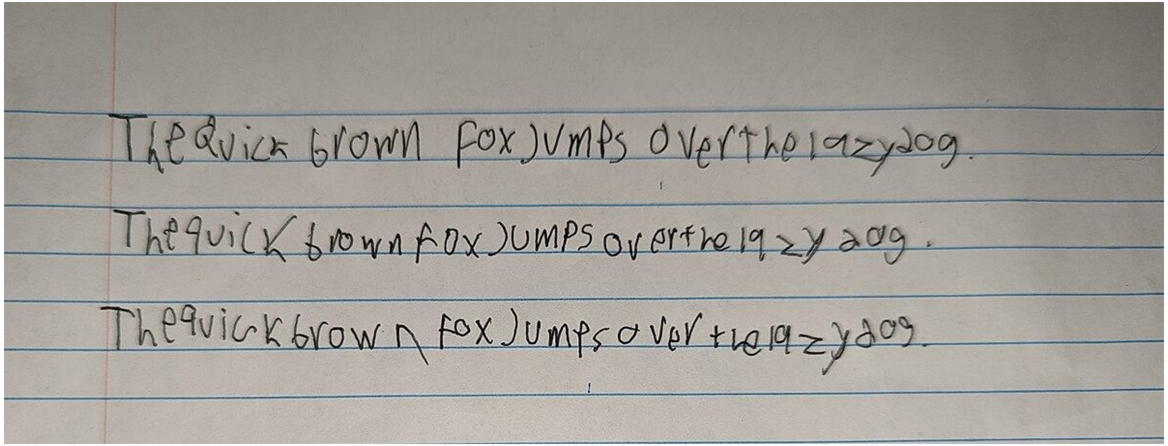
Machine Learning for Dysgraphia Detection

Victor Micha - Research Project

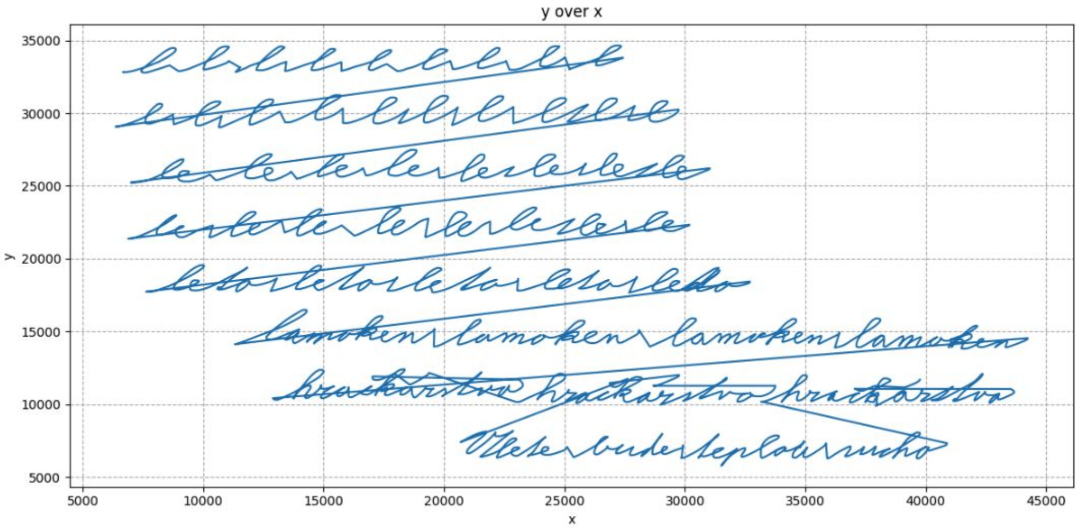
Intro

- Dysgraphia diagnosis is costly and time consuming
- Need for **reliable** automation



Background

- Kids were asked to write on a tablet
- 1 sequence of 7D data per child



1st Approach - Shallow ML

- aggregate each student's handwriting data into a single feature vector by extracting statistical features
- Loss of information :(
- Easy and fast :)

```
X
```

	sex	hand	age	1	2	3	4
0	0	1	15	53767	11888	31024.919611	29470.5
1	1	1	15	47465	9195	26634.640821	26582.0
2	1	1	14	48469	12428	28809.235427	27966.0
3	1	1	8	49767	7524	24606.965965	22472.5
4	1	1	14	42992	11855	26333.640530	25533.5
...	...	...	...	...	...	...	...
115	1	1	15	41355	8091	22671.780803	21653.0
116	1	0	15	42528	9470	23900.884446	22470.0
117	0	1	15	39601	9750	22195.464463	20755.0
118	0	1	15	39210	11885	22608.391800	21653.0
119	1	1	15	46737	11510	26492.197607	24748.0

120 rows × 22 columns

```
# y --> (labels 0 or 1 for dysgraphia for each user)
y
```

```
0      1
1      1
2      1
3      1
4      1
      ..
115    0
116    0
117    0
118    0
119    0
Name: diag, Length: 120, dtype: object
```

Results

- Results show that there is no meaningful performance change with or without SHA
- Shallow ML performs poorly ~74-75% accuracy

TABLE I: Performance of SVM Models for Dysgraphia Classification

Kernel	CV Accuracy (All)	CV Accuracy (No SHA)	Test Accuracy (All)	Test Accuracy (No SHA)	CV Diff	Test Diff
Linear	0.6644 ± 0.1457	0.6856 ± 0.1432	0.7333	0.6333	0.0211	-0.1000
Poly	0.6678 ± 0.1559	0.6722 ± 0.1551	0.8000	0.7333	0.0044	-0.0667
RBF	0.5800 ± 0.1339	0.5922 ± 0.1352	0.8333	0.8000	0.0122	-0.0333
Sigmoid	0.5111 ± 0.0544	0.5111 ± 0.0544	0.5667	0.5667	0.0000	0.0000

TABLE II: Performance of Random Forest Models for Dysgraphia Classification

Trees	CV Accuracy (All)	CV Accuracy (No SHA)	Test Accuracy (All)	Test Accuracy (No SHA)	CV Diff	Test Diff
10	0.7322 ± 0.1580	0.7300 ± 0.1550	0.5333	0.6333	-0.0022	0.1000
50	0.7589 ± 0.1516	0.7400 ± 0.1642	0.6333	0.6000	-0.0189	-0.0333
100	0.7544 ± 0.1408	0.7478 ± 0.1514	0.6000	0.6000	-0.0067	0.0000
200	0.7500 ± 0.1417	0.7444 ± 0.1560	0.6000	0.6333	-0.0056	0.0333

TABLE III: Performance of AdaBoost Models for Dysgraphia Classification

Estimators	CV Accuracy (All)	CV Accuracy (No SHA)	Test Accuracy (All)	Test Accuracy (No SHA)	CV Diff	Test Diff
50	0.7444 ± 0.1519	0.7500 ± 0.1590	0.7000	0.7333	0.0056	0.0333
100	0.7456 ± 0.1417	0.7422 ± 0.1594	0.6333	0.7000	-0.0033	0.0667
200	0.7422 ± 0.1439	0.7511 ± 0.1533	0.7000	0.7667	0.0089	0.0667
500	0.7378 ± 0.1453	0.7478 ± 0.1522	0.6667	0.7667	0.0100	0.1000

TABLE IV: Mean CV Accuracies and Differences Across Shallow ML Classifiers

Classifier	CV Accuracy (All)	CV Accuracy (No SHA)	Mean CV Diff	Mean Test Diff
SVM	60.58%	61.53%	0.94%	-5.00%
Random Forest	74.89%	74.06%	-0.83%	2.50%
AdaBoost	74.25%	74.78%	0.53%	6.67%

2nd approach - Deep ML

- train (VAE) on a feature matrix (120 students, about 20 features) to learn the distribution and generate synthetic data
- tuned SVM, trained on real features and dysgraphia labels to predict labels for the synthetic data
- train a deep (NN) for binary classification using

  1. real data

  2. real plus synthetic data

  3. synthetic data

Results

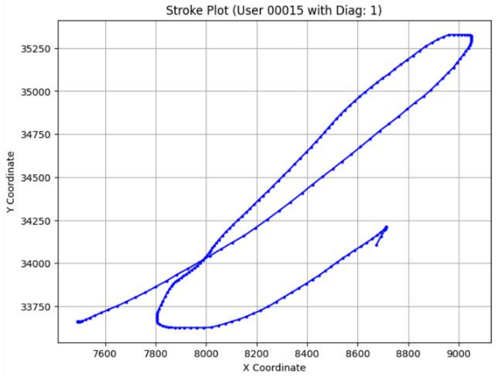
- Poor accuracy for real data only (expected)
- Higher accuracy for real+fake vs just fake
- ~75% accuracy slightly higher than ~74% for shallow ML
- High std for CV accuracy → less robust

TABLE I: Performance of Neural Network Models for Dysgraphia Classification

Training Data	Training Accuracy	CV Accuracy
Real Data Only	45.83%	54.42% ± 17.87%
Fake Data Only	77.88%	74.95% ± 27.14%
Combined Real + Fake Data	83.65%	75.74% ± 25.56%

3rd approach - Clustering

- segment handwriting time series from 120 users into 7628 individual strokes based on pen-down transitions
- preserving temporal dynamics
- K-medoids + Dynamic Time Warping (DTW) as the dissimilarity measure to cluster groups reflecting dysgraphia-specific patterns
- For diagnosis: a new stroke is assigned to the nearest cluster (via DTW to medoids) → the cluster's dysgraphia proportion IS the diagnosis
    - Averaged for multiple strokes

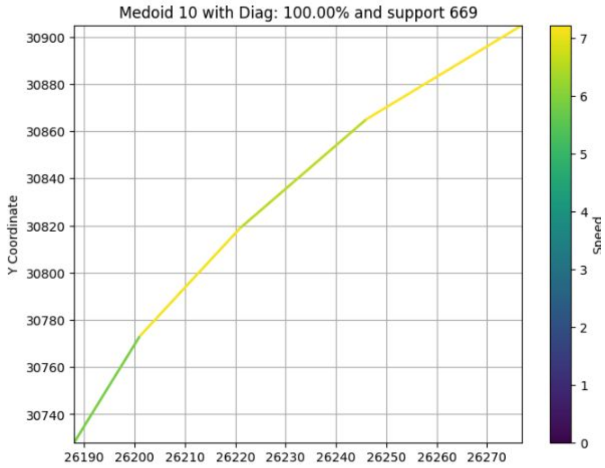


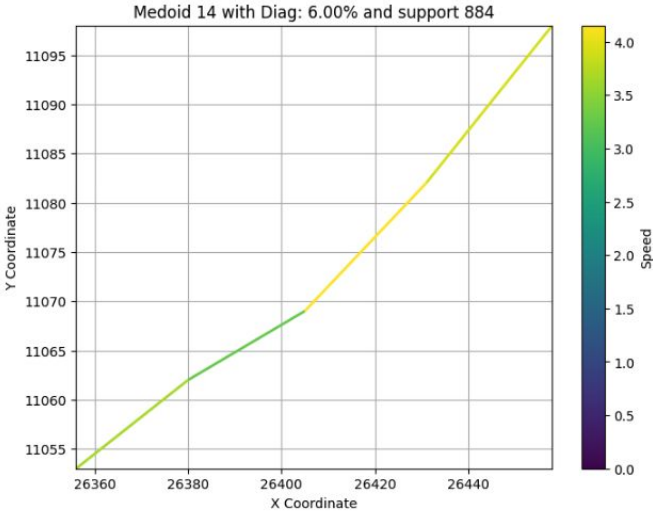
Results

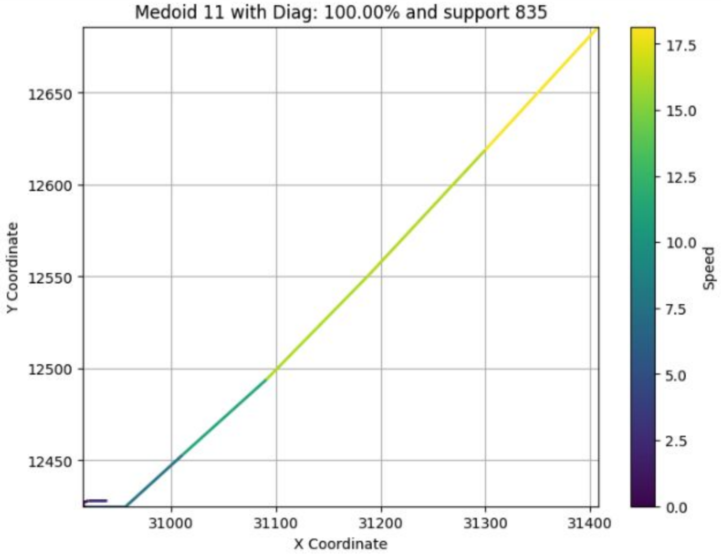
Final Cluster 0: 26 strokes, Dysgraphia: 26 (100.00%)
Final Cluster 1: 33 strokes, Dysgraphia: 33 (100.00%)
Final Cluster 2: 399 strokes, Dysgraphia: 0 (0.00%)
…
Final Cluster 12: 1437 strokes, Dysgraphia: 150 (10.44%)
Final Cluster 13: 370 strokes, Dysgraphia: 370 (100.00%)
Final Cluster 14: 884 strokes, Dysgraphia: 53 (6.00%)

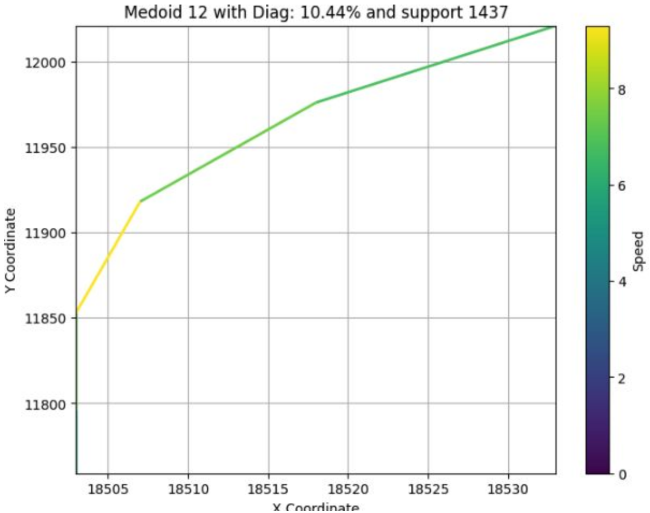
- Test Continuous Accuracy: 85.65%
-
```
test_predictions = [predict_dysgraphia([s], final_medoids, final_cluster_summary) for s in test_sequences]
test_true = [m['diag'] for m in test_meta]
test_ca = np.mean([1 - abs(p - int(t)) for p, t in zip(test_predictions, test_true)])
```

Better than 1st & 2nd approaches by ~10%


Medoid 10 with Diag: 100.00% and support 669
Y Coordinate
Speed
30740 30760 30780 30800 30820 30840 30860 30880 30900
26190 26200 26210 26220 26230 26240 26250 26260 26270
0 1 2 3 4 5 6 7


Medoid 14 with Diag: 6.00% and support 884
Y Coordinate
X Coordinate
Speed
11055 11060 11065 11070 11075 11080 11085 11090 11095
26360 26380 26400 26420 26440
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0


Medoid 11 with Diag: 100.00% and support 835
Y Coordinate
X Coordinate
Speed
12450 12500 12550 12600 12650
31000 31100 31200 31300 31400
0.0 2.5 5.0 7.5 10.0 12.5 15.0 17.5


Medoid 12 with Diag: 10.44% and support 1437
Y Coordinate
Speed
11800 11850 11900 11950 12000
18505 18510 18515 18520 18525 18530
0 2 4 6 8

Next Steps

-Instead of fixed number of clusters: need to maximize CV accuracy while minimize the number of clusters so we do not overfit (by minimizing the entropy of each clusters dysgraphia rate)

-scaling: collect more data from kids' handwritings

-integrate as background tasks in iPads/tablets for hassle-free diagnosis