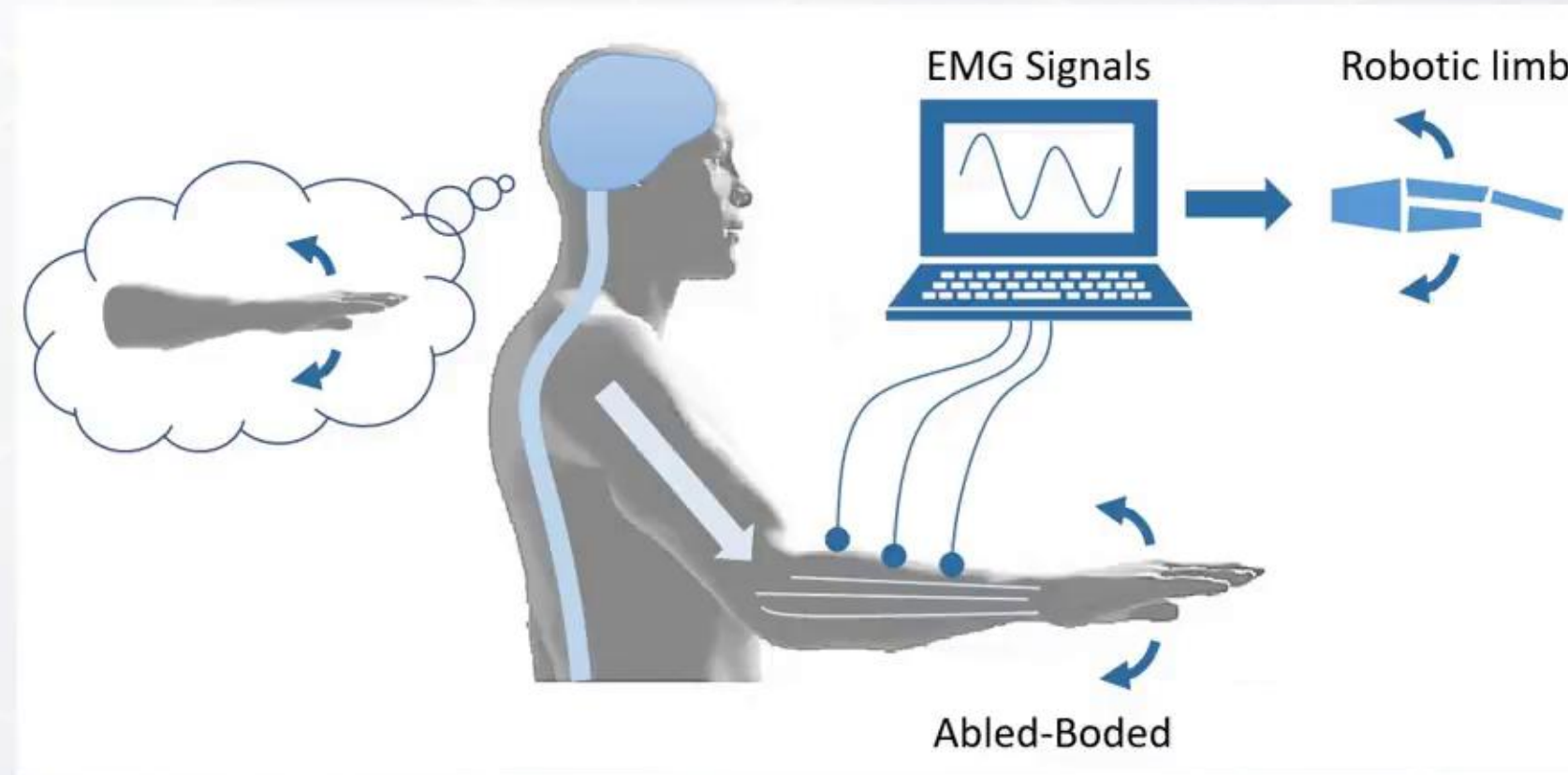# Feature Selection in Machine Learning

Guohui Song
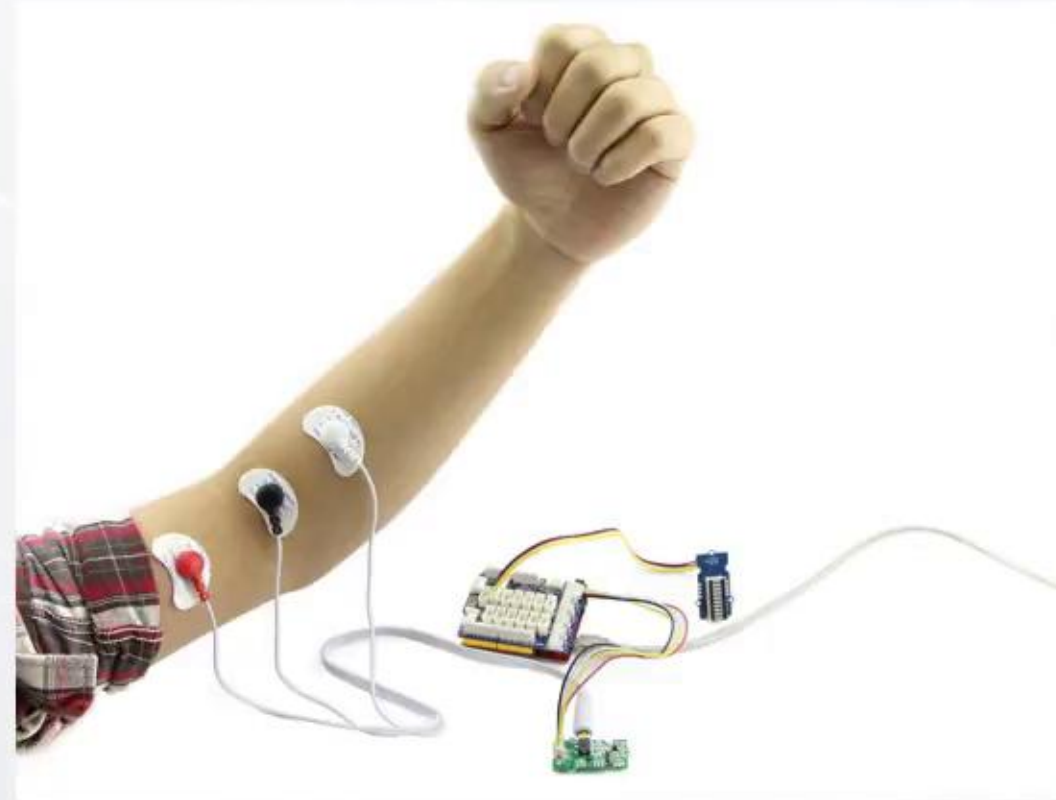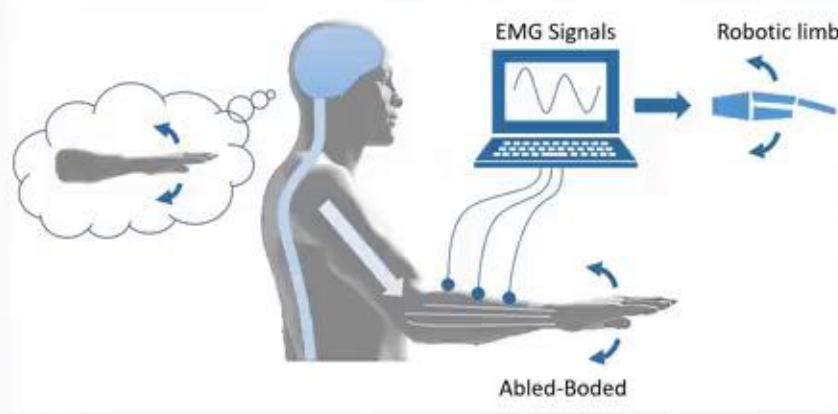
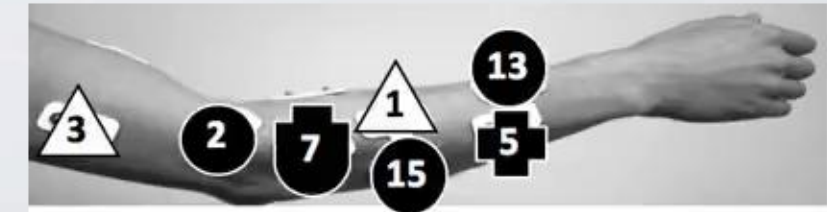Old Dominion University

Designing an EMG-Controlled Robotic Hand
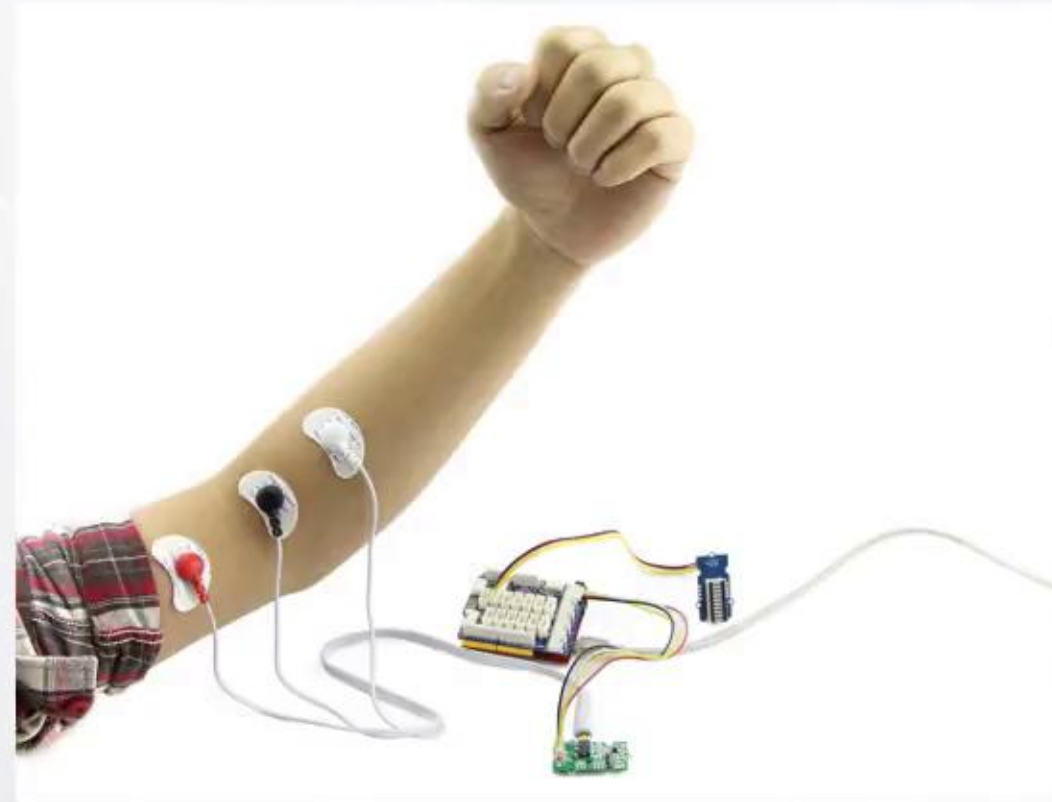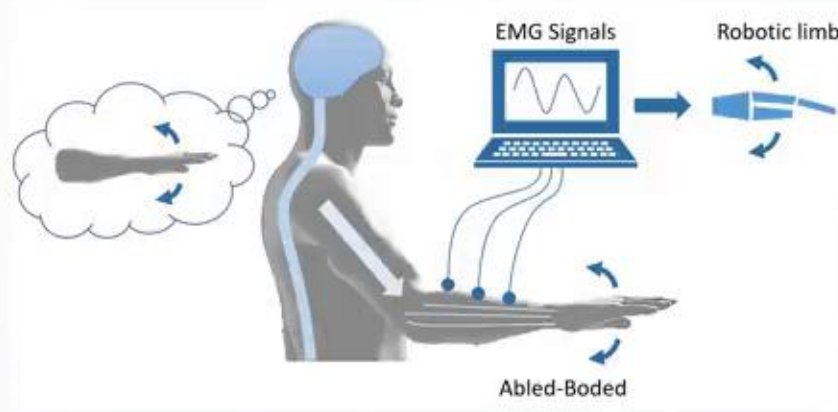

A controller converts EMG signals into hand movement.

Designing an EMG-Controlled Robotic Hand

Designing an EMG-Controlled Robotic Hand

Designing an EMG-Controlled Robotic Hand
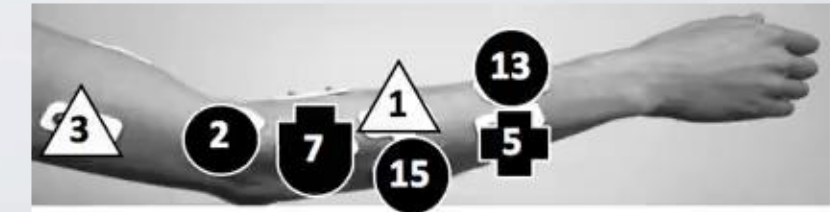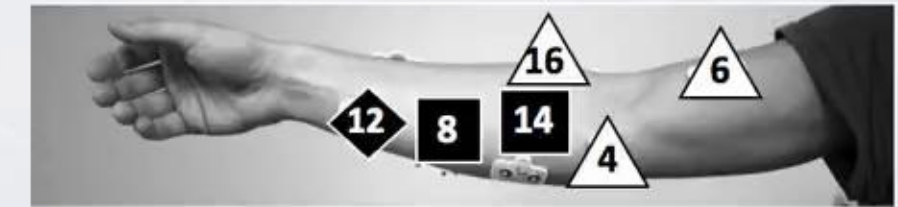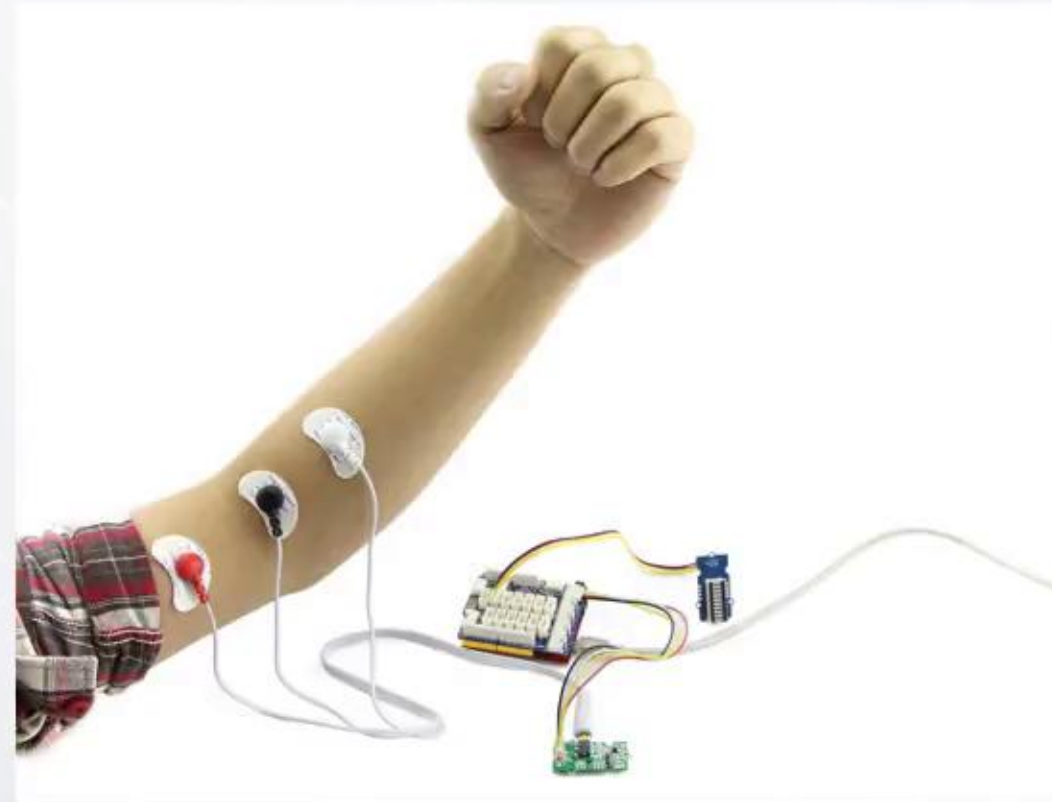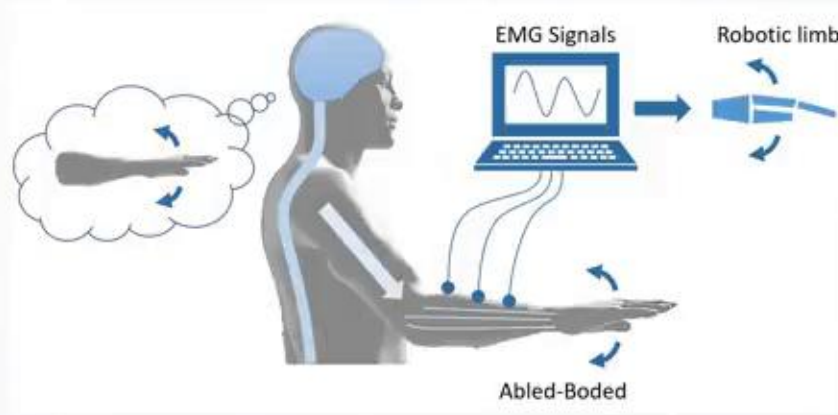
How to choose the sensor locations?

Disease Discoveries Using Biomarkers

- Identifying biomarkers that are indicative of a specific disease, such as breast cancer or Alzheimer's disease.

Disease Discoveries Using Biomarkers

- Identifying biomarkers that are indicative of a specific disease, such as breast cancer or Alzheimer's disease.

Features

- gene expressions
- protein levels
- metabolites
- microbiome composition

Disease Discoveries Using Biomarkers

- Identifying biomarkers that are indicative of a specific disease, such as breast cancer or Alzheimer's disease.

Features

- gene expressions
- protein levels
- metabolites
- microbiome composition

You might have data on the expression levels of 20,000 genes for each patient sample.

Feature Selection Methods

Feature Selection Methods

- filter methods: use statistical tests to measure the relation between each feature and the output.
  - Pearson correlation, mutual information, Chi-Square etc.
  - Fast but not very accurate. They do not consider interactions among features.

Feature Selection Methods

- filter methods: use statistical tests to measure the relation between each feature and the output.
  - Pearson correlation, mutual information, Chi-Square etc.
  - Fast but not very accurate. They do not consider interactions among features.

- wrapper methods: search the space of all possible subsets of features.
  - Forward selection, Backward elimination, Recursive feature elimination, Exhaustive feature selection.
  - Slow but more accurate. They consider interactions among features.

Feature Selection Methods

- embedded methods: embed the feature selection process into the model learning process.
  - Lasso, Ridge regression, Elastic net, Random forest, Gradient boosting.
  - Combine the benefits of both the wrapper and filter methods. Include interactions of features but also maintain reasonable computational costs

Feature Selection Methods

- embedded methods: embed the feature selection process into the model learning process.
  - Lasso, Ridge regression, Elastic net, Random forest, Gradient boosting.
  - Combine the benefits of both the wrapper and filter methods. Include interactions of features but also maintain reasonable computational costs

We assign a weight to each feature and find the best weights under the constraints of sparseness.

Multiscale LASSO Feature Selection on EMG Data

Multiscale LASSO Feature Selection on EMG Data

Multiscale LASSO

Movement	Selected sensors	CV MSE	Time (m)
Finger	7,12	7.2E-2	47
Wrist	11,15	3.9E-2	70

Multiscale LASSO Feature Selection on EMG Data

Multiscale LASSO

Movement	Selected sensors	CV MSE	Time (m)
Finger	7,12	7.2E-2	47
Wrist	11,15	3.9E-2	70

It does not work well if the data is too few comparing to the number of features.

Bayesian Methods

Bayesian Methods

- When the data lacks of quality/quantity, we need to incorporate prior knowledge.

Bayesian Methods

- When the data lacks of quality/quantity, we need to incorporate prior knowledge.

Prior + Data $\longrightarrow$ Posterior

Bayesian Methods

- Estimate the probability of head in a coin toss.
  - toss the coin 10 times and get 7 heads

Bayesian Methods

- Estimate the probability of head in a coin toss.
  - toss the coin 10 times and get 7 heads

- Prior comes from the domain knowledge

- Prior is in particular useful for limited data

- Prior will be "corrected" by the data

- Do not eliminate the possibility unless absolutely impossible

Thank you for your attention!

Questions?

Guohui Song, gsong@odu.edu