Prediction of Clinical Parameters for Gait Disorder Recognition

Mehrad Moradshahi
Soheil Esmaeilzadeh
Ouassim Khebzegga

Goals

- Reducing diagnosis time after a visit
- Predicting Cadence and other clinical parameters
  - Using patient's clinical data
  - Using video recordings
- Determining whether patient needs treatment (Classification)

Data

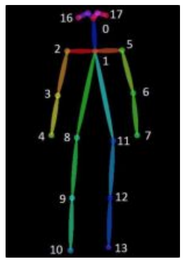
- Clinical data gathered by physicians in Mobilize center by measurements and recording videos
- Data are from patients diagnosed with gait disorders
- Gathering data is time consuming and expensive
- Not cleaned (N/A, mismatched, etc.)

mobilize
Center for Mobility Data Integration to Insight

Scott Delp, PhD
Department of Bioengineering
Stanford University

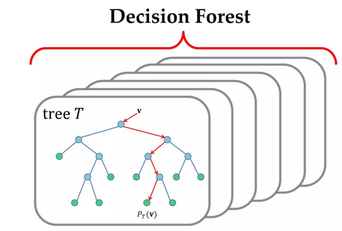
Feature Extraction

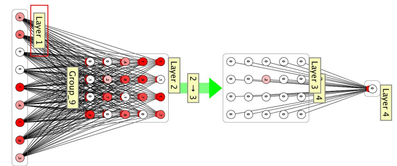
- Raw data set containing many different data types including cadence, step length, speed, different joint and body parts attributes, etc.
- A huge number of .json files extracted from videos via Openpose
- Limited features to reduce overfitting: picking features which seems relevant to target value using cross validation
- Parsed raw datasets for each patient and extracted relevant feature data into one large, shuffled dataset
- Creating indicator features
- grouping target values via binning for classification methods



Models

- Regularized Least Squares Logistic Regression
- Linear Regression and Classification
  - SVM minimizing squared loss / hinge loss via SGD.
  - Optimized with small step sizes, limited iterations, and regularization penalties.
- Random Forest Regression and Classification
  - Averages over decision trees of dataset subsamples
- Clustering, in conjunction with the above models
- Neural Network: Extract feature relationships for smarter prediction using layered set of neurons





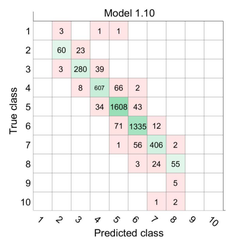
Results

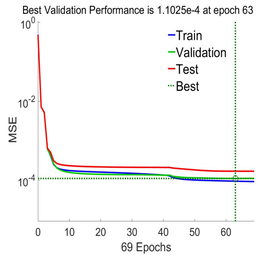
Regression Methods for Cadence

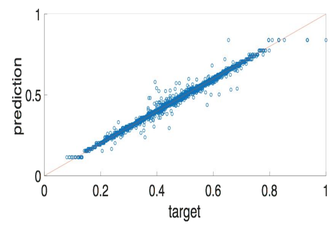
Method:	Linear	RF	SVM	Gaussian Process	NN
RMSE (min):	0.018937	0.013645	0.015415	0.014739	0.01038
R^2 :	0.97	0.99	0.98	0.98	0.99

Classification Methods for Cadence

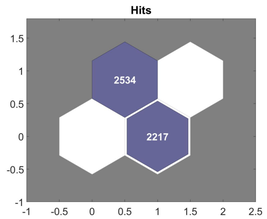
Method:	Linear Discriminant	RF	SVM	KNN	Ensemble (Boosted trees)
Accuracy:	83.3%	95.9%	91.6%	69.1%	96.2%







- Predicting if patient needs treatment
  - Used same data features
  - SOM network with 1 layer and 2x2 map
  - After training, data, only hit 2 clusters with (46%, 53%) ratio



Discussion and Suggestions

- The best results are achieved by Random Forest and Neural Network
- Ensemble learning has a boosting effect after binning cadence into 10 different categories
- Neural network can be fed into random forest to improve our best results further
- Predictions using features from clinical data work well but predictions from processing video files is challenging due to extremely noisy data
- Recording data needs careful treatment as presence of more than one person and shadows might interfere with patient's position detection
- Additional computing power would enable the predictions to be done using the entire patient's features, and possibly exposing more feature relationships

References

- **[1].** *Andrea Mannini, Diana Trojaniello; A machine learning framework for gait classification using inertial sensors: Application to elderly, post-stroke and huntingtons disease patients. 2016.*
- **[2].** *Alex Fu.; Seizure prediction from intracranial eeg recordings. 2014.*
- **[3].** *Tomas Simon, Hanbyul Joo.; Hand keypoint detection in single images using multiview bootstrapping. 2017.*