memory, auditory or visual discrimination, comprehension. There is a need of a standard mechanism which quantifies these disabilities, that’s where the Information Educational assessment tools come in, these include tests like Schonell’s Spelling Test, Burt’s Reading Test, Wepman’s Auditory Discimination Test, MAPS etc.

We decided on a 2-phase approach to address this problem by means of Machine and Deep learning techniques.

Phase 1: The first phase includes research and study of the different tools and occurrence of LD

1. We are initially going to perform PCA, for dimensionality reduction yet containing most information.
2. The occurrence of LD could be in any combination in any age group, so we are going to perform unsupervised clustering techniques, out of which Self Organizing Maps is a potential technique that we might use depending on it’s feasibility thus identifying potential clusters.
3. By using these class labels we identified earlier, we can like try to find tools most suitable for these clusters with less error for the same model and data since there are a bunch of tools that could be used and we only need an optimal set of tools just right to determine LD.

Phase 2: In Phase-2,

1. After Phase 1, we will be having the cluster labels and the shortlisted tools.
2. So, we are going to customize a model such that it classifies new data into these clusters from the data of these assessment tools we have selected.
3. We will deploy the model and perform inference using a web interface with HTML and CSS front end and python backend connecting them by flask which is a micro web framework and that’s how we are going to perform inference with our model.

(So, we haven’t decided which classification model we are going to be using, it could be a Bayesian Classifier or even an Ensembled Learning Framework that works parallelly or sequentially which ever is best for the model)

The last but the most important step is to generate an assessment report, that states all these facts like the test scores, reading age/ spelling age etc.. and if the child has LD or not, and if they have LD, like what things need to be taken care of. SO this report will either be mailed to the person’s mail which we will input from our interface or it will be available for download which ever is convenient.

Add ons : Which is attention detection, usually in the report it is also stated how the child behaved or how attentive the child is during the assessment process, so there is this idea that we have where the camera opens up in random intervals of time for 5-10 secs and detects the sentiment from face recognition, since if it is from the beginning, it might take hours to process, and this is just a further implementation idea if we are willing to like take this project forward.

Innovations:

The first innovation is automating the entire assessment process, which traditionally is a manual process with the help of a trained remedial practioner or a special educator. This provides the advantage of making no prior assumptions about the nature of the data obtained from the subject who is the child here and eliminates natural bias when carried out by humans.

The second innovation is training and performing inference with the help of Informal Educational assessment tools.

The other innovations include identifying potential tools for diagnosis, so in the process of doing the assessment we can determine which tools are more precise in predicting LD. Plus we can also study the new classes and combination of occurrence of LD, since it is not necessary that it occurs singly. By applying algorithms such as SOM, we can identify clusters in the population of subjects each corresponding to a new class of LD. Using the clusters identified by the SOM as class labels for the evaluation of Bayesian classifiers using the dataset, and enabling the identification of the tests that can be considered the best for that cluster identification.

Contribution: We are a team of two members and I’ve done all the sections except the related work and references. SO related work and references part is done by my teammate.

Extras:

1. A **self**-**organizing map** (SOM) is a grid of neurons which adapt **to** the topological shape of a dataset, allowing us **to** visualize large datasets and identify potential clusters. They are like a rectangualar lattice of hexagonal structures.

Naive **Bayes** is a kind of **classifier** which uses the **Bayes** Theorem. It predicts membership probabilities for each class such as the probability that given record or data point belongs to a particular class. The class with the highest probability is considered as the most likely class.

1. Mental Attributes Profiling System (MAPS) identification and classifying to a certain type.