MACHINE LEARNING APPROACHES USING MATLAB

PREMANANDA INDIC, PH.D.
DEPARTMENT OF ELECTRICAL ENGINEERING
OUTLINE

- INTRODUCTION
- DIFFERENT MACHINE LEARNING APPROACHES
- DISCUSSION
OUTLINE

- INTRODUCTION
- DIFFERENT MACHINE LEARNING APPROACHES
- DISCUSSION
INTRODUCTION

What is Machine Learning?

- Machine Learning is a field of study that gives computers the ability to “learn” without being explicitly programmed
  - Prediction
  - Classification

Samuel AL, IBM J. Research & Development, 1959, vol. 3 (3), 210-229
INTRODUCTION

- Too many books spoil the curiosity

- Start with Andrew Ng, Machine Learning, Stanford University available on YouTube

  Some Statistics & Programming Knowledge Helps!
INTRODUCTION

Machine Learning with MATLAB

You have a complex problem involving a large amount of data and lots of variables. You know that machine learning would be the best approach—but you’ve never used it before. How do you deal with data that’s messy, incomplete, or in a variety of formats? How do you choose the right model for the data?

Sounds daunting? Don’t be discouraged. A systematic workflow will help you get off to a smooth start.
INTRODUCTION

MathWorks

University of Texas at Tyler

Get Software | Learn MATLAB | Teach with MATLAB | What's New

MATLAB Access for Everyone at

University of Texas at Tyler


ORS Webinar Recordings
https://www.uttler.edu/research/ors-research-design-data-analysis-lab/ors-research-design-data-analysis-lab-resources/ors-consultant-recordings.php

30 Minute Feature Extraction

Data ➔ Preprocessing ➔ Feature Extraction ➔ Feature Selection

Statistical or Machine Learning Models
INTRODUCTION

Preprocessing

Feature Extraction

Feature Selection

Data

Command window

Text Editor

Workspace

30 Minute Feature Extraction

Statistical or Machine Learning Models

Data → Preprocessing → Feature Extraction → Feature Selection
OUTLINE

 INTRODUCTION

 DIFFERENT MACHINE LEARNING APPROACHES

 DISCUSSION
Statistical vs. Machine Learning Models

Purpose:

Statistical models are used for inference (To find association between features and an outcome). Results should be interpretable.

Machine Learning models are used for prediction (Use features that can predict an outcome). Results may not be interpretable.
Statistical vs. Machine Learning Models

Association vs. Prediction

Philips Actiwatch

Healthy Individual

Individual with depression

\[ N = 24 \]

\[ r = -0.73; p < 0.001 \]

\[ VI = m \times SI + C \]

\[ m = r \frac{\sigma_{VI}}{\sigma_{SI}} \]

\[ C = \mu_{VI} - m\mu_{SI} \]

\[ \hat{SI} = a \times VI + b \]

Sensitivity & Specificity
LEARNING APPROACHES

- **Supervised Learning**
  Learning a relationship between features and the outcome using a training set

- **Unsupervised Learning**
  Learning underlying structures in features
LEARNING APPROACHES

- Supervised Learning
  - Linear Regression
  - Logistic Regression
  - Support Vector Machine
  - Artificial Neural Network
  - ........................
  - ........................
  - ........................
  - ........................
LEARNING APPROACHES

- Unsupervised Learning

  Clustering
  
  • Principal Component Analysis
  • Independent Component Analysis
  • Singular Value Decomposition
  • .......
  • .........
LEARNING APPROACHES

Do machines actually “learn”?

\[ VI = m \times SI + C \]
LEARNING APPROACHES

Do machines actually “learn”?

\[ e(N = 1) = \overline{VI}(N = 1) - VI(N = 1) \]
\[ e(N = 2) = \overline{VI}(N = 2) - VI(N = 2) \]
\[ \ldots \ldots \]
\[ e(N = 24) = \overline{VI}(N = 24) - VI(N = 24) \]

\[ E = \sum_{n=1}^{N} e^2 \]

\[ \overline{VI} = m \times SI + C \]
LEARNING APPROACHES

Do machines actually “learn”?

How do we find minimum E?

\[ \bar{VI} = m \times SI + C \]
LEARNING APPROACHES

- Do machines actually “learn”? 

How do we find minimum E?

- Gradient Descent

by Louis Augustin Cauchy in 1847

\[
\overline{VI} = m \times SI + C
\]

Linear Regression

\[
\hat{SI} = a \times VI + b
\]
LEARNING APPROACHES

Do machines actually “learn”?  

Classification of High Risk (n=43) vs. Low Risk (n=95)  

0 = Low Risk, 1 = High Risk

\[ p = \frac{1}{1 + e^{-(a \times VI + b)}} \]  

Linear Regression

\[ p = \frac{1}{1 + e^{-(A^T F + B)}} \]  

Logistic Regression  

Accuracy ~73%
LEARNING APPROACHES

➢ How to implement in MATLAB?

Step 1: Create an excel sheet with features with class assignments
LEARNING APPROACHES

How to implement in MATLAB?

Step 2: Open MATLAB and drag the excel file to workspace.
LEARNING APPROACHES

- How to implement in MATLAB?

**Step 3:** Click Import Selection and import data
LEARNING APPROACHES

- How to implement in MATLAB?

**Step 4:** Features are in workspace and ready
LEARNING APPROACHES

- How to implement in MATLAB?

**Step 5:** Go to Apps,
- click classification learner,
- select Logistic Regression from Model Type
- click New Session,
- select from Workspace
LEARNING APPROACHES

How to implement in MATLAB?

Step 6: Set 10 fold Cross validation
- Start the session
LEARNING APPROACHES

- NONLINEAR FEATURES

LEARNING APPROACHES

2: No Stress
3: Stress

Stress Detection Algorithm

Sloke Shrestha
LEARNING APPROACHES

Stress Detection Algorithm

2: No Stress
3: Stress

Model 1 (Fine Gaussian SVM)

True class

2

27%

73%

3

24%

76%

1

27%

76%

False class

Predicted class

True Positive Rate

73%

False Negative Rate

24%
SUMMARY

ORS Research Design & Data Analysis Lab
Office of Research and Scholarship

- Identification of Features
- Develop MATLAB code for feature extraction
- Set up database
- Assist with experiment protocol and data analysis
- Machine Learning

![Diagram of data processing stages: Preprocessing, Feature Extraction, Feature Selection, leading to Statistical or Machine Learning Models]
SUMMARY

- Ready to go features and Machine Learning Models

The University of Texas at Tyler Center for Health Informatics & Analytics

Data → Preprocessing → Feature Extraction → Feature Selection

CHIA FEATURES

Statistical or Machine Learning Models
THANK YOU

Current Students:

Sloke Shrestha, UG
Mohammed Alenazi, Graduate
Pravitha Ramanand, PhD, Postdoc

Former Students:

Apurupa Amperayani (PhD Student, Arizona State University)
Jonathan Wells (PhD Student, UT Austin)
Pallavi Atluri
Keerthi Chinthia (Data Scientist, Wabtec Corporation)
Selorm Darkey (Business Intelligent Analyst, Taylor Solutions)
THANK YOU

ORC Research Design & Data Analysis Lab
Office of Research and Scholarship

SBIR: RAE (Realize, Analyze, Engage) - A digital biomarker based detection and intervention system for stress and carvings during recovery from substance abuse disorders. 
**PIs: M. Reinhardt, S. Carreiro, P. Indic**

**Pre-Vent**
National Institute Of Health Grant
**P. Indic (Analytical Core PI, UT-Tyler)**
**N. Ambal (PI, Univ. of Alabama, Birmingham)**

**Design of a wearable sensor system and associated algorithm to track suicidal ideation from movement variability and develop a novel objective marker of suicidal ideation and behavior risk in veterans.**
Clinical Science Research and Development Grant (approved for funding).
**P. Indic (site PI, UT-Tyler)**
**E.G. Smith (Project PI, VA)**
**P. Salvatore (Investigator, Harvard University)**

**Design of a wearable biosensor sensor system with wireless network for the remote detection of life threatening events in neonates.**
National Science Foundation Smart & Connected Health Grant
**P. Indic (Lead PI, UT-Tyler)**
**D. Paydarfar (Co PI, UT-Austin)**
**H. Wang (Co PI, UMass Dartmouth)**
Y. Kim (Co PI, UMass Dartmouth)

**SBIR: RAE (Realize, Analyze, Engage) - A digital biomarker based detection and intervention system for stress and carvings during recovery from substance abuse disorders.**

**P. Indic (PI, UT-Tyler)**
**M. Reinhart (PI, ContinueYou, LLC)**
**S. Carriero, (PI, Univ. of Mass. Med. School)**
DISCUSSION