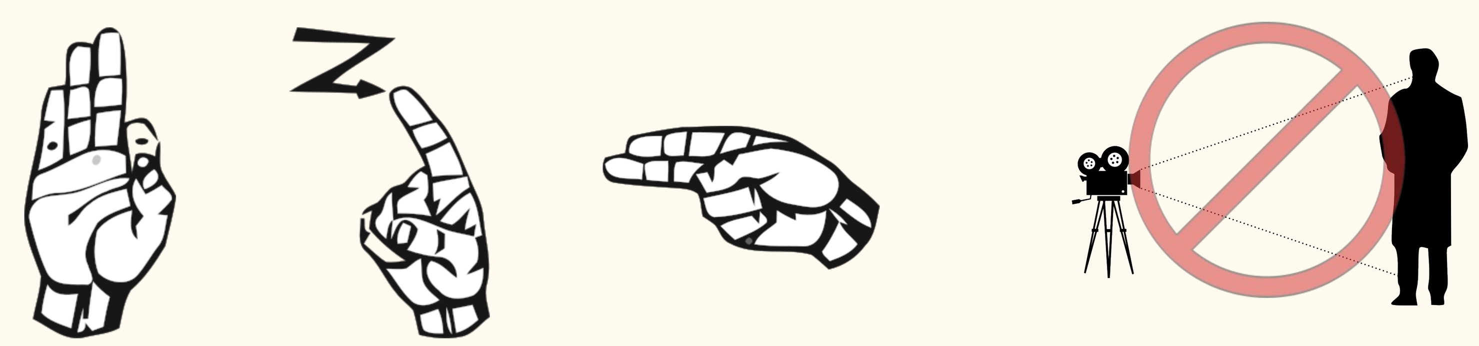
Sign Language Translation using Temporal Classification

Hardie Cate, Fahim Dalvi, Zeshan Hussain

Motivation

Hearing-impaired people often find it difficult to communicate with non-signers. Most technologies aimed at sign to natural language translation rely on cameras, forcing signers to carry around equipment and set up a proper translation environment. Instead, our approach utilizes non-invasive sensors to track the movements of a signer's hands and fingers and to predict signs using this multivariate time-series data.



Data

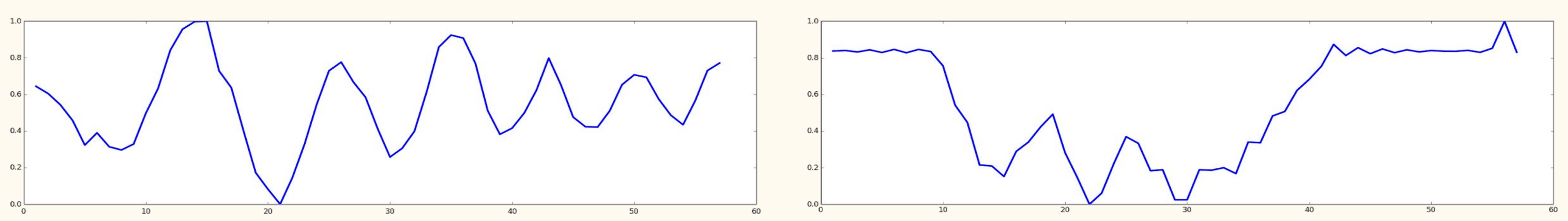


Figure 1: High and low quality data for hand movement along y-axis for sign "soon"

	High Quality	Low Quality
Number of features	22	8
Number of examples	27 (per sign)	70 (per sign)
Frequency of Sensing	200 Hz	50 Hz
Number of signs	95 signs	

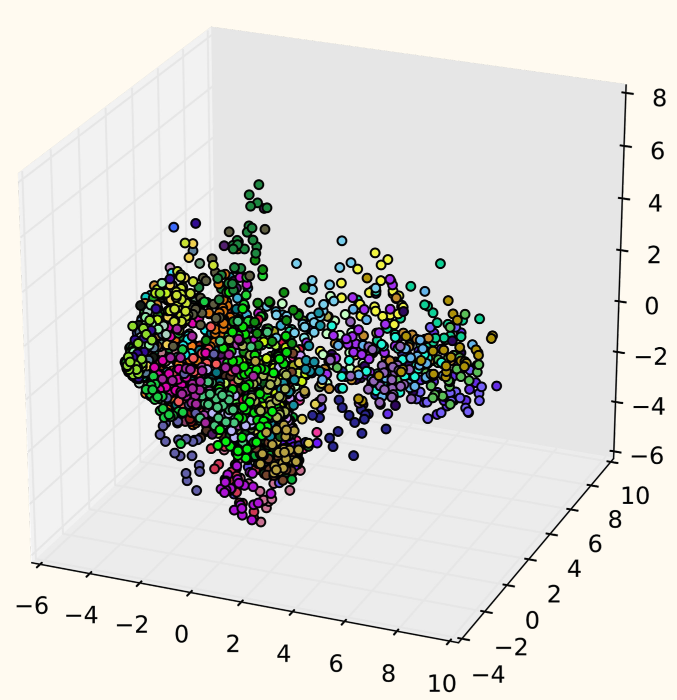
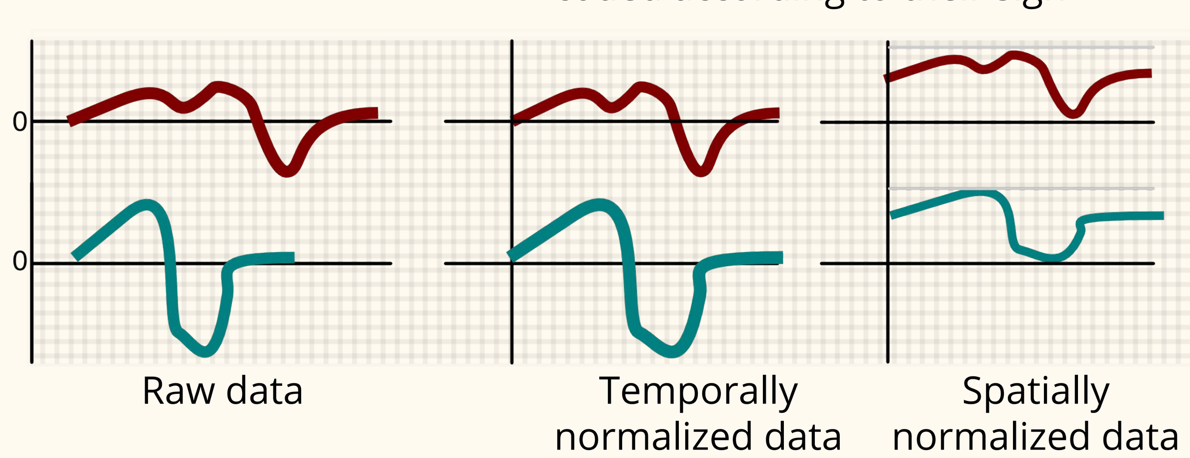
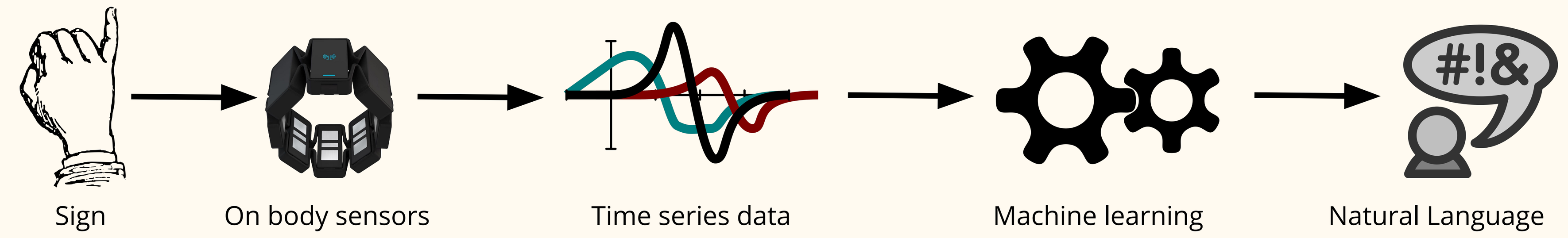


Figure 2: Depiction of PCA reduction of multi-dimensional sign space to 3D. Individual examples are color coded according to their sign

Figure 3: Illustration of temporal and spatial normalization of data of two different instances of the same sign



Methodology



Baseline

SVM + Logistic Regression:

- Does not take into account temporal nature of data
- Accurate prediction for high quality data, poor performance on low quality
- We try to improve on SVM's performance on low-quality data

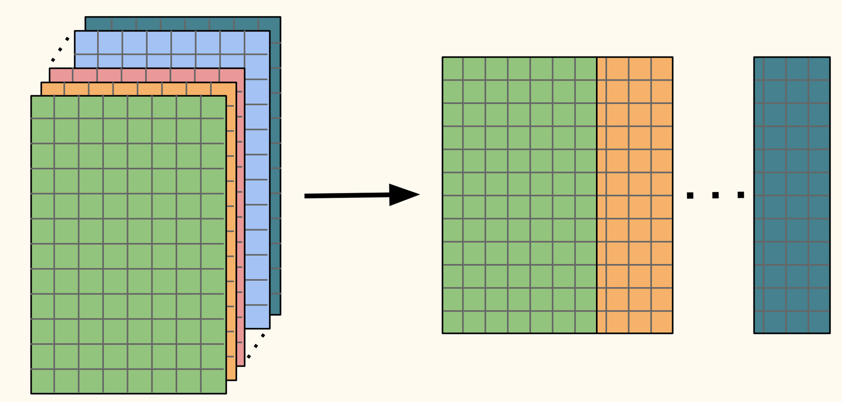


Figure 4: Flattening of 3D data tensor into 2D matrix. Each row in the matrix is a flattened vector containing signals across all time steps. Flattened matrix is input to SVM.

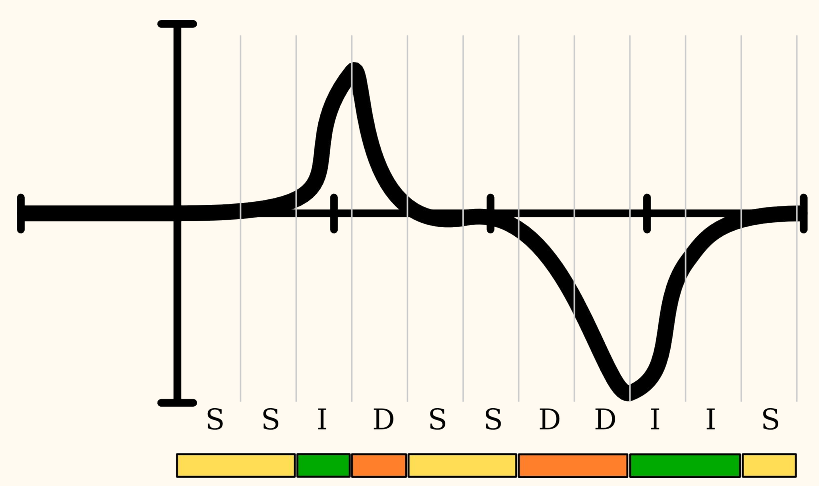


Figure 5: Discretization of signal into increasing, decreasing and steady states for SPM algorithm

Alternate Strategies

Neural Network Architecture + Sequential Pattern Mining (SPM):

- Long Short-Term Memory
  - backpropagation through time
- Apriori SPM
  - discretization
  - spatial/temporal scaling

Final Approach

SPM to find Temporal Abstraction Patterns (Batal et al.)

- Data segmentation + temporal abstraction
- Pattern mining
- Chi-square test to determine most relevant patterns
- Feature selection from patterns

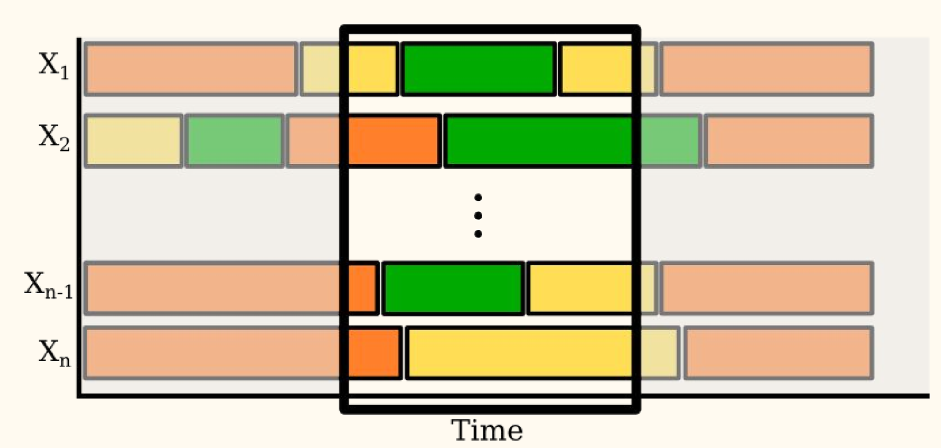


Figure 6: Graphical depiction of pattern identification across multiple signals within a given time window

Results & Analysis

SVM + Logistic Regression Results

	SVM high quality	Log. reg. high quality	SVM low quality	Log. reg. low quality
Precision	0.942	0.938	0.566	0.444
Recall	0.936	0.933	0.55	0.444
F1	0.936	0.932	0.549	0.436

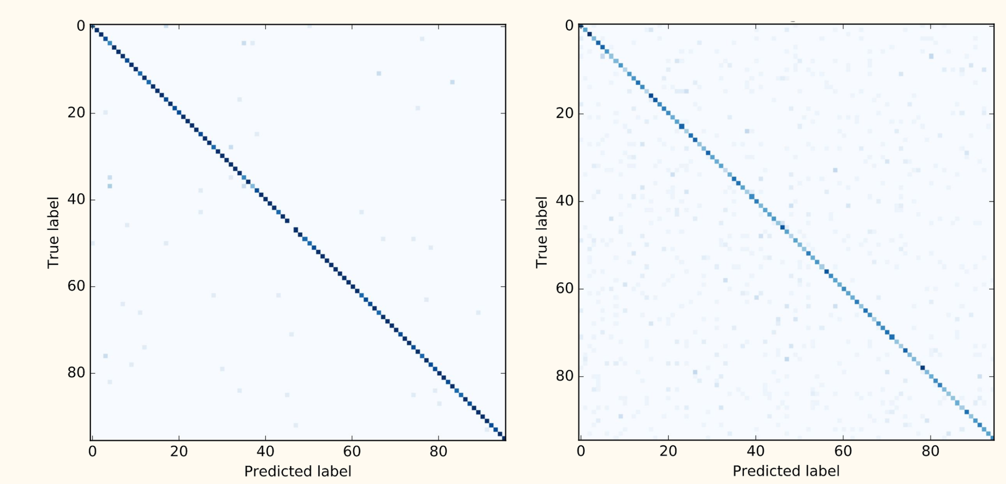


Figure 7: Confusion matrices for SVM sign classification for high quality (left) and low quality data on all 95 signs

- Signs "confused" more frequently on low quality data
- SVM performs slightly better than logistic regression

LSTM Results

Training Error	0.518
Testing Error	0.778

- Structure of model causes high error
- Incorrect backpropagation

Future Work

We plan to adjust our discretization and temporal abstraction (state definition) to speed up pattern generation. There are also several hyperparameters in our SPM algorithm that we can tune to improve speed and results. These include the window size used in pattern generation and the minimum support used to filter patterns. We hope that these changes will result in higher performance on the low quality dataset.