MACHINE LEARNING APPROACHES USING MATLAB

PREMANANDA INDIC, PH.D.

DEPARTMENT OF ELECTRICAL ENGINEERING

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

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

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

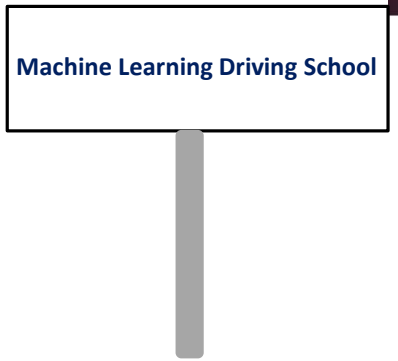
Some Statistics & Programming Knowledge Helps !

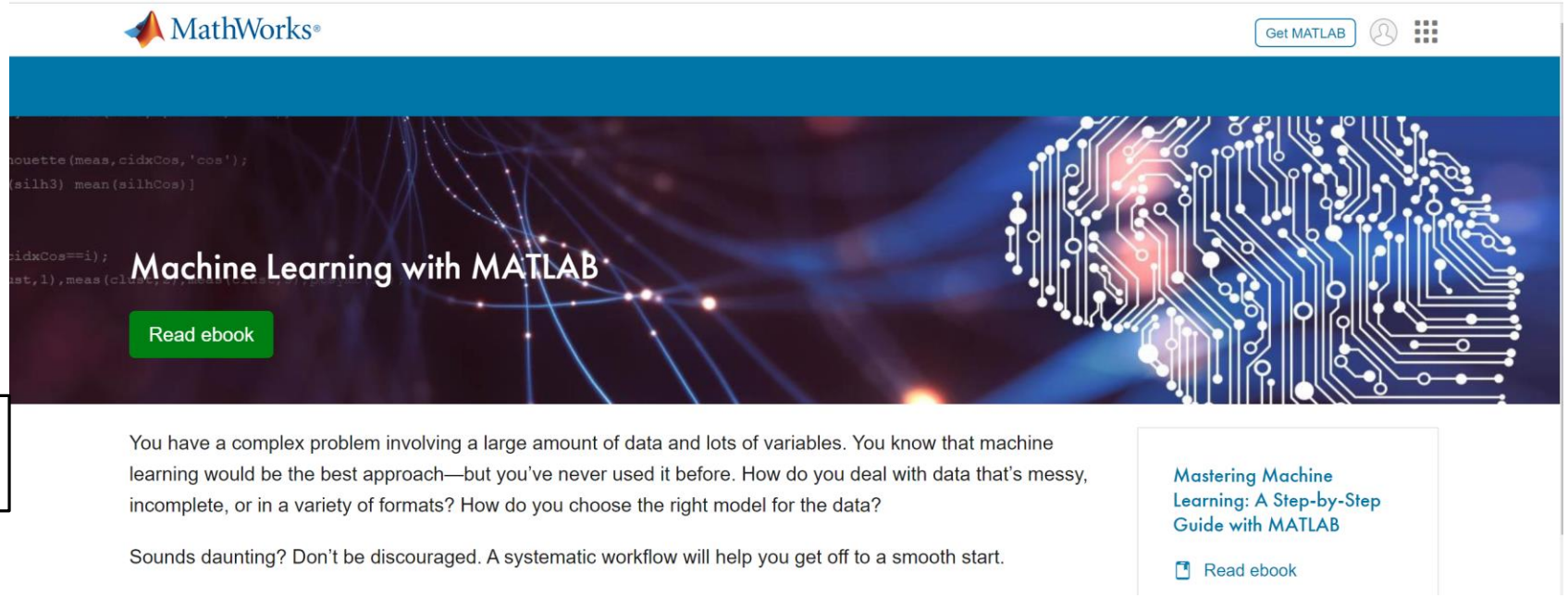
INTRODUCTION

- Machine Learning with MATLAB

https://commons.wikimedia.org/wiki/File:Man_Driving_Car_Cartoon_Vector.svg

http://clipart-library.com/mechanic-cliparts.html





INTRODUCTION

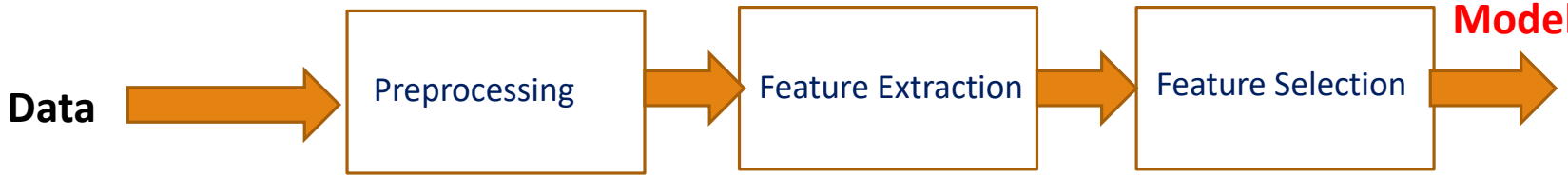
MATLAB Access for Everyone at

University of Texas at Tyler

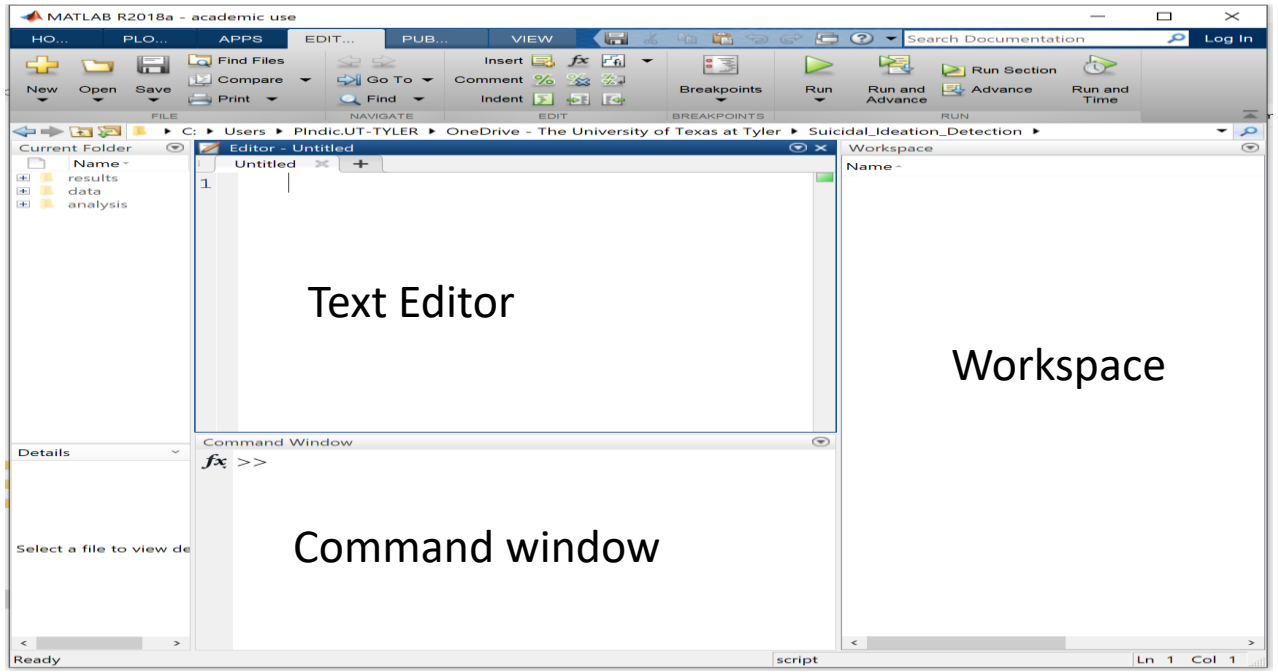
https://www.mathworks.com/academia/tah-portal/university-of-texas-at-tyler-1108545.html

30 Minute Feature Extraction

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



INTRODUCTION



30
Minute
Feature Extraction

Data → Preprocessing → Feature Extraction → Feature Selection → **Statistical or Machine Learning Models**

OUTLINE

- INTRODUCTION
- DIFFERENT MACHINE LEARNING APPROACHES
- DISCUSSION

Statistical vs. Machine Learning Models

Purpose:

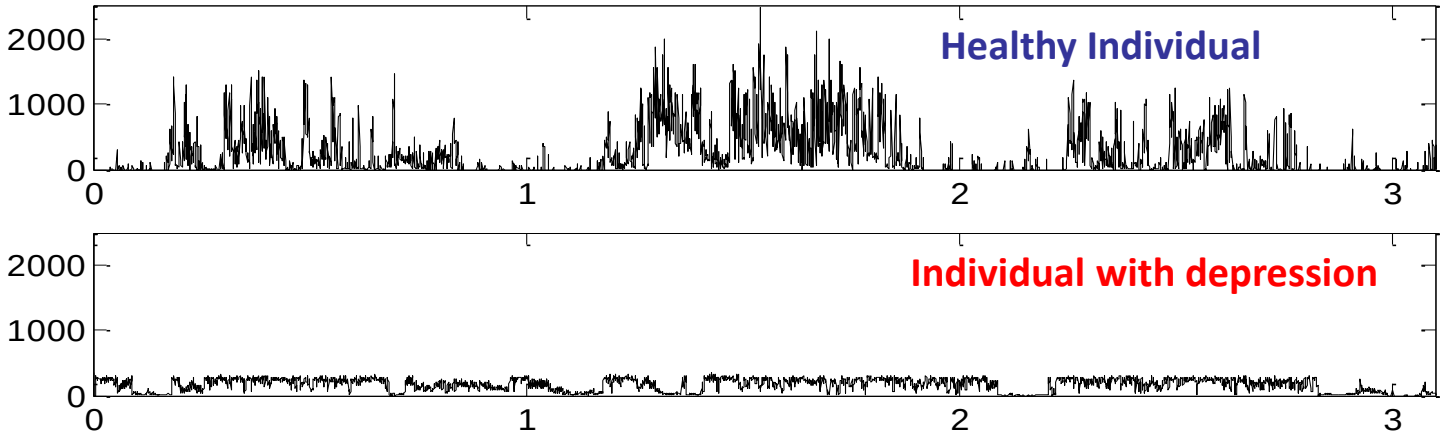
Statistical models are used for inference (To find association between features and an outcome). Results should be interpretable.

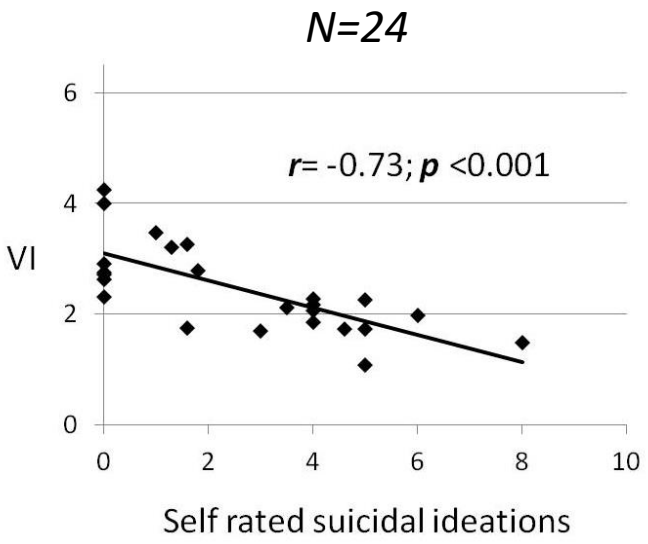
Machine Learning models are used for prediction (Use features that can predict an outcome). Results may not be interpretable.

Statistical vs. Machine Learning Models

Association vs. Prediction

Philips Actiwatch 2





$$VI = m \times SI + C$$

$$m = r\frac{\sigma_{VI}}{\sigma_{SI}}$$

$$C = \mu_{VI} - m\mu_{SI}$$

$$\widetilde{SI} = a \times VI + b$$

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

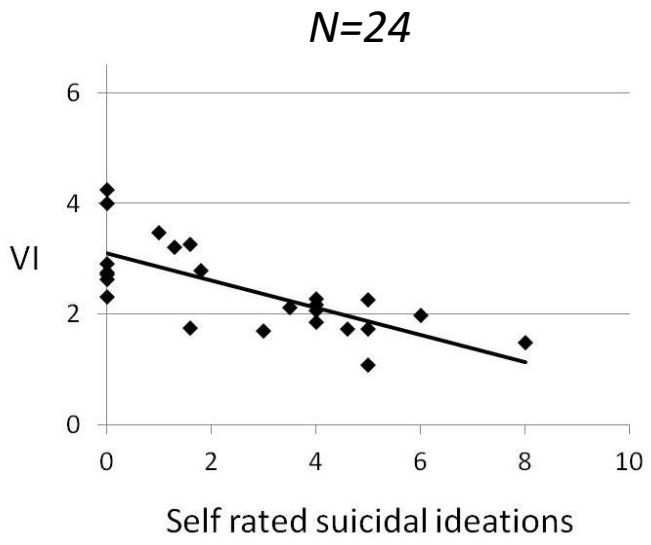
- Unsupervised Learning

  Clustering

  - Principal Component Analysis
  - Independent Component Analysis
  - Singular Value Decomposition
  - …….
  - …….

LEARNING APPROACHES

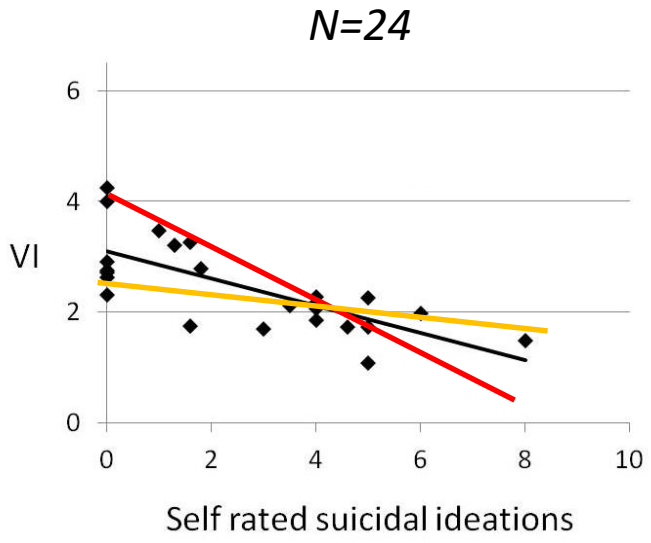
- Do machines actually “learn” ?



$$VI = m \times SI + C$$

LEARNING APPROACHES

- Do machines actually “learn” ?



$$e(N=1) = \widetilde{VI}(N=1) - VI(N=1)$$

$$e(N=2) = \widetilde{VI}(N=2) - VI(N=2)$$

……

……

$$e(N=24) = \widetilde{VI}(N=24) - VI(N=24)$$

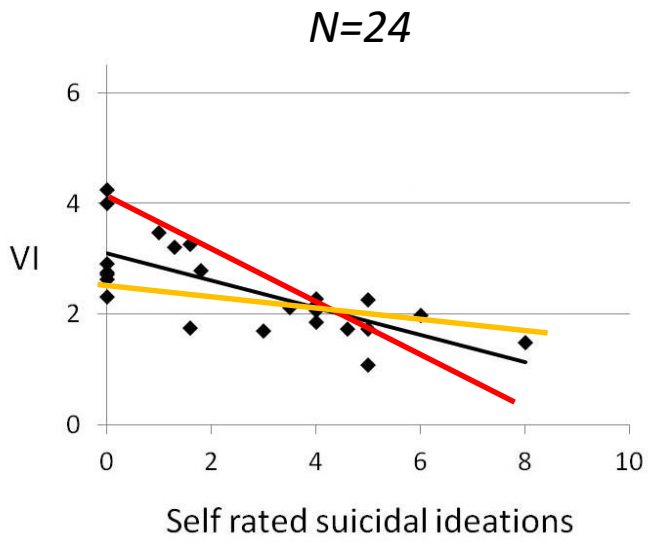
$$\widetilde{VI} = m \times SI + C$$

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

LEARNING APPROACHES

- Do machines actually “learn” ?

How do we find minimum E ?



m →

C ↓

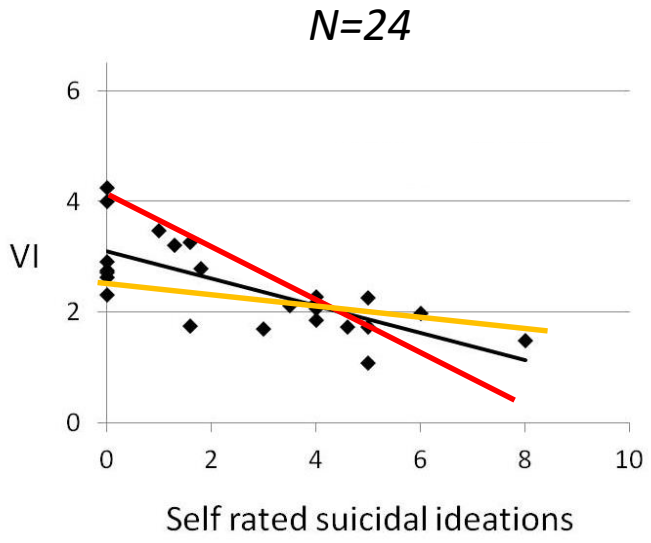
0.1	0.6	0.8	0.01	0.5
1	10	0.01	0.001	0.002
8	7	0.0006	0.03	0.55
100	12	0.1	12	0.89
2	1	2	0.5	0.05

$$\widetilde{VI} = m \times SI + C$$

LEARNING APPROACHES

- Do machines actually “learn” ?

How do we find minimum E ?



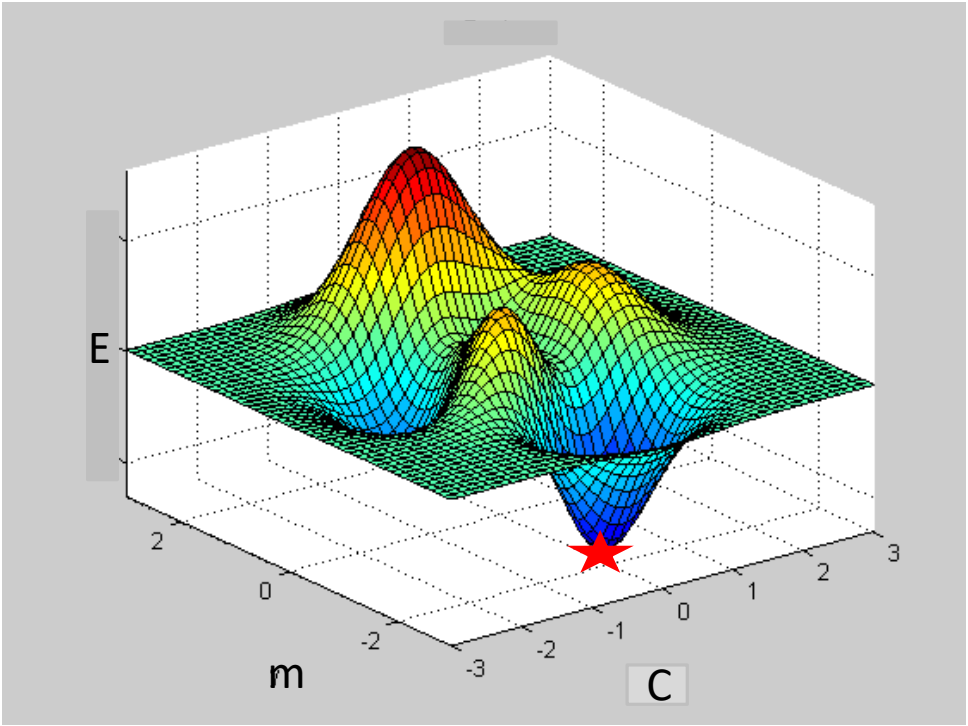
- Gradient Descent

by **Louis Augustin Cauchy** in 1847

$$\widetilde{VI} = m \times SI + C$$

Linear Regression

$$\widetilde{SI} = a \times VI + b$$



LEARNING APPROACHES

- Do machines actually “learn” ?

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

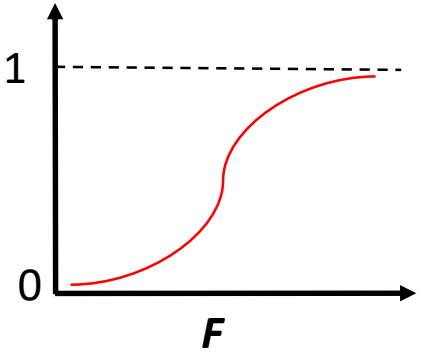
0 = Low Risk, 1 = High Risk

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

Mean
Variance
Skewness
Kurtosis
Power
Period

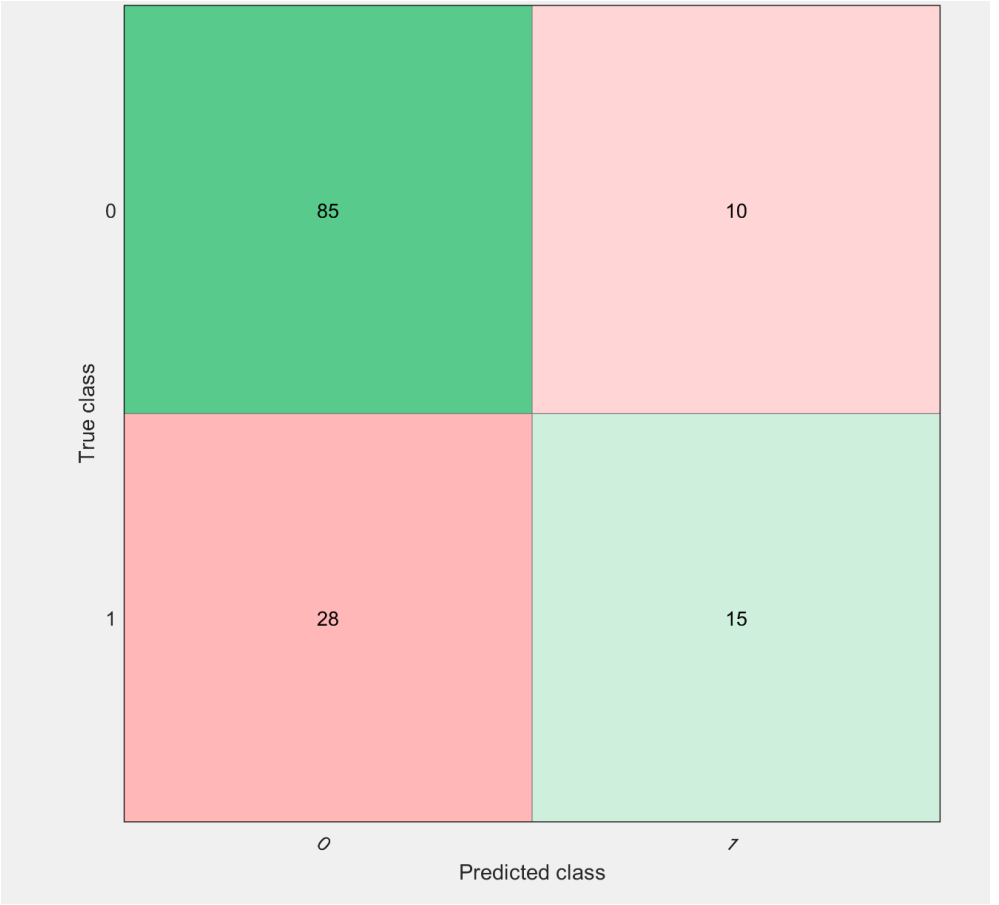
Linear Regression

$P = \frac{1}{1 + e^{-(A^T F + B)}}$



Logistic Regression

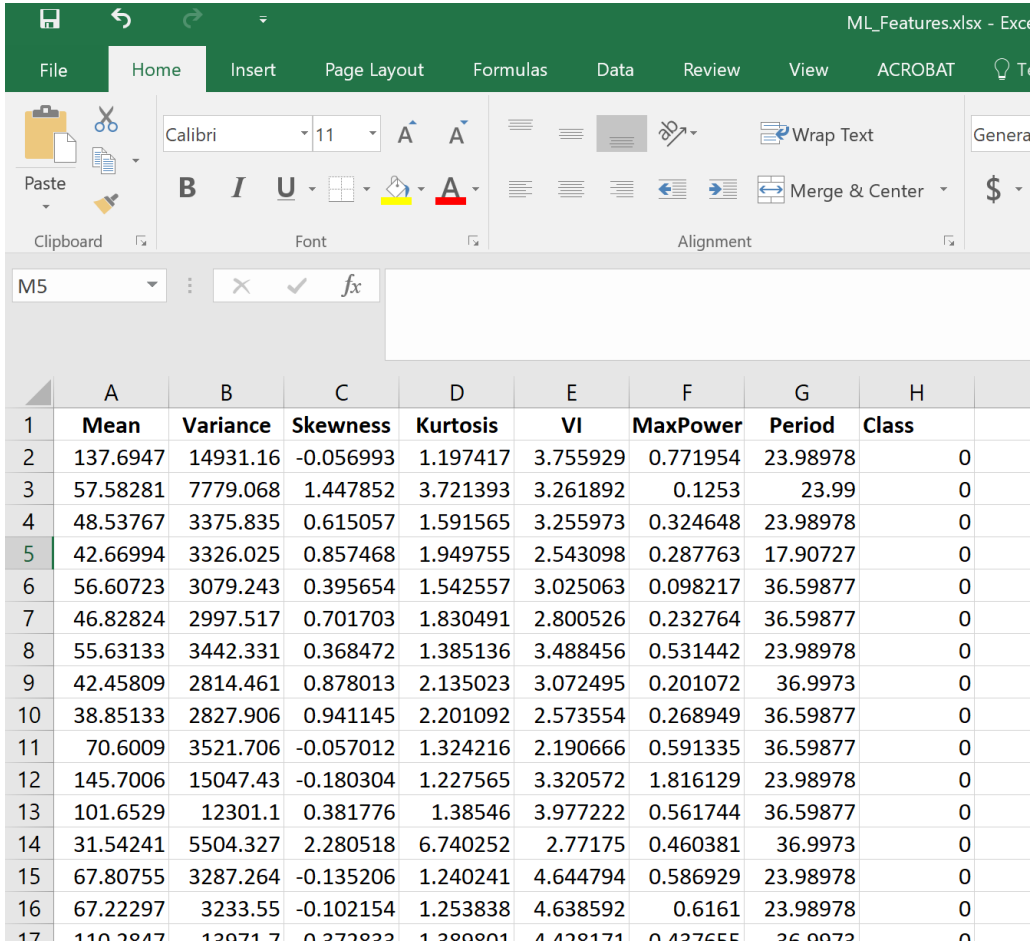
Accuracy ~73%



LEARNING APPROACHES

- How to implement in MATLAB ?

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

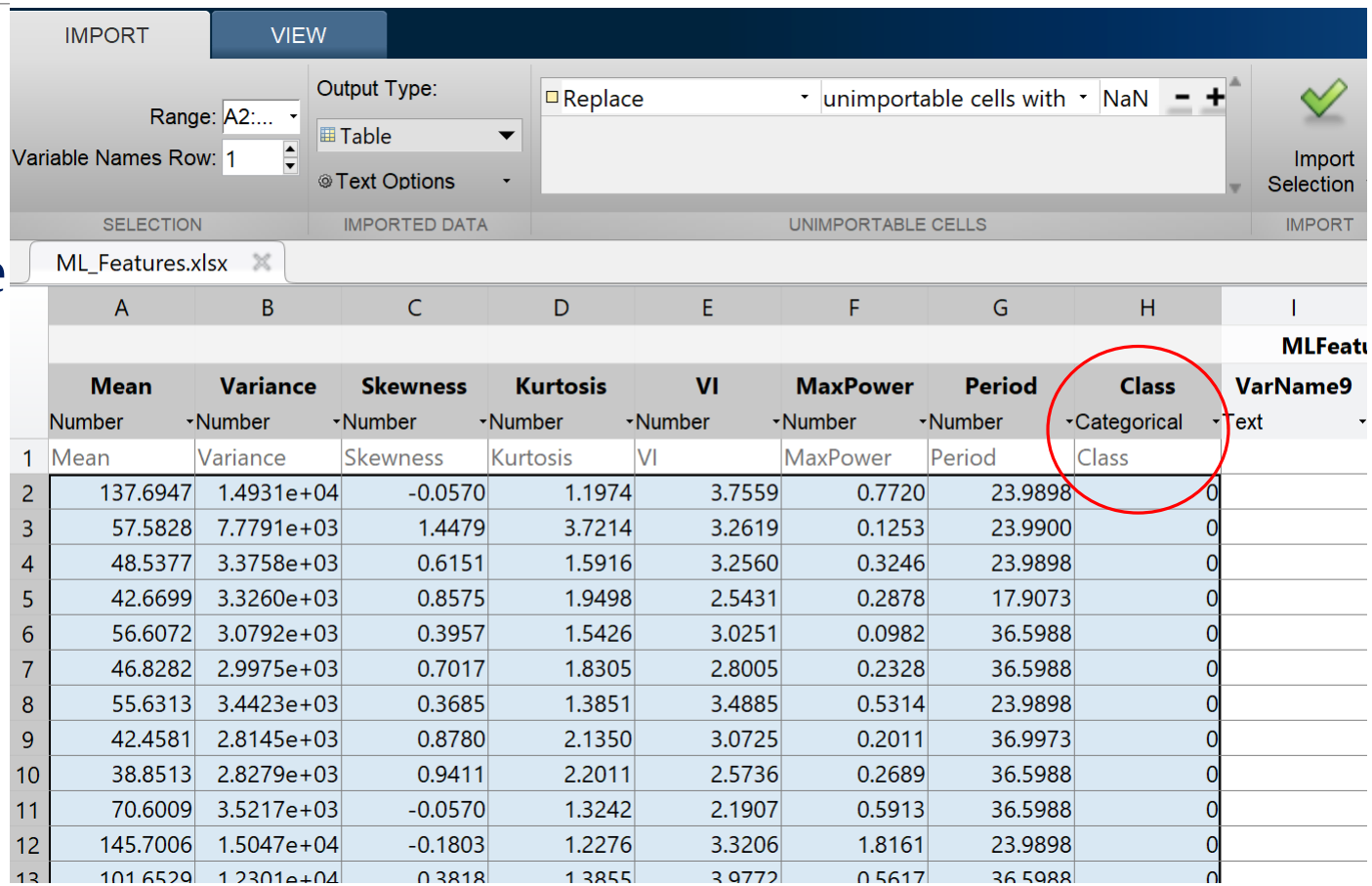


	A	B	C	D	E	F	G	H
1	Mean	Variance	Skewness	Kurtosis	VI	MaxPower	Period	Class
2	137.6947	14931.16	-0.056993	1.197417	3.755929	0.771954	23.98978	0
3	57.58281	7779.068	1.447852	3.721393	3.261892	0.1253	23.99	0
4	48.53767	3375.835	0.615057	1.591565	3.255973	0.324648	23.98978	0
5	42.66994	3326.025	0.857468	1.949755	2.543098	0.287763	17.90727	0
6	56.60723	3079.243	0.395654	1.542557	3.025063	0.098217	36.59877	0
7	46.82824	2997.517	0.701703	1.830491	2.800526	0.232764	36.59877	0
8	55.63133	3442.331	0.368472	1.385136	3.488456	0.531442	23.98978	0
9	42.45809	2814.461	0.878013	2.135023	3.072495	0.201072	36.9973	0
10	38.85133	2827.906	0.941145	2.201092	2.573554	0.268949	36.59877	0
11	70.6009	3521.706	-0.057012	1.324216	2.190666	0.591335	36.59877	0
12	145.7006	15047.43	-0.180304	1.227565	3.320572	1.816129	23.98978	0
13	101.6529	12301.1	0.381776	1.38546	3.977222	0.561744	36.59877	0
14	31.54241	5504.327	2.280518	6.740252	2.77175	0.460381	36.9973	0
15	67.80755	3287.264	-0.135206	1.240241	4.644794	0.586929	23.98978	0
16	67.22297	3233.55	-0.102154	1.253838	4.638592	0.6161	23.98978	0

LEARNING APPROACHES

- How to implement in MATLAB ?

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

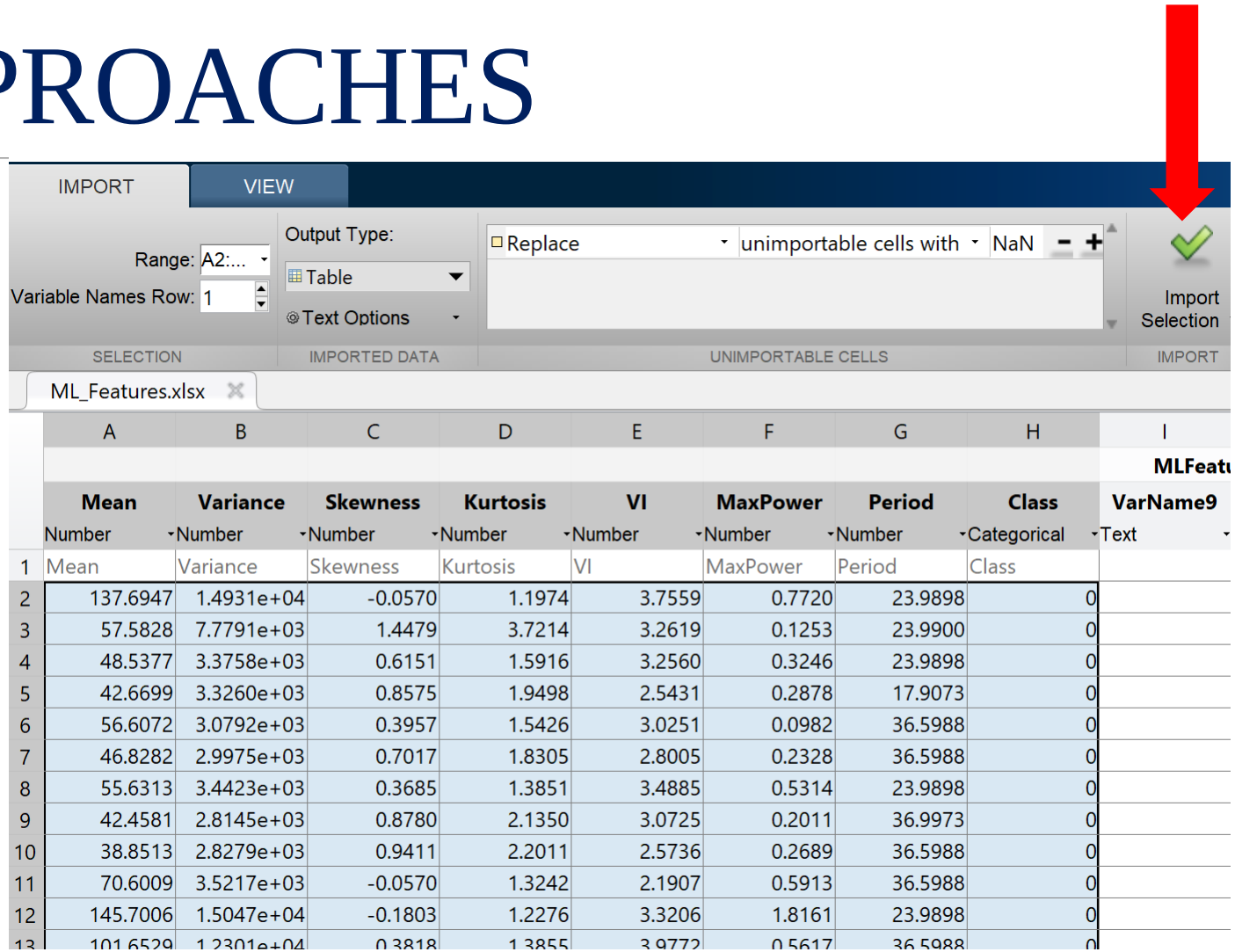


	A	B	C	D	E	F	G	H	I
									MLFeatu
	Mean	Variance	Skewness	Kurtosis	VI	MaxPower	Period	Class	VarName9
	Number	Number	Number	Number	Number	Number	Number	Categorical	Text
1	Mean	Variance	Skewness	Kurtosis	VI	MaxPower	Period	Class	
2	137.6947	1.4931e+04	-0.0570	1.1974	3.7559	0.7720	23.9898	0	
3	57.5828	7.7791e+03	1.4479	3.7214	3.2619	0.1253	23.9900	0	
4	48.5377	3.3758e+03	0.6151	1.5916	3.2560	0.3246	23.9898	0	
5	42.6699	3.3260e+03	0.8575	1.9498	2.5431	0.2878	17.9073	0	
6	56.6072	3.0792e+03	0.3957	1.5426	3.0251	0.0982	36.5988	0	
7	46.8282	2.9975e+03	0.7017	1.8305	2.8005	0.2328	36.5988	0	
8	55.6313	3.4423e+03	0.3685	1.3851	3.4885	0.5314	23.9898	0	
9	42.4581	2.8145e+03	0.8780	2.1350	3.0725	0.2011	36.9973	0	
10	38.8513	2.8279e+03	0.9411	2.2011	2.5736	0.2689	36.5988	0	
11	70.6009	3.5217e+03	-0.0570	1.3242	2.1907	0.5913	36.5988	0	
12	145.7006	1.5047e+04	-0.1803	1.2276	3.3206	1.8161	23.9898	0	
13	101.6529	1.2301e+04	0.3818	1.3855	3.9772	0.5617	36.5988	0	

LEARNING APPROACHES

- How to implement in MATLAB ?

Step 3: Click Import Selection and import data

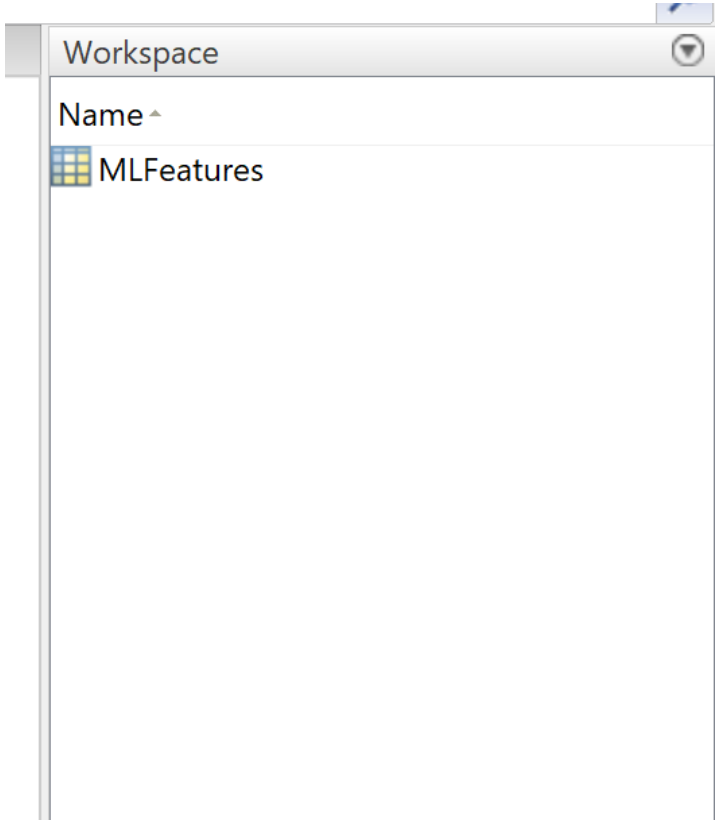


	A	B	C	D	E	F	G	H	I
									MLFeatu
	Mean	Variance	Skewness	Kurtosis	VI	MaxPower	Period	Class	VarName9
	Number	Number	Number	Number	Number	Number	Number	Categorical	Text
1	Mean	Variance	Skewness	Kurtosis	VI	MaxPower	Period	Class	
2	137.6947	1.4931e+04	-0.0570	1.1974	3.7559	0.7720	23.9898	0	
3	57.5828	7.7791e+03	1.4479	3.7214	3.2619	0.1253	23.9900	0	
4	48.5377	3.3758e+03	0.6151	1.5916	3.2560	0.3246	23.9898	0	
5	42.6699	3.3260e+03	0.8575	1.9498	2.5431	0.2878	17.9073	0	
6	56.6072	3.0792e+03	0.3957	1.5426	3.0251	0.0982	36.5988	0	
7	46.8282	2.9975e+03	0.7017	1.8305	2.8005	0.2328	36.5988	0	
8	55.6313	3.4423e+03	0.3685	1.3851	3.4885	0.5314	23.9898	0	
9	42.4581	2.8145e+03	0.8780	2.1350	3.0725	0.2011	36.9973	0	
10	38.8513	2.8279e+03	0.9411	2.2011	2.5736	0.2689	36.5988	0	
11	70.6009	3.5217e+03	-0.0570	1.3242	2.1907	0.5913	36.5988	0	
12	145.7006	1.5047e+04	-0.1803	1.2276	3.3206	1.8161	23.9898	0	
13	101.6529	1.2301e+04	0.3818	1.3855	3.9772	0.5617	36.5988	0	

LEARNING APPROACHES

- How to implement in MATLAB ?

Step 4: Features are in workspace and ready



LEARNING APPROACHES

- How to implement in MATLAB ?

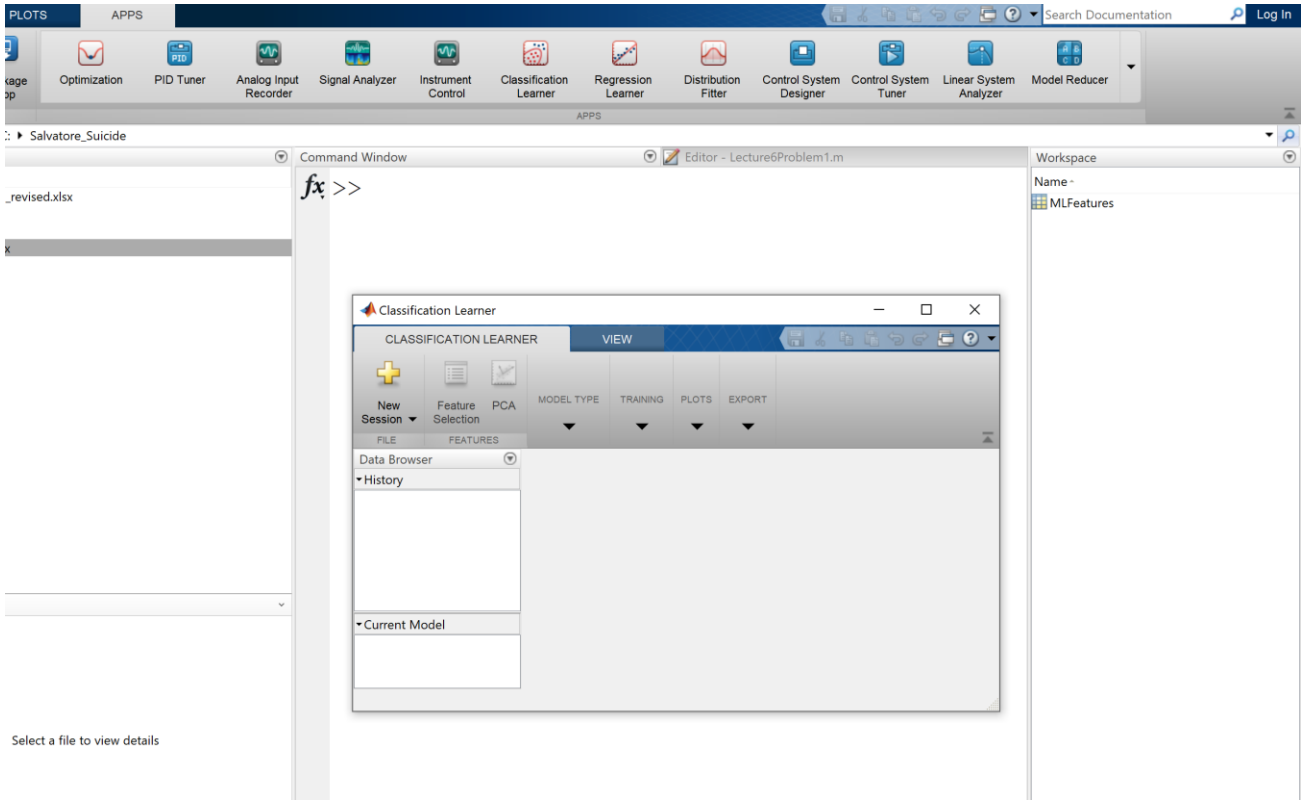
Step 5: Go to Apps,

-click classification learner,

- -select Logistic Regression from Model Type

-click New Session,

-select from Workspace

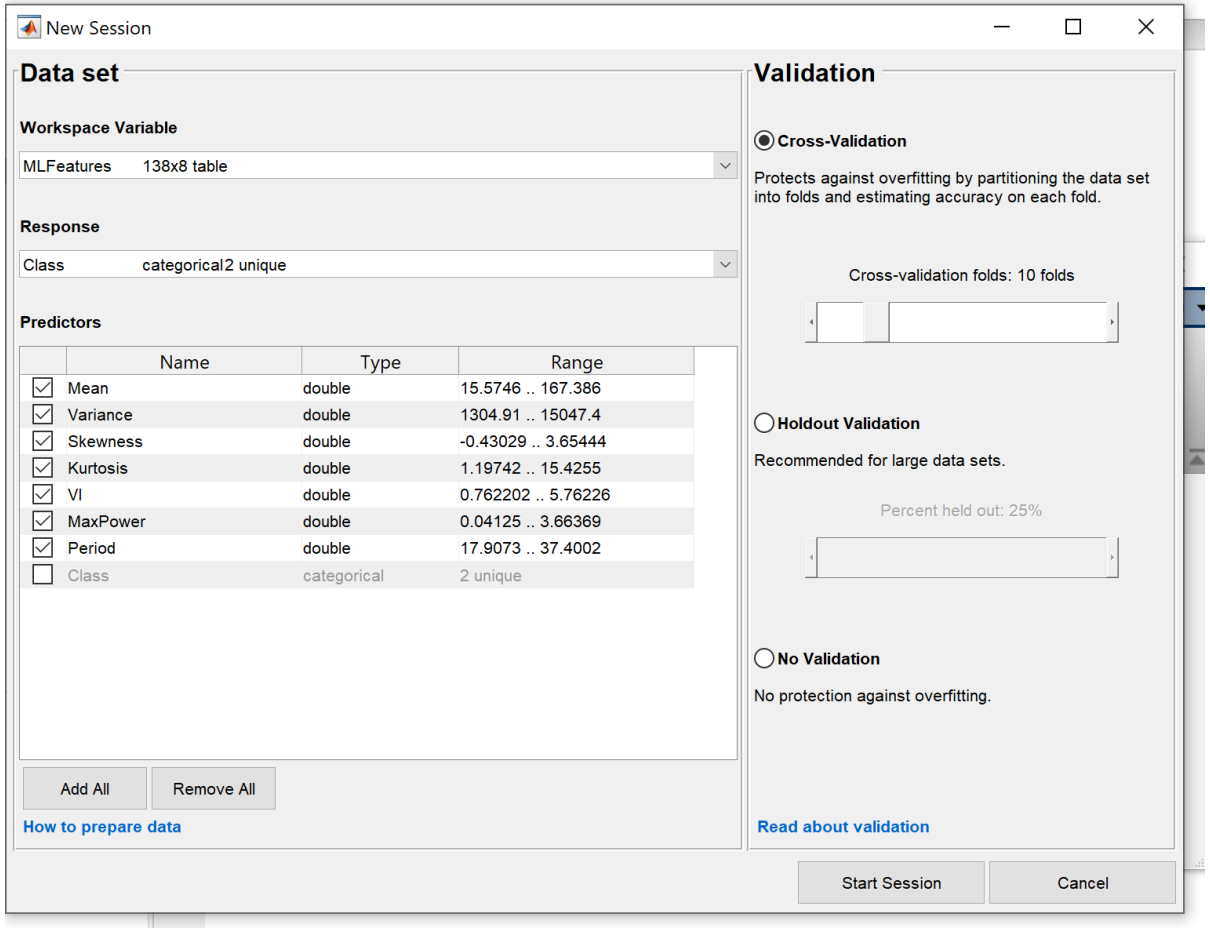


LEARNING APPROACHES

- How to implement in MATLAB ?

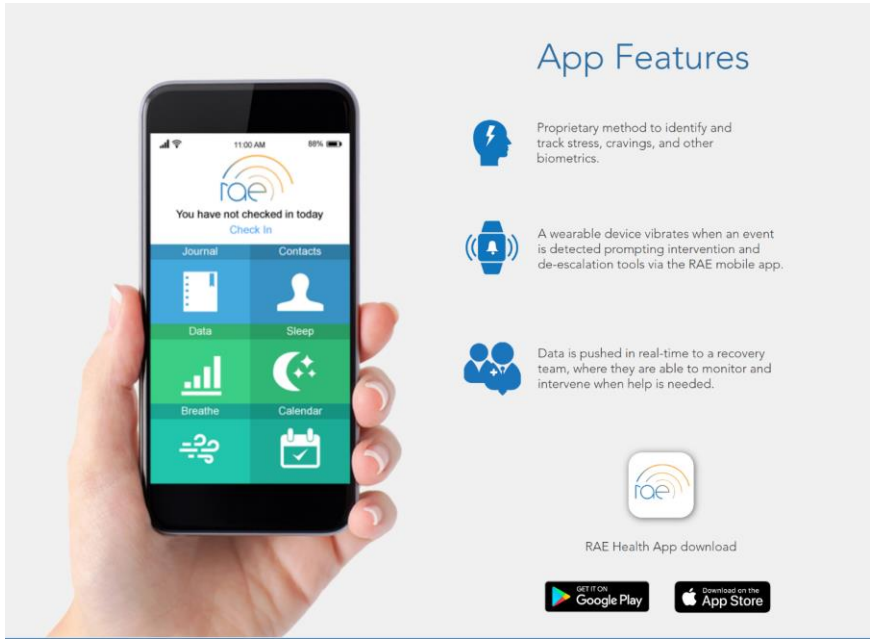
Step 6: Set 10 fold Cross validation

- Start the session

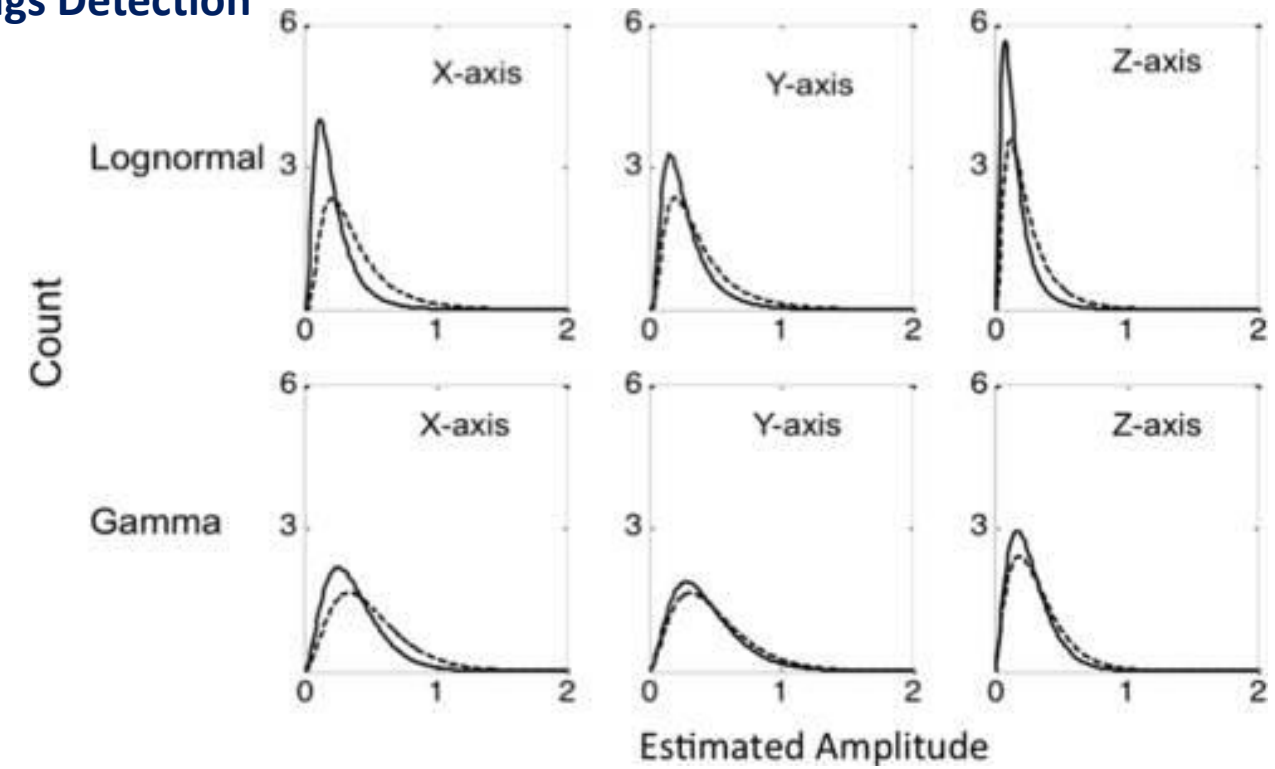


LEARNING APPROACHES

- NONLINEAR FEATURES



Cravings Detection

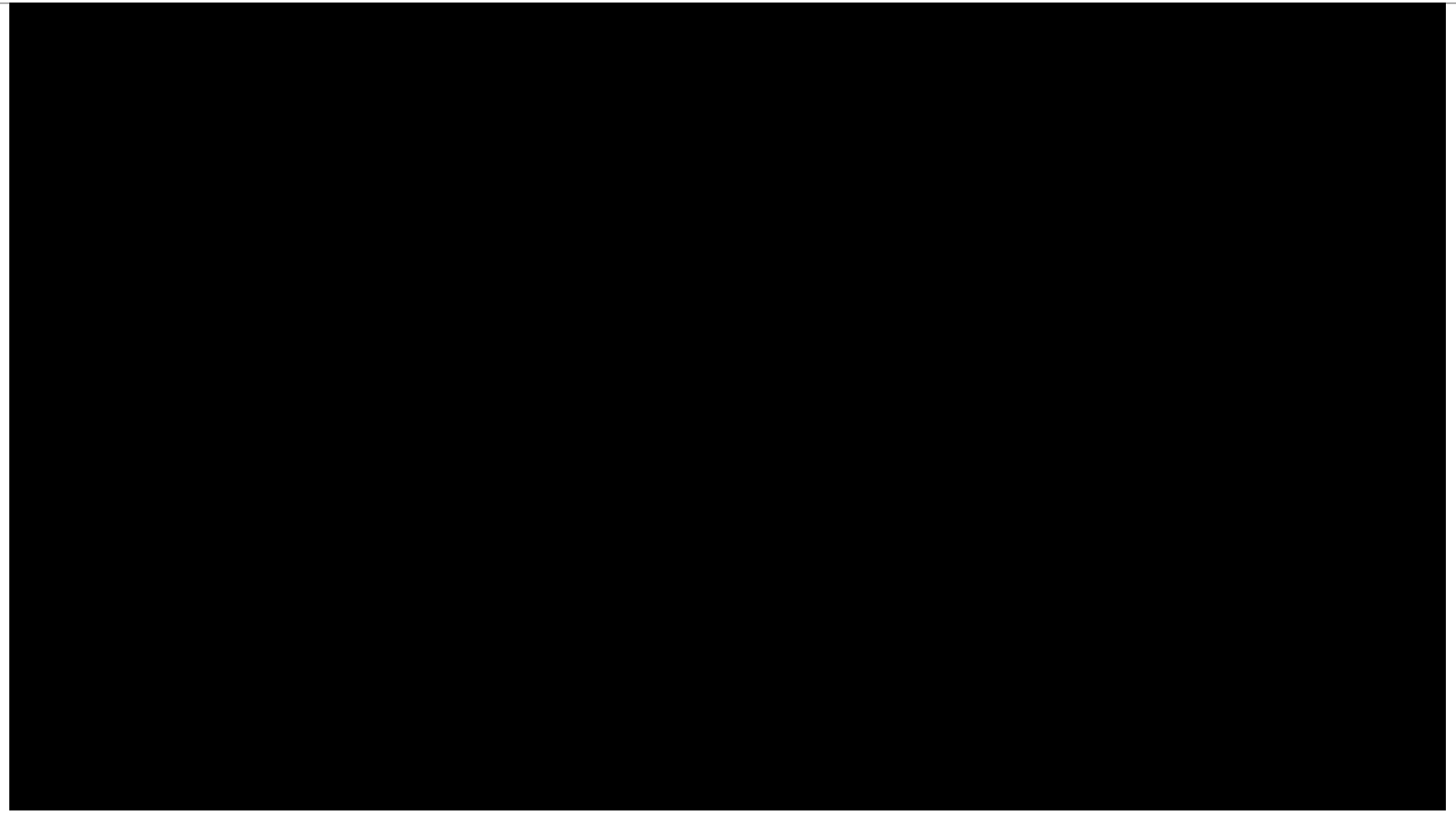


Carreiro, S, Chintha KK, **Shrestha S**, Chapman B, Smelson D, Indic P. Wearable sensor based detection of stress and craving in patients during treatment for substance use disorder: A mixed methods pilot study. Drug and Alcohol Dependence. 2020, 107929

LEARNING APPROACHES

2 : No Stress

3: Stress



Stress Detection Algorithm

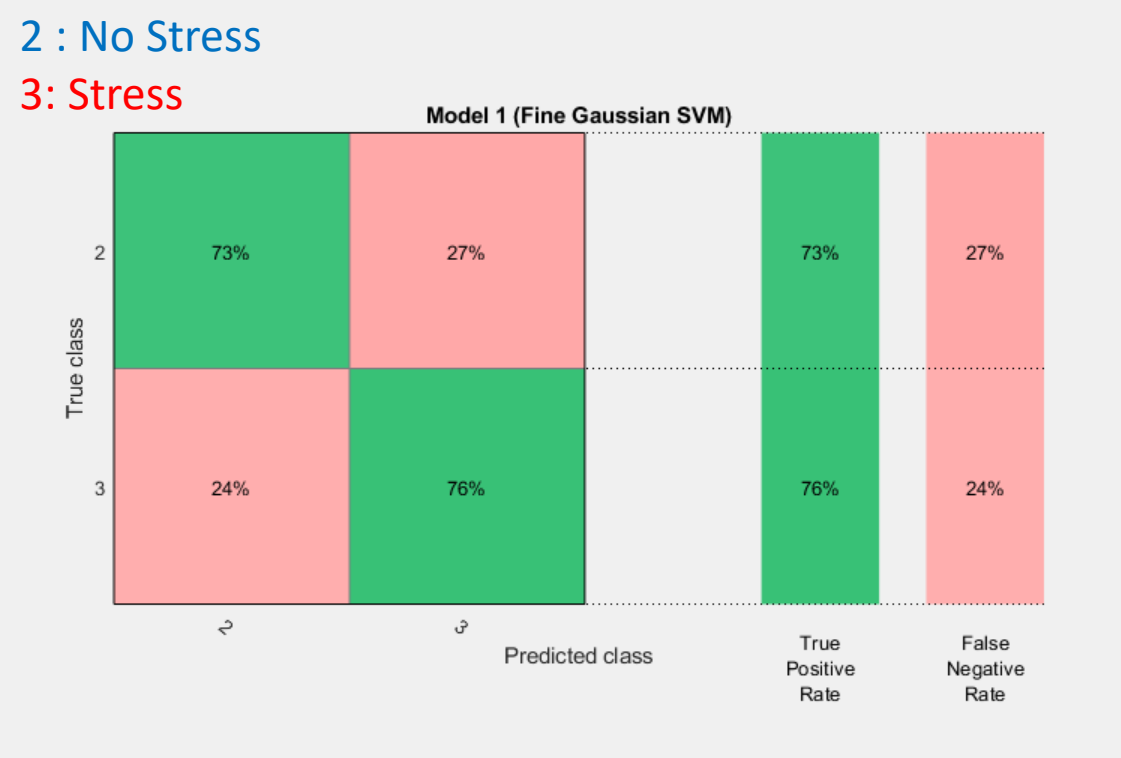
Sloke Shrestha

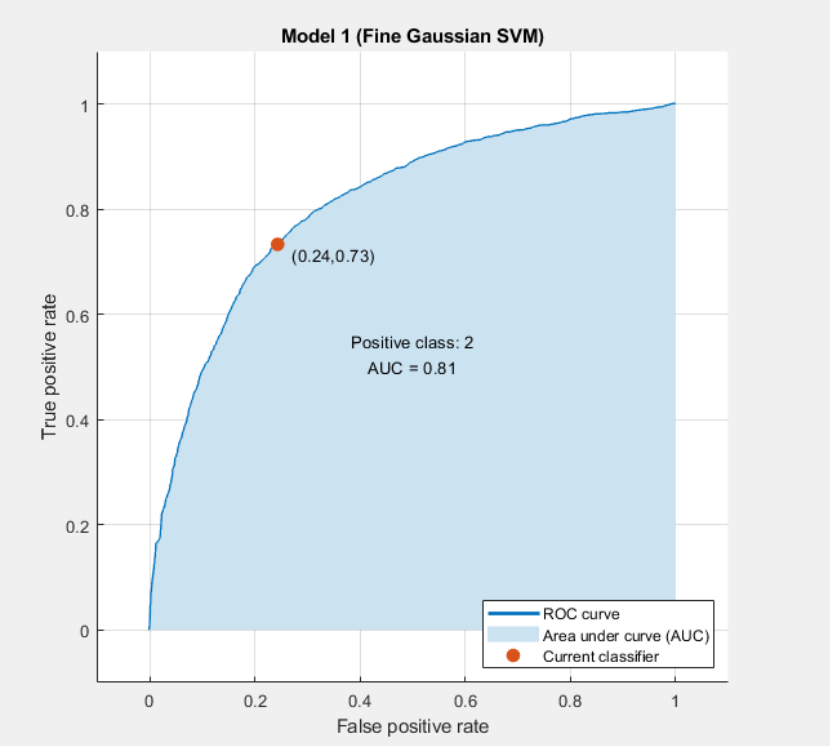
LEARNING APPROACHES

Stress Detection Algorithm

2 : No Stress

3: Stress





SUMMARY

ORS Research Design & Data Analysis Lab

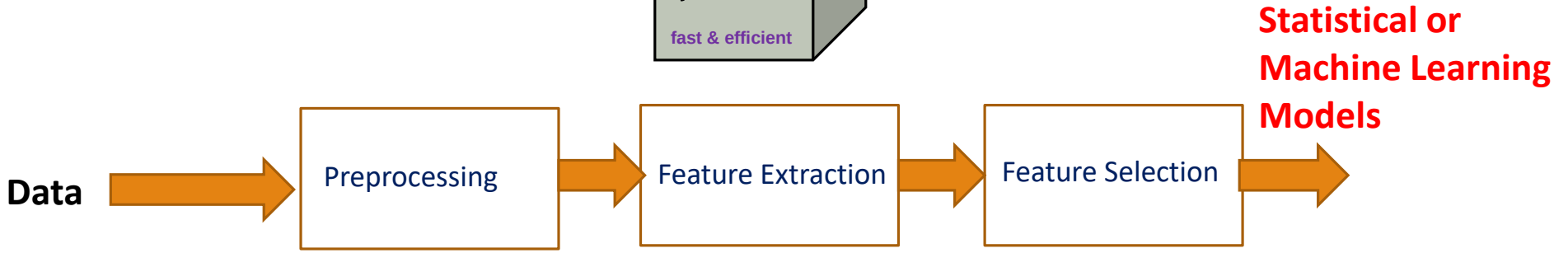
Office of Research and Scholarship

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

SUMMARY

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

- **Ready to go features and Machine Learning Models**



THANK YOU

Current Students:

Sloke Shrestha, UG

Mohammed Alenazi, Graduate

Pravitha Ramanand, PhD, Postdoc

Joshua Stapp, Graduate

Former Students:

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

THANK YOU

ORS 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

STARs Award

The University of Texas System
P. Indic (PI, UT Tyler)

Department of Veterans Affairs

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)

Pre-Vent

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

Wearable system for the detection of addiction
P. Indic (PI, UT-Tyler)
M. Reinhart (PI, ContinueYou, LLC
S. Carriero, (PI. Univ. of Mass. Med. School)

DISCUSSION