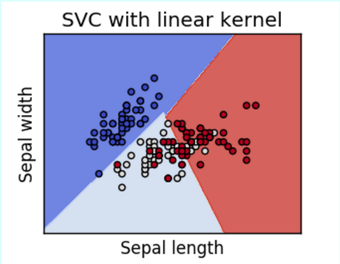
Voice

Sign Language Translation System

Theory

- Segments of glove sensor data get converted into 64 features (maximum absolute deviation of the acceleration, variance of angular acceleration, etc.) best representing the hands' movements
- Every model learns values for relevant constants allowing it to be used for prediction later. These constants are stored in the backend to be used on the go for real-time prediction
- Support Vector Machine (highest practical accuracy) uses a kernel trick to transform the data into higher dimensions to make it linear classifiable.
- The graph below shows an example of a SVM classifying three classes of data by constructing three hyperplanes in a higher dimension. A similar transformation can be imagined for n points representing n signs in a higher dimension



Design Considerations

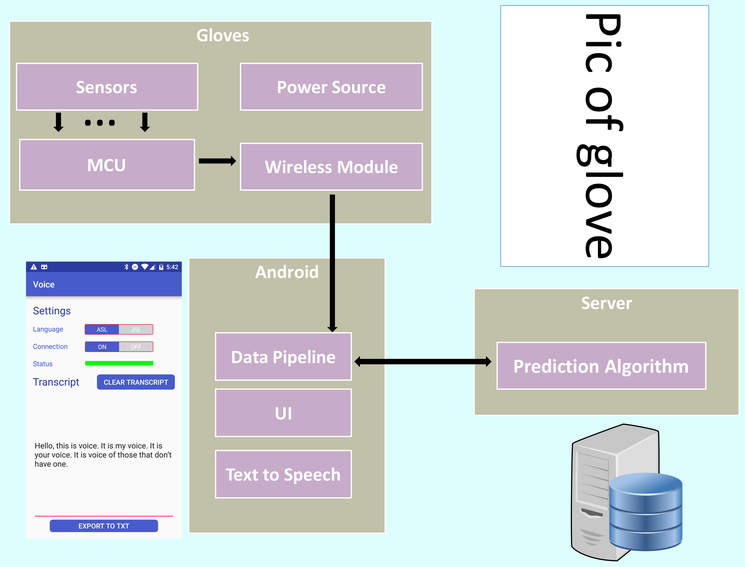
- Computer vision (requires a camera, very difficult to capture just a person's hands in changing environments)
- Data gloves (very expensive)
- Support Vector Machines vs. Naive Bayes vs. Decision Trees (SVM has highest accuracy)

Acknowledgements

- Group Members: Akshay Budhkar, Eliot Chan, Biraj Kapadia, Amish Patel
- Consultant: Dana Kulić
- Course Coordinator: Dan Davison

[1] "Statistics on deaf Canadians," Canadian Association of the Deaf, 2015. [Online]. Available: http://cad.ca/issues-positions/statistics-on-deaf-canadians/.
[2] T. Harrington, "Deaf population of the U.S. - Local and regional deaf populations," Gallaudet University Library Library, 2010. [Online]. Available: http://libguides.gallaudet.edu/content.php?pid=119476&sid=1029190.
[3] "Plot different SVM classifiers in the iris dataset — scikit-learn 0.18.1 documentation", Scikit-learn.org, 2017. [Online]. Available: http://scikit-learn.org/stable/auto_examples/svm/plot_iris.html.

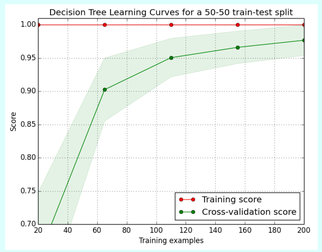
Principle of Operation

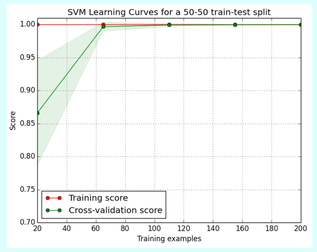


Results and Analysis

- Machine Learning (ML) models were trained on 100 iterations/sign of three users and used to predict on 100 iterations/sign of one user to calculate accuracy
- Logistic Regression and Support Vector Machines (SVM) perform better than other ML models like Naive Bayes and Decision Tree Learning
- More the input data, better the accuracy of the models. SVM converges with fewer iterations and a higher accuracy compared to Decision Tree Learning as seen in the graphs below

Model Name	ASL (15 Signs)	JSL (20% of the symbols)
SVM	96.3%	100%
Logistic Regression	99.6%	100%





Motivation

- ~1,500,000 culturally deaf in the United States and Canada [1][2] with no efficient and natural way to communicate with the hearing
- Communication barriers between sign language speakers and the general populace remain high

Objective

- Build a system to allow sign-language speakers to communicate with non-speakers in an unobtrusive and natural way
- Support a subset of two sign language dialects, namely American Sign Language and Japanese Sign Language

Advantages

- Allows users to communicate in sign language in a mobile environment
- Portable and minimal setup time
- Easily integrated into daily routine
- Cheaper than other alternatives
- Seamless extension for supporting more signs in the future

Features

- Supports 15 American Sign Language (ASL) signs
- Supports ~20% of Japanese Sign Language (JSL) symbols
- Theoretical accuracy of over 99%
- Weighs under X grams
- Can run for more than 10 hour on a single charge
- Translation latency less than 500 ms
- Bluetooth communication between the gloves and mobile application