Upper Extremity Symmetry Differences as Measured by Machine Learning Models

Stephen MacNeille, Benjamin Park, Aleksandra M Kostic, David Ahn, Matthew Weintraub, Jean Francois Daneault, Alice Chu,
Rutgers New Jersey Medical School, NJ
smm668@njms.rutgers.edu

Introduction:

Handedness, or hand laterality, plays a significant role in Activities of Daily Living (ADL) and conditions like stroke and cerebral palsy can induce additional asymmetry. We used 3D motion analysis, submovement decomposition, and ten popular machine learning models to determine relative dominance along the upper extremity. By training the models on data from participants' dominant and nondominant hands, we derived an accuracy score for each model's prediction of dominant versus non dominant status on unseen data. Our results indicate that tree based models, particularly RandomForest, outperform the others—with preliminary results suggesting highest sensor accuracy scores ranging from .93 for the ring task to .95 for the block task. The accuracy tends to increase as one moves from proximal to distal sensor locations. Additionally, we believe the differences in the top-performing model scores across the upper limb can serve as a benchmark for evaluating musculoskeletal interventions.

Methods:

Using the Vicon Vero motion capture system, we documented the movements of 25 healthy participants executing ADLs at a sampling rate of 100Hz. We placed 10 markers on the upper extremities to track the three-dimensional positions of the shoulder, elbow, and wrist joints. Participants were tasked with stacking six blocks and transferring a large ring from one hand to the other, subsequently positioning it on a contralateral cone. Each task was repeated thrice, using both the self-reported dominant and non-dominant hands. We then processed the positional data, eliminated noise, and established planes in x (anterior-posterior), y (medial-lateral), and z (superior-inferior) directions. After deriving velocities, we demarcated the beginning and end of each trial based on wrist velocity criteria and segmented each trial into submovements.

Subsequently, we extracted features from each participant's task and discarded redundant ones, specifically those with a correlation greater than .95 with another or those with empty values. This process retained between 30 to 60 features for each sensor and task. We trained these features on the classifiers, grouping them by patient number and the status of the dominant hand. We implemented k-fold cross-validation (with 5 folds) and maintained an 80/20 training-to-test split.

For hyperparameter optimization, we employed a grid search technique, where we experimented with a spectrum of values for each sensor and model combination, retaining only the best-performing ones.

Results:

Preliminary results suggest that tree based classifiers show the strongest performance and laterality becomes more pronounced as one moves distally. Generally speaking, acceptable (over .90) accuracy measurements were obtained in the distal extremities, while the models struggled to differentiate proximal sensor data. Based on the heatmaps below (Figure 1), one can see the trend favoring distal sensors. Additionally, the block task resulted in higher accuracy scores overall.

When looking at the combined data (Figure 2), the RandomForest model performed well in terms of accuracy (0.831), but had a relatively lower AUC-ROC (0.793) and F1-score (0.725). The DecisionTree model had the highest accuracy (0.839) but performed poorly in other metrics, particularly log loss (5.033). The CatBoost model showed a balanced performance with an accuracy of 0.784, AUC-ROC of 0.812, and F1-score of 0.718. The XGBoost and GradientBoosting models had similar performances with accuracies of 0.757 and 0.761, respectively. The AdaBoost model had a comparable performance with an accuracy of 0.752 but had a high log loss (0.962). The LogisticRegression and SVM models had the lowest accuracies of 0.697 and 0.679, respectively, with the SVM model performing the worst in terms of AUC-ROC (0.605) and F1-score (0.602).

In terms of precision and recall, the RandomForest and ExtraTrees models performed the best with values of 0.763 and 0.761 for precision, and 0.735 and 0.737 for recall, respectively. The LogisticRegression and SVM models had low precision and recall values, indicating their limited ability to correctly classify hand dominance.

Discussion:

We believe that tree based models can help quantify relative dominance in the distal extremities, particularly in tasks involving fine motor skills, like stacking blocks—where models showed higher accuracy scores. That said, a few interesting trends emerge. First, with both tasks, DecisionTree exhibited smaller drops in performance compared to the other models when going from distal to proximal. Also, sensors on the elbow had a smaller range between the best and worst performing models. Finally, different features had different relative importances across models. Although the best performing model, RandomForest, did not assign weights to features, further analysis will be needed to understand why particular features might be important to different models.

Significance and Clinical Relevance:

As musculoskeletal conditions often disproportionately affect one side of the body, establishing a novel framework to quantify upper extremity dominance can help diagnose these conditions and assess intervention efficacy.

Figure 1: Model Accuracy Heatmap Across Sensors and Tasks

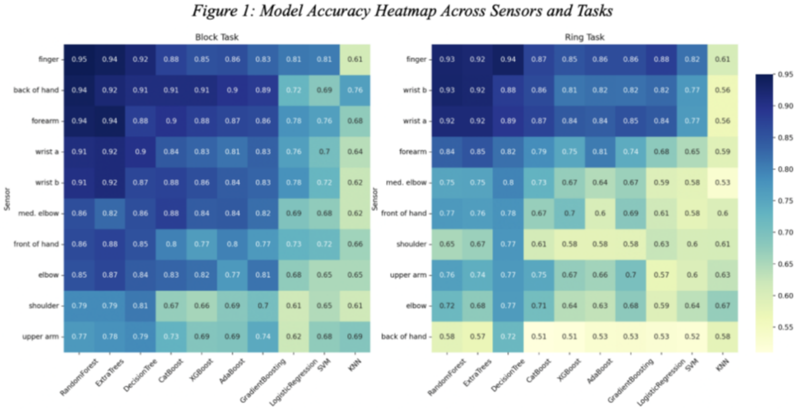

Figure 2: Combined Model Performance Across Tasks and Sensors

	Accuracy	AUC-ROC	F1-score	Log loss	Precision	Recall
RandomForest	0.831	0.793	0.725	0.567	0.763	0.735
ExtraTrees	0.828	0.804	0.726	0.585	0.761	0.737
DecisionTree	0.839	0.64	0.605	5.033	0.641	0.63
CatBoost	0.784	0.812	0.718	0.607	0.749	0.727
XGBoost	0.757	0.77	0.702	0.574	0.733	0.711
GradientBoosting	0.761	0.758	0.694	0.825	0.731	0.705
AdaBoost	0.752	0.751	0.693	0.962	0.731	0.706
LogisticRegression	0.697	0.696	0.634	1.411	0.66	0.648
SVM	0.679	0.605	0.602	0.67	0.628	0.624