Model Card: Hand Position Estimation

Model Details

Model Name: Hand Position Estimation
Developed By: Nicholas Stranges
Framework Used: scikit-learn
Version: 1.0
Date Created: January 20, 2025

Model Overview

This model is a **Random Forest Classifier** trained to predict hand positions based on the combinations of fully-open or fully-closed in three degrees of freedom. The three degrees of freedom are the thumb, index/middle finger, and ring/pinky finger. Input comes from three electromyography (EMG) sensors and a pulse blood concentration sensor. It is built using the `sklearn.ensemble.RandomForestClassifier`.

Model Architecture

The Random Forest model consists of an ensemble of decision trees, where each tree is trained on a subset of the data using bootstrap aggregation (bagging). The model averages the predictions of multiple decision trees to reduce overfitting and improve generalization.

Hyperparameters:

- `n_estimators`: 50 (number of trees in the forest)
- `max_depth`: None (trees grow until all leaves are pure or contain fewer than min_samples_split samples)
- `max_features`: None (default: auto selection)
- `random_state`: 42 (ensures reproducibility)
- `n_jobs`: -1 (utilizes all available CPU cores for parallel training)

Data

- **Dataset:** Processed from `processed_data.csv`
- **Target Variable:** `Position`
- **Feature Selection:**
    - Accelerometer and gyroscope signals were removed.
    - Derived features such as RMS voltage and derivatives were eliminated.
- **Data Split:**

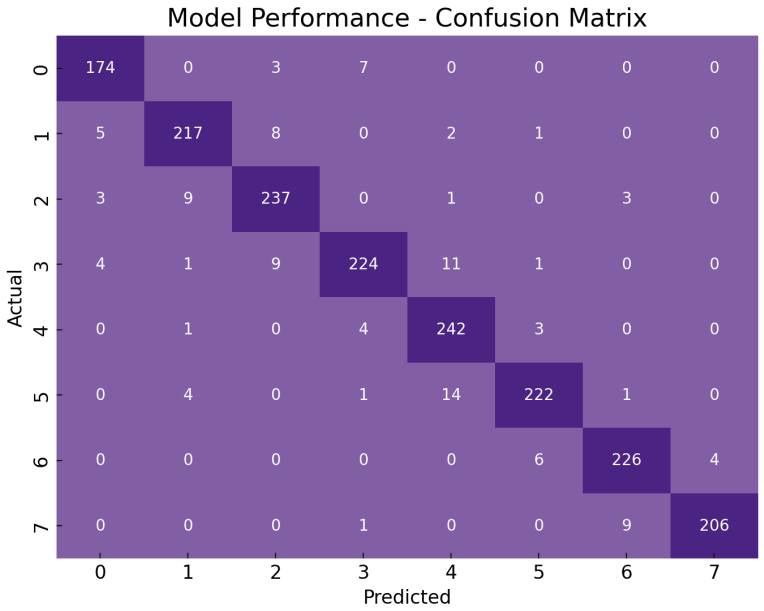
- Training Set: 80%
- Test Set: 20%
- Stratified splitting for balanced class representation.

Performance Metrics

Evaluation on Test Data:

- **Accuracy:** 93.78%
- **Precision (Weighted):** 93.85%
- **Recall (Weighted):** 93.78%
- **F1 Score (Weighted):** 93.78%
- **Matthews Correlation Coefficient (MCC):** 92.89%

Confusion Matrix:



Feature Importance

The top features contributing to the model's predictions are:

1. emg2_var: 17.67%
2. emg3_var: 16.69%
3. emg1_var: 15.04%
4. pulse_avg: 12.22%

Limitations and Risks

- **Potential Biases:** Model performance may be partially due to data features that do not determine actual hand position.

- **Overfitting:** Ensuring regularization via tree depth constraints or feature selection.
- **Feature Engineering Sensitivity:** Changes in feature extraction or data preprocessing may affect performance.

Model Deployment

The trained model can be exported using `m2cgen` to generate C-compatible code for deployment in embedded systems or real-time applications.

Conversion to C Code:

- Uses `m2cgen.export_to_c(model, function_name="predict")`
- Adjustments include variable renaming and formatting optimizations.

Future Improvements

- Fine-tuning hyperparameters
- Exploring other model types for this problem. Other models were tested with limited success, and all performed worse than this model.
- Expanding feature selection to incorporate new derived metrics. Especially in the frequency domain because of the noisy output of the EMG sensors.

Author: Nicholas Stranges
Affiliation: MSE 4499 Group 6
Date of Last Update: March 2, 2025