SIGNAL & DATA ANALYTICS IN IoMT
Tech-in-Med Summer Camp

PREMANANDA INDIC, PH.D.

DEPARTMENT OF ELECTRICAL ENGINEERING

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

NSF Award OAC-1924117: Easy-Med: Interdisciplinary Training in Security, Privacy-Assured Internet of Medical Things

OUTLINE

1. Different physiological signals
2. Features of the signals associated with health
3. Differentiating signals and data
5. Development of algorithms
6. Processing of signals
7. Data analytics (Machine Learning)
8. Converting algorithms into software code
9. Embedding the code in the sensors.

MACHINE LEARNING

➢ 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

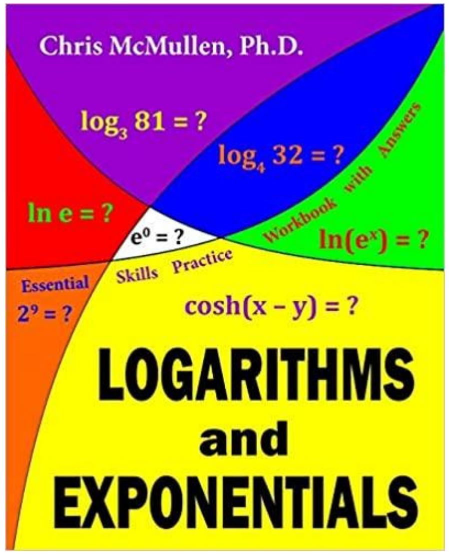
MACHINE LEARNING

- Too many books spoil the curiosity

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

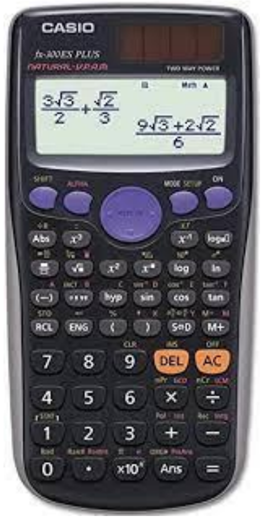
Some Statistics & Programming Knowledge Helps !

MACHINE LEARNING



Analytical Tools

Simple Calculator (Boolean Algebra)



Scientific Calculator (Series Expansion, Boolean Algebra)

Computer (Programming Language, Assembly Language, Series Expansion, Boolean Algebra)

Smart Devices (ML Models, Programming Language, Assembly Language, Series Expansion, Boolean Algebra)

MACHINE LEARNING

- Always there is a mathematical foundation

  Analytical Tools (Logarithm, Laplace Transform, Fourier Transform…….)

  Computational Tools (Boolean Algebra, Taylor Series Expansion,……..)

  Programming Languages (Basic, Fortran, C, C++, Java, …….)

  Assembly Languages (depending upon the computer processors)

  Machine Learning Models

  Artificial Intelligence

MACHINE LEARNING

- Examples of Smart Systems

  Voice Recognition

  Tumor Detection

  Weather Forecast

  Driverless Cars

MACHINE LEARNING

- Training Data
- Appropriate Model
- Procedure to Train (Make a machine to “learn”)

(Learning Algorithms, Online vs Batch Learning, Instance Based vs Model Based

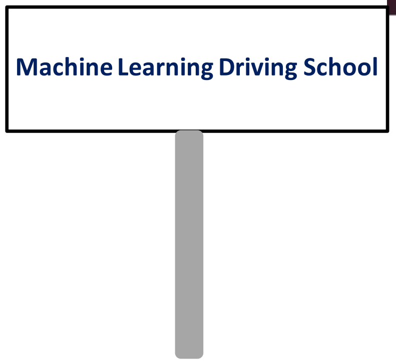
- Test Data

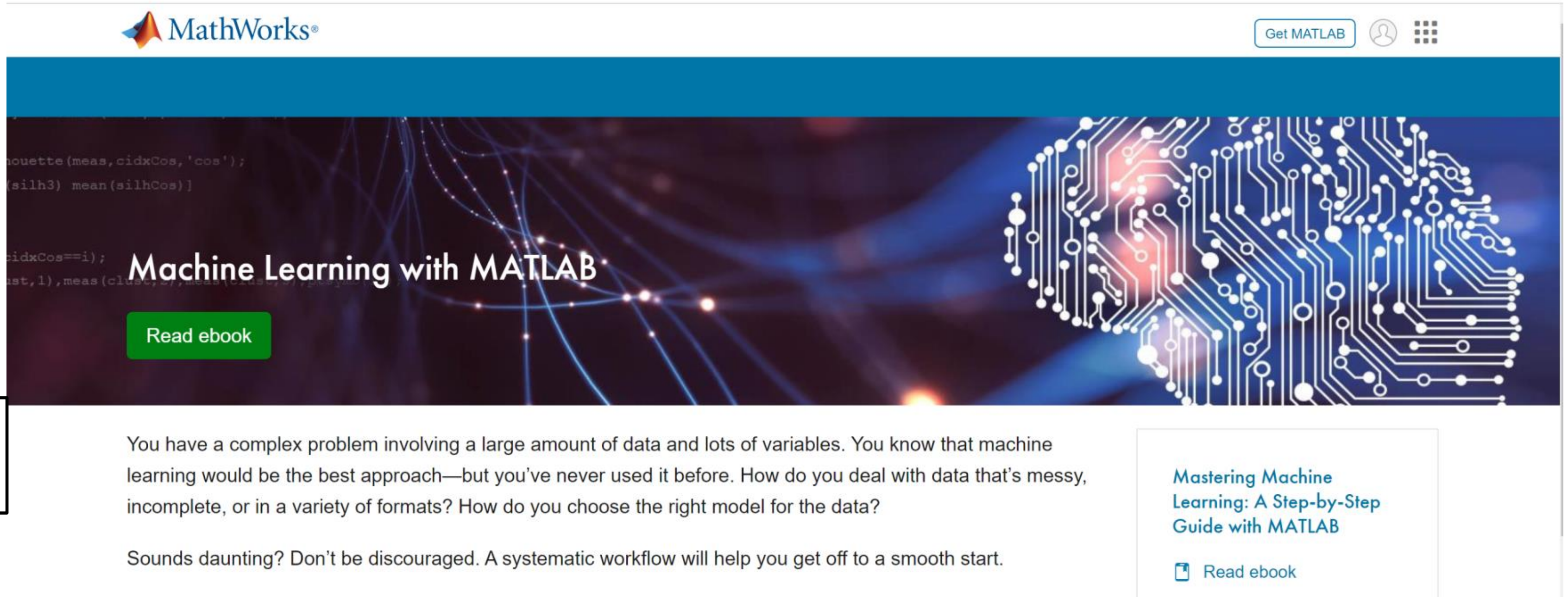
MACHINE LEARNING

- Machine Learning with MATLAB

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

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



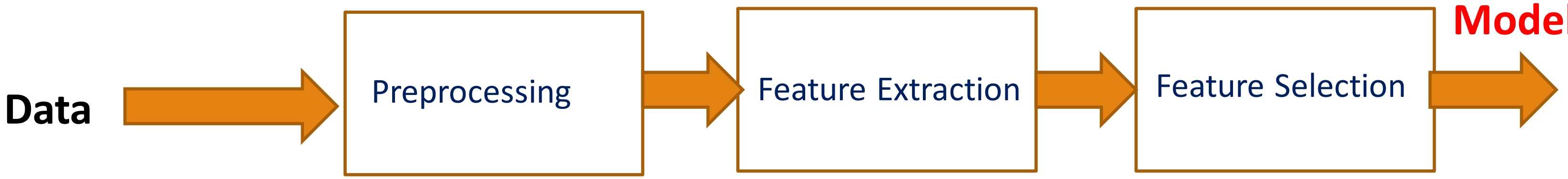


MACHINE LEARNING

- ARTIFACT
- NOISE
- SIGNAL

- VARIABLE
- STATISTICAL FEATURES
- SPECTRAL FEATURES
- NONLINEAR FEATURES

Statistical or Machine Learning Models



DIFFERENT TYPES OF FEATURES

➢ VARIABLE

fitlm

Weight	Acceleration	MPG
3504	12	18
3693	11.5	15
3436	11	18
3433	12	16
3449	10.5	17

```
lm = fitlm(tbl,'MPG~Weight+Acceleration')
```

```
lm = 


Linear regression model:
    MPG ~ 1 + Weight + Acceleration

Estimated Coefficients:
                     Estimate        SE          tStat       pValue  
                    __________    __________    _______    __________

    (Intercept)        45.155        3.4659     13.028     1.6266e-22
    Weight         -0.0082475    0.00059836    -13.783     5.3165e-24
    Acceleration      0.19694       0.14743     1.3359        0.18493
```

MPG = a + b Weight + c Acceleration

```
Number of observations: 94, Error degrees of freedom: 91
Root Mean Squared Error: 4.12
R-squared: 0.743,  Adjusted R-Squared 0.738
```

DIFFERENT TYPES OF FEATURES

- VARIABLE

Real Estate Data

Command Window

```
>> lm=fitlm(housing)

lm =


Linear regression model:
    median_house_value ~ [Linear formula with 9 terms in 8 predictors]

Estimated Coefficients:
```

	Estimate	SE	tStat	pValue
(Intercept)	-3.5854e+06	62901	-57.001	0
longitude	-42730	717.09	-59.588	0
latitude	-42510	676.95	-62.796	0
housing_median_age	1157.9	43.389	26.687	2.9463e-154
total_rooms	-8.2497	0.79426	-10.387	3.2948e-25
total_bedrooms	113.82	6.9306	16.423	3.1889e-60
population	-38.386	1.0841	-35.407	1.4597e-266
households	47.701	7.5466	6.3209	2.6535e-10
median_income	40298	337.21	119.5	0

Workspace

Name

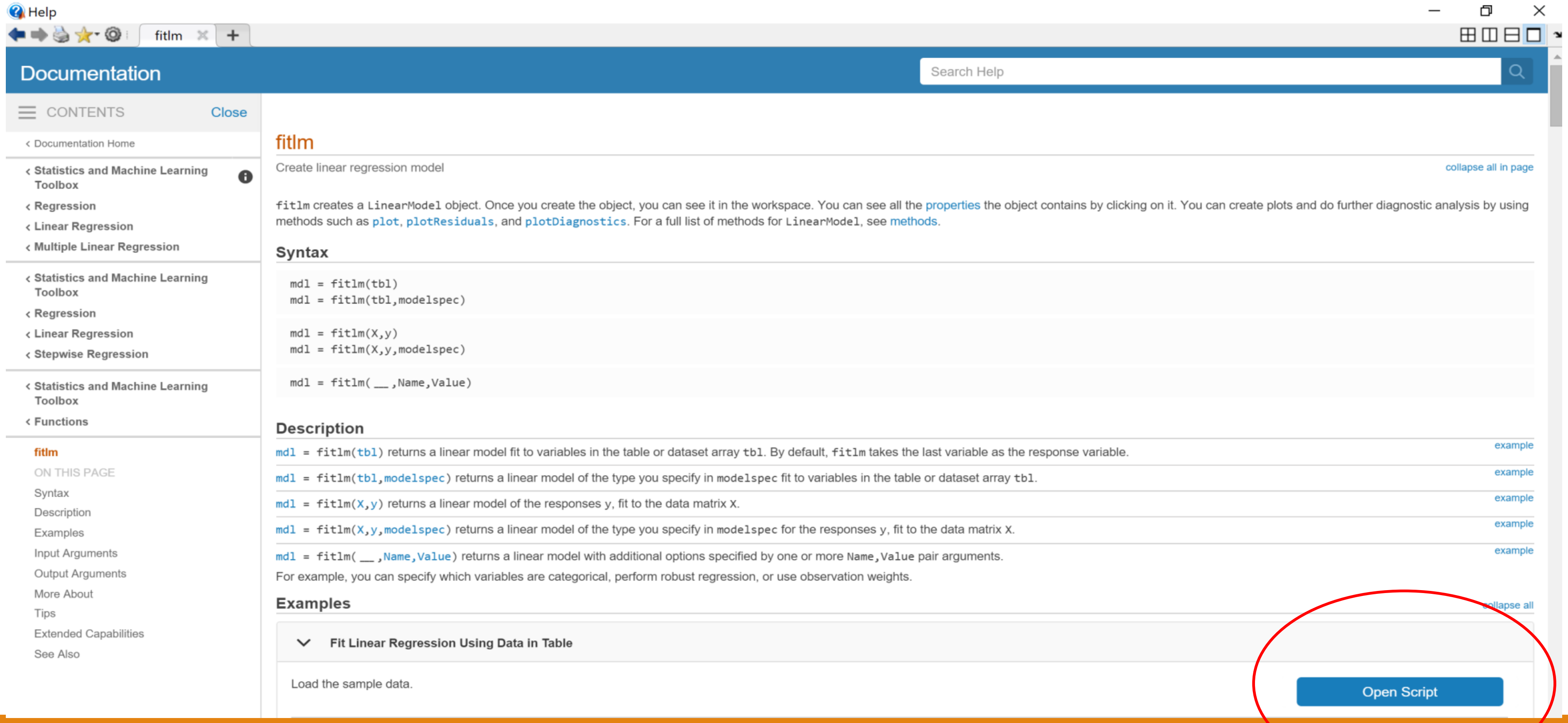
housing

lm

DIFFERENT TYPES OF FEATURES

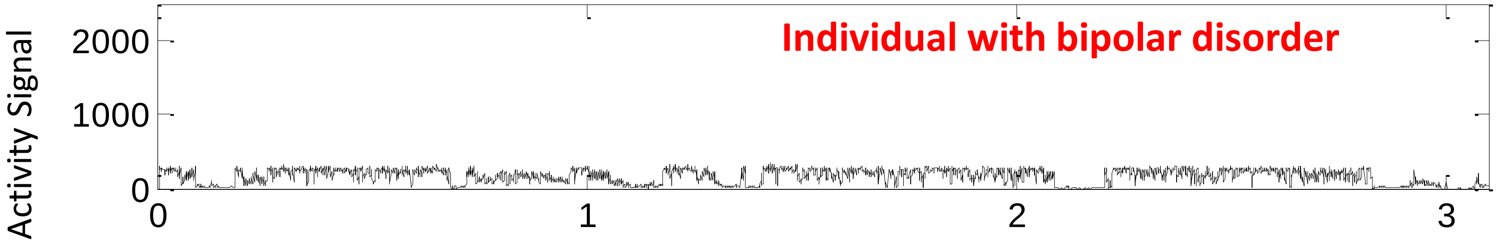
- VARIABLE

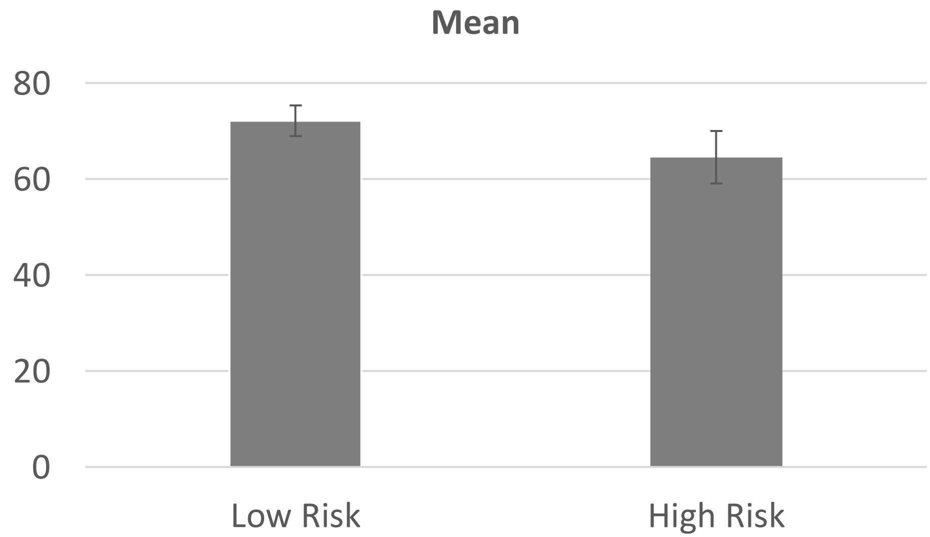
Don’t want to write the code?



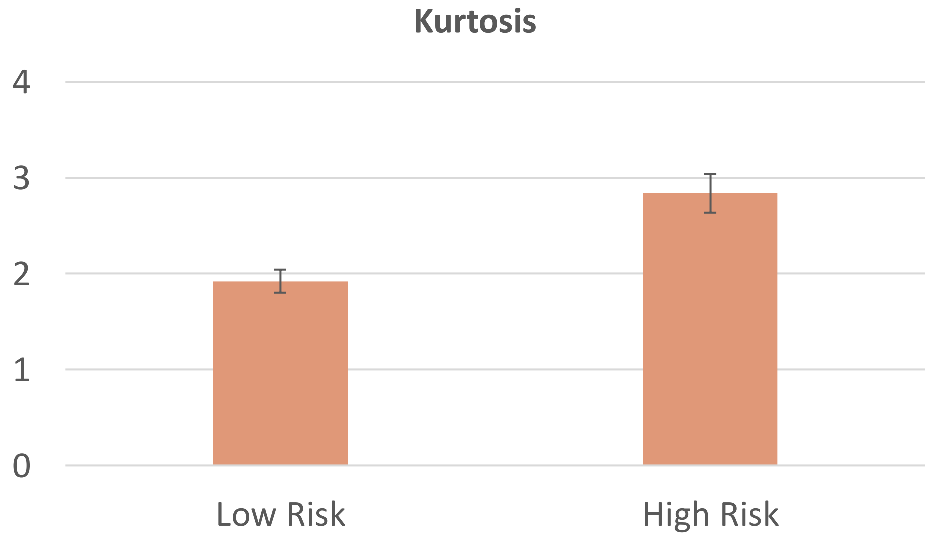
DIFFERENT TYPES OF FEATURES

- STATISTICAL FEATURES



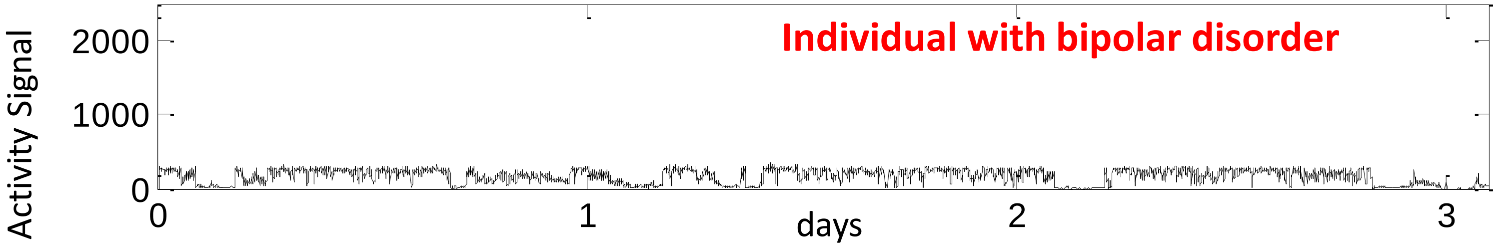


N =128



DIFFERENT TYPES OF FEATURES

➢ STATISTICAL FEATURES



Correlation with Self Reported Suicidal Ideation

Mean : ***r*** = -0.17 ***p*** = 0.05

Variance : ***r*** = -0.05 ***p*** = 0.53

Skewness : ***r*** = 0.23 ***p*** = 0.007

Kurtosis : ***r*** = 0.18 ***p*** = 0.03

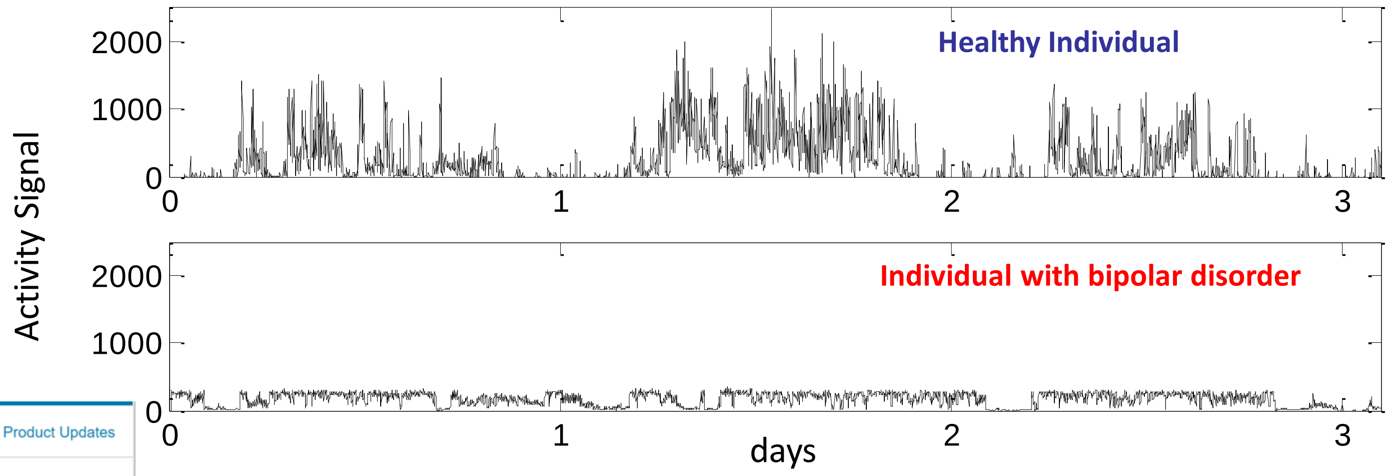
MATLAB functions

mean(filename)
variance(filename)
skewness (filename)
kurtosis(filename)
[r,p]=corr(resultsfilename(:,1),resultsfilename(:,5));

Atluri, Salvatore and Indic (unpublished)

DIFFERENT TYPES OF FEATURES

- SPECTRAL FEATURES



Documentation | Examples | Functions | Apps | Videos | Answers | Trial Software | Product Updates

Spectral Features

R2020b

Spectral features provide frequency-domain metrics on your data. To compute spectral features, you must already have a power spectrum or an order spectrum variable.

Spectrum

- **Spectrum** — Choose from the available spectrum variables. The software brings up the plot of that variable for reference, and converts the plot from log scale to linear scale.

Spectral Peaks

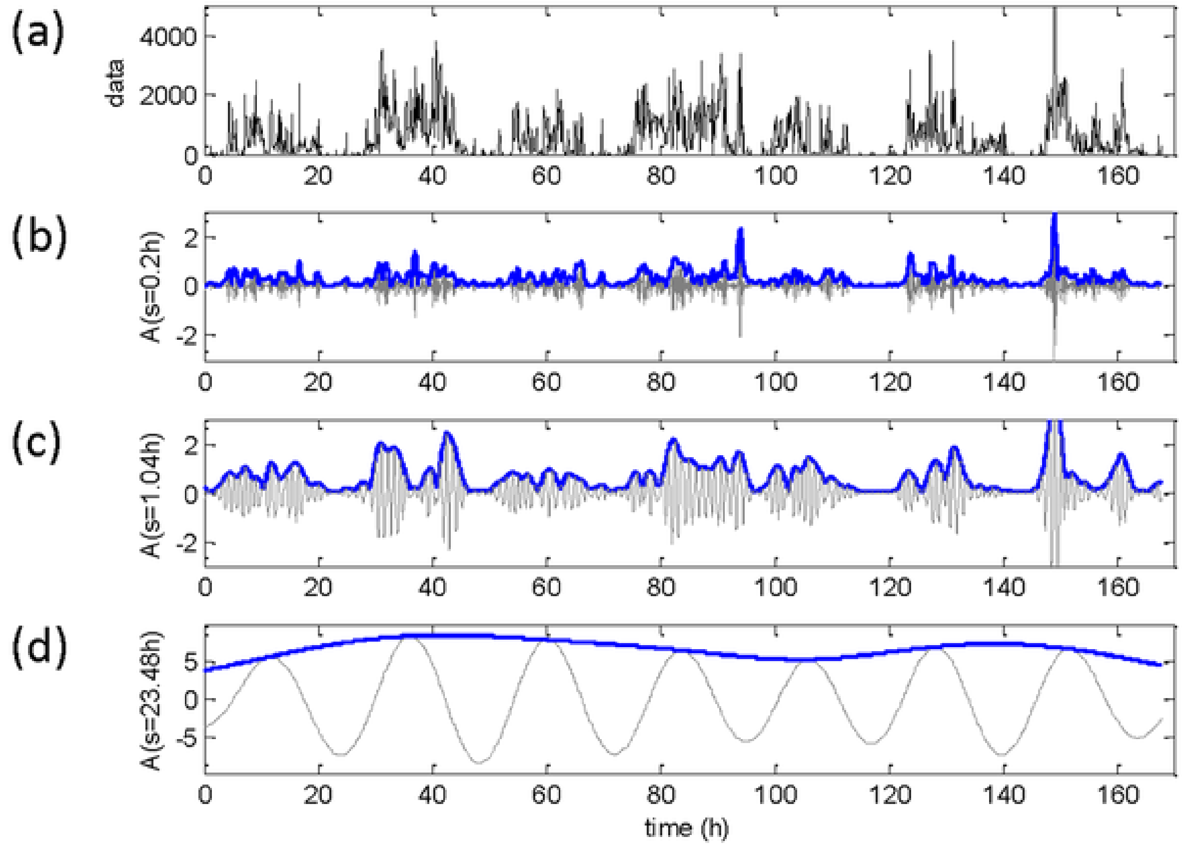
- **Peak amplitude** — Generate a feature based on the amplitude of the peaks.
- **Peak frequency** — Generate a feature based on the frequency of the peaks.
- **Peak value lower threshold** — Constrain peak size to exclude low-amplitude peaks. For more information, in `findpeaks`, see the `MinPeakHeight` name-value pair argument.
- **Number of peaks** — Number of peaks to generate features for. The software selects N most prominent peaks in the chosen frequency band, going in the descending amplitude order. For more information, in `findpeaks`, see the `NPeaks` name-value pair argument.
- **Minimum frequency gap** — Specify a minimum frequency gap. If the gap between two peaks is less than this specification, the software ignores the smaller peak of the pair. For more information, in `findpeaks`, see the `MinPeakDistance` name-value pair argument.
- **Peak excursion tolerance** — Specify the minimum prominence of a peak. The prominence of a peak measures how much the peak stands out due to its intrinsic height and its location relative to other peaks. For more information, in `findpeaks`, see the `MinPeakProminence` name-value pair argument.

DIFFERENT TYPES OF FEATURES

- SPECTRAL FEATURES

Wavelet transform

wavelets(filename)



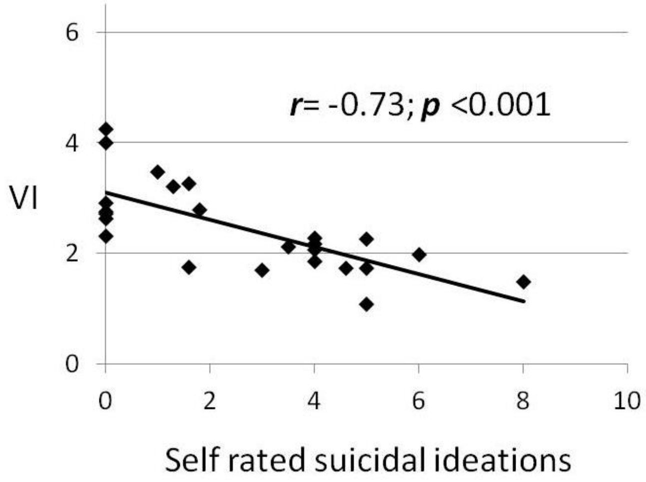
Indic P, Salvatore P, Maggini C, et al. (2011) Scaling Behavior of Human Locomotor Activity Amplitude: Association with Bipolar Disorder. PLOS ONE 2011 6(5): e20650.

DIFFERENT TYPES OF FEATURES

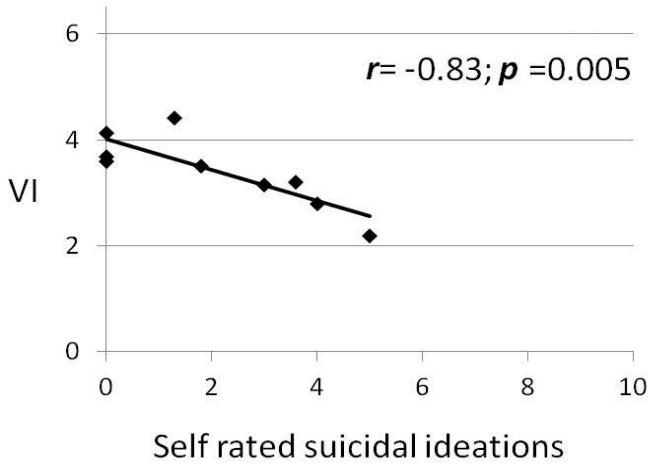
- SPECTRAL FEATURES

Individuals during major depression phase

N=24



N=1



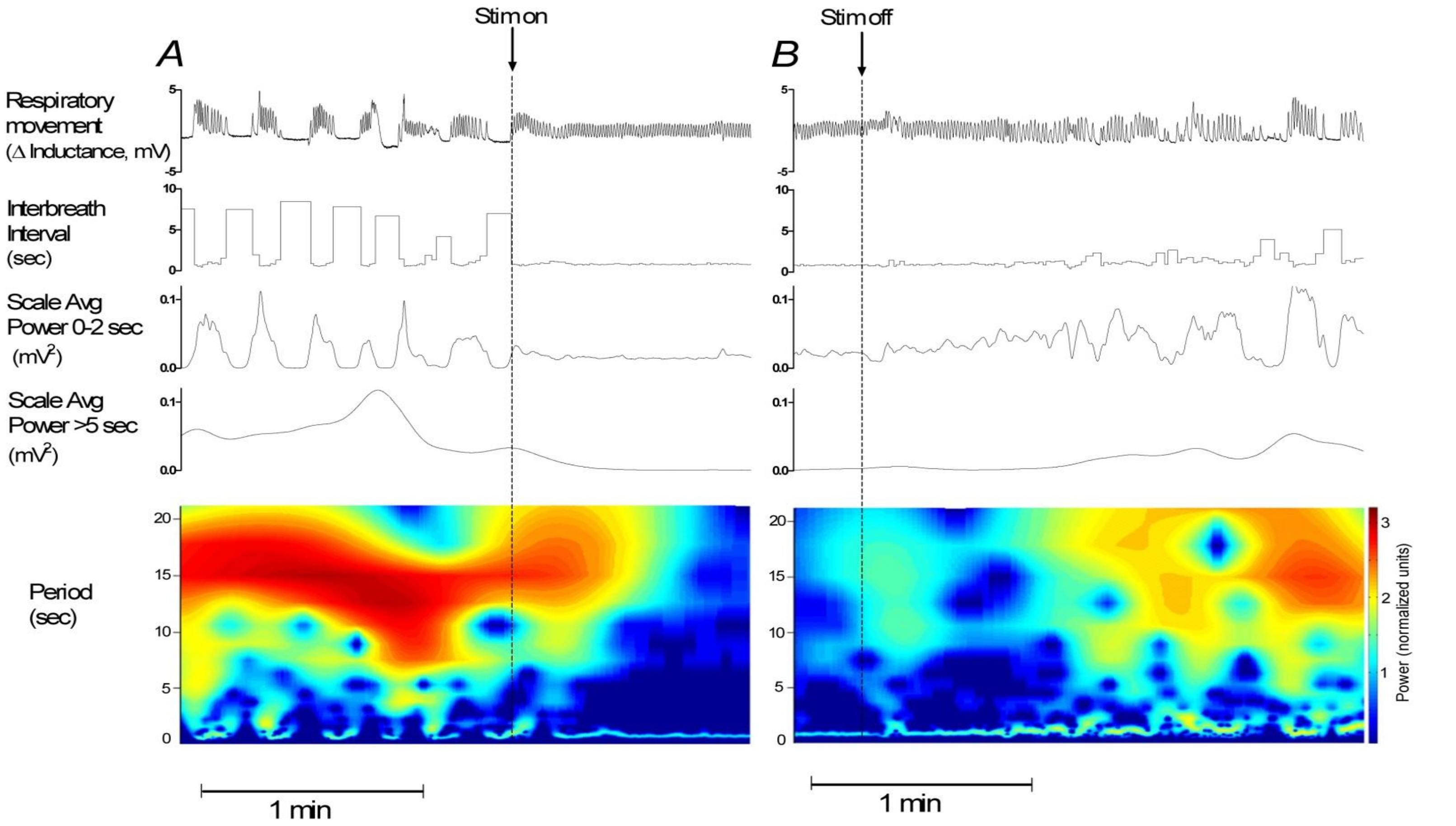
MATLAB functions:

Wavelets
corr

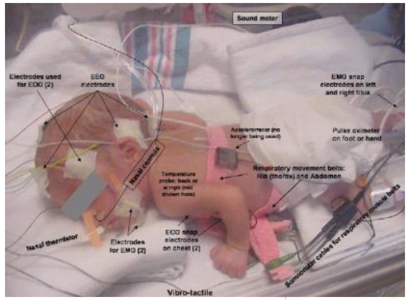
Indic P, Murray G, Maggini C, et al. (2012) Multiscale Motility Amplitude Associated with Suicidal Thoughts in Major Depression. PLOS ONE 2012, 7: e38761

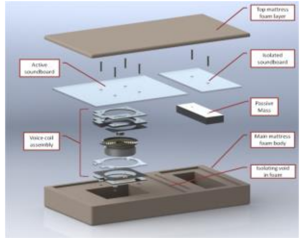
DIFFERENT TYPES OF FEATURES

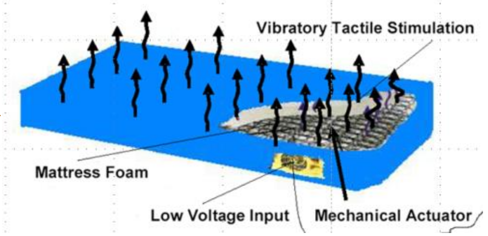
- SPECTRAL FEATURES



Bloch-Salisbury E, Indic P, Bednarek F, and Paydarfar D, J Appl Physiol., 2009, 107: 1017-1027



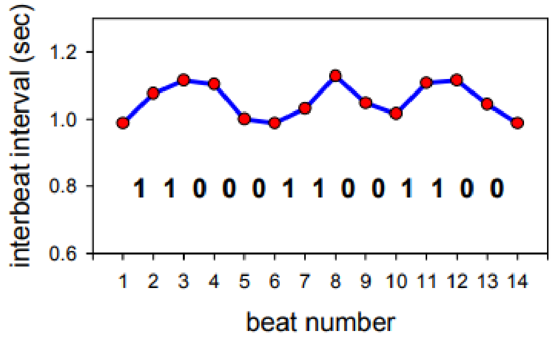




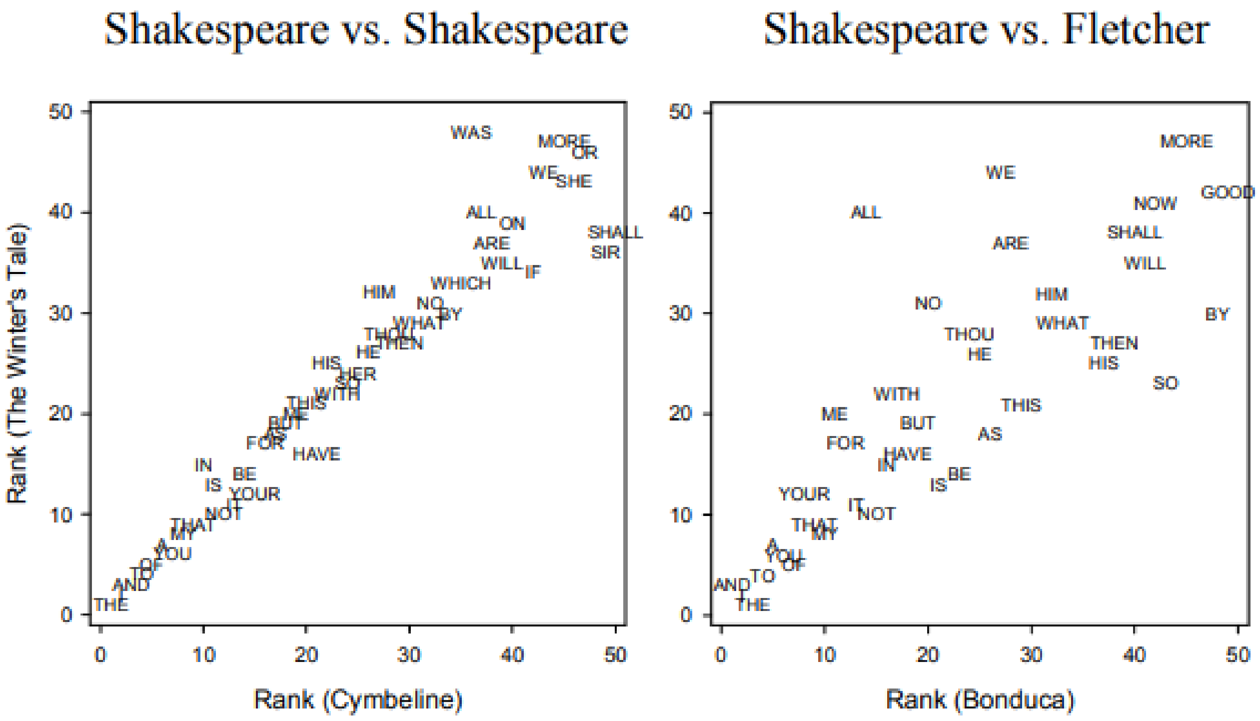
prapela

DIFFERENT TYPES OF FEATURES

➢NONLINEAR FEATURES



- Fluctuation Analysis
- Pattern Analysis
- Fractal Analysis
- Information Categorization Approach
- Power Law
- Entropy
- Dimension

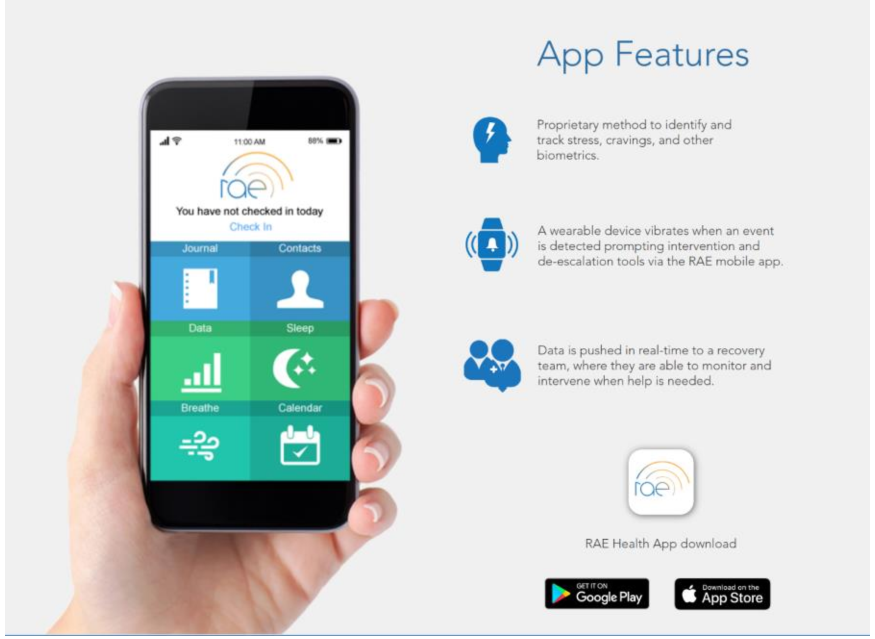


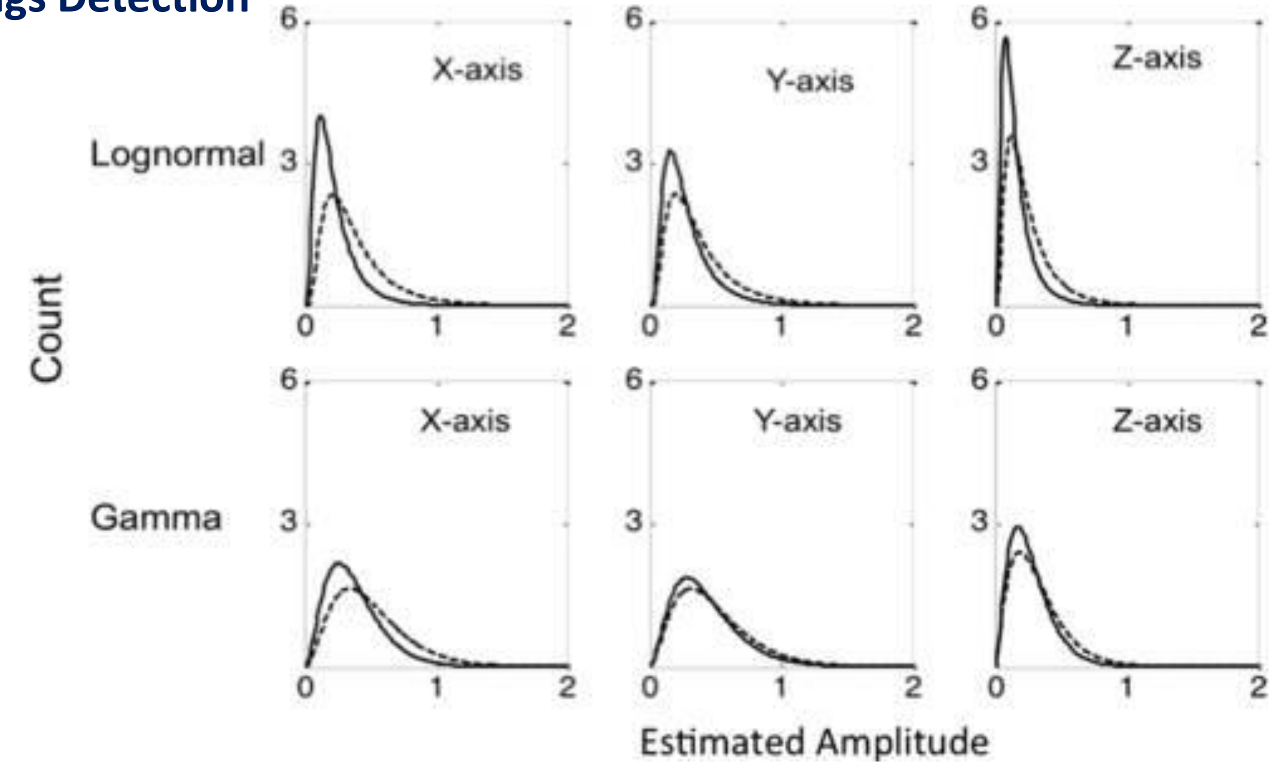
Yang CC, Peng CK, Yien HW, Goldberger AL. Information categorization approach to literary authorship disputes. Physica A. 2003; 329:473-483.

DIFFERENT TYPES OF FEATURES

➢ NONLINEAR FEATURES

Cravings Detection





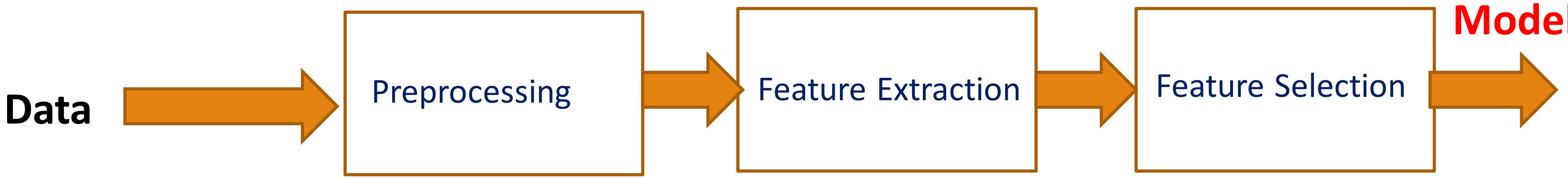
Carreiro, S, Chintha KK, Shrestha S*, Chapman B, Smelson D, Indic P. Wearable sensor based detection of stress and craving in patients during treatment for substance use disorder: A mixed methods pilot study. Drug and Alcohol Dependence. 2020, 107929

Statistical vs. Machine Learning Models

- ARTIFACT
- NOISE
- SIGNAL

- VARIABLE
- STATISTICAL FEATURES
- SPECTRAL FEATURES
- NONLINEAR FEATURES

Statistical or Machine Learning Models



Statistical vs. Machine Learning Models

Purpose:

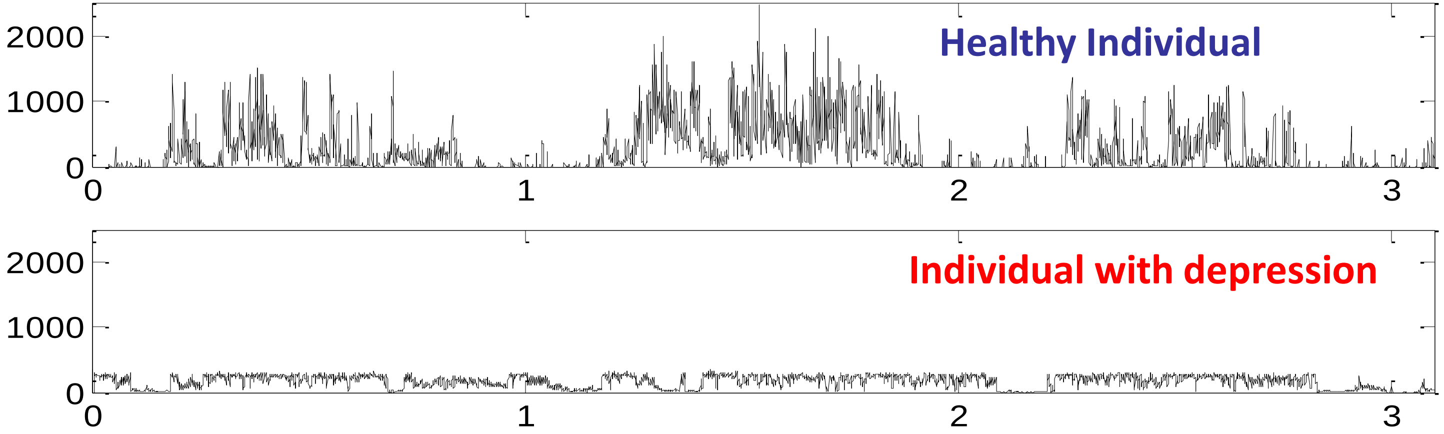
Statistical models are used for inference (To find association between features and an outcome). Results should be interpretable.

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

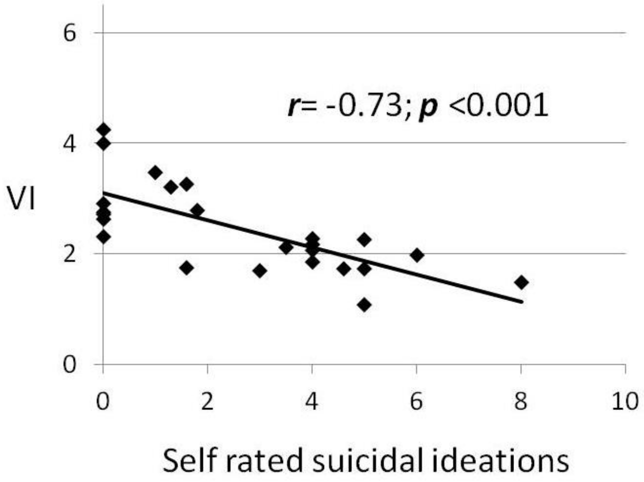
Statistical vs. Machine Learning Models

Association vs. Prediction

Philips Actiwatch 2



N=24



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

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

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

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

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

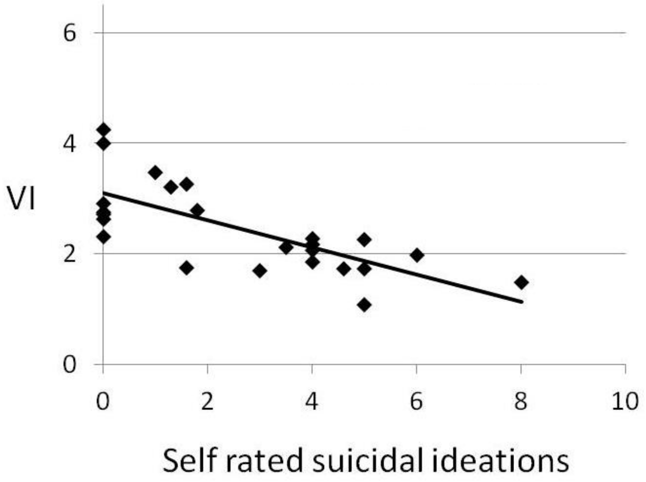
  Clustering

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

LEARNING APPROACHES

- Do machines actually "learn" ?

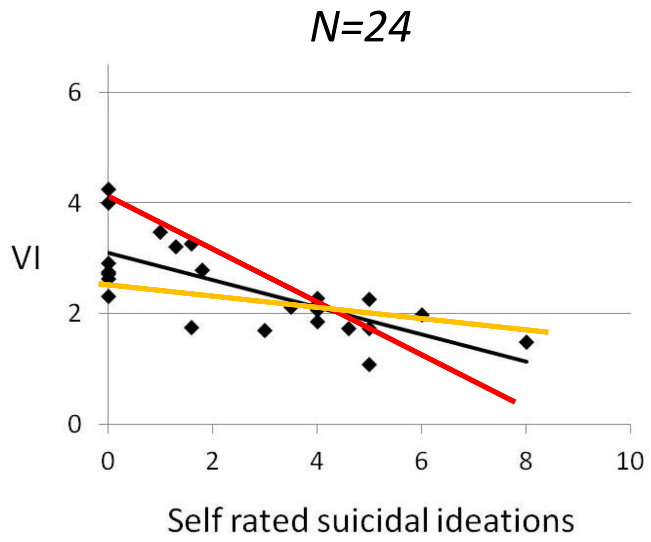
N=24



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

LEARNING APPROACHES

- Do machines actually "learn" ?



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

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

....

....

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

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

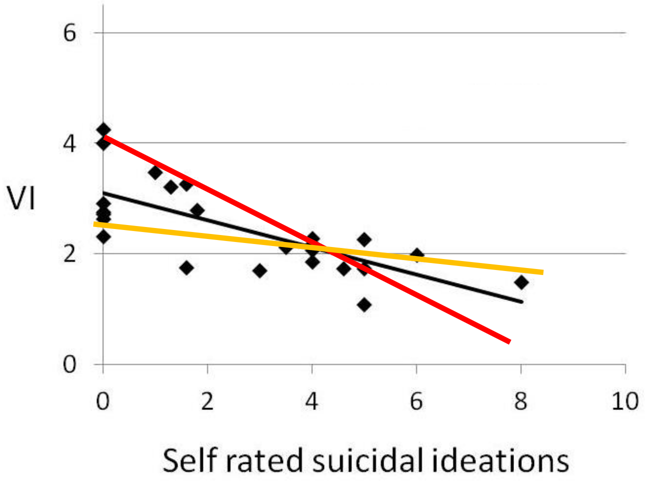
$$\widetilde{VI} = m \times SI + C$$

LEARNING APPROACHES

- Do machines actually “learn” ?

How do we find minimum E ?

N=24



m →

C ↓

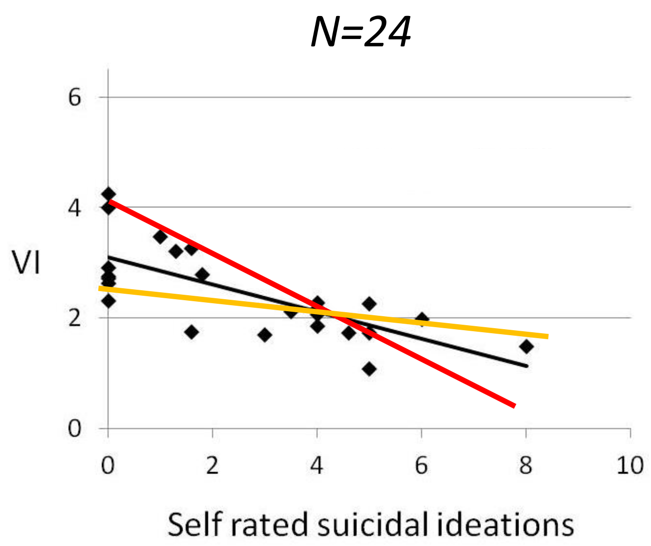
0.1	0.6	0.8	0.01	0.5
1	10	0.01	0.001	0.002
8	7	0.0006	0.03	0.55
100	12	0.1	12	0.89
2	1	2	0.5	0.05

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

LEARNING APPROACHES

- Do machines actually "learn" ?

How do we find minimum E ?



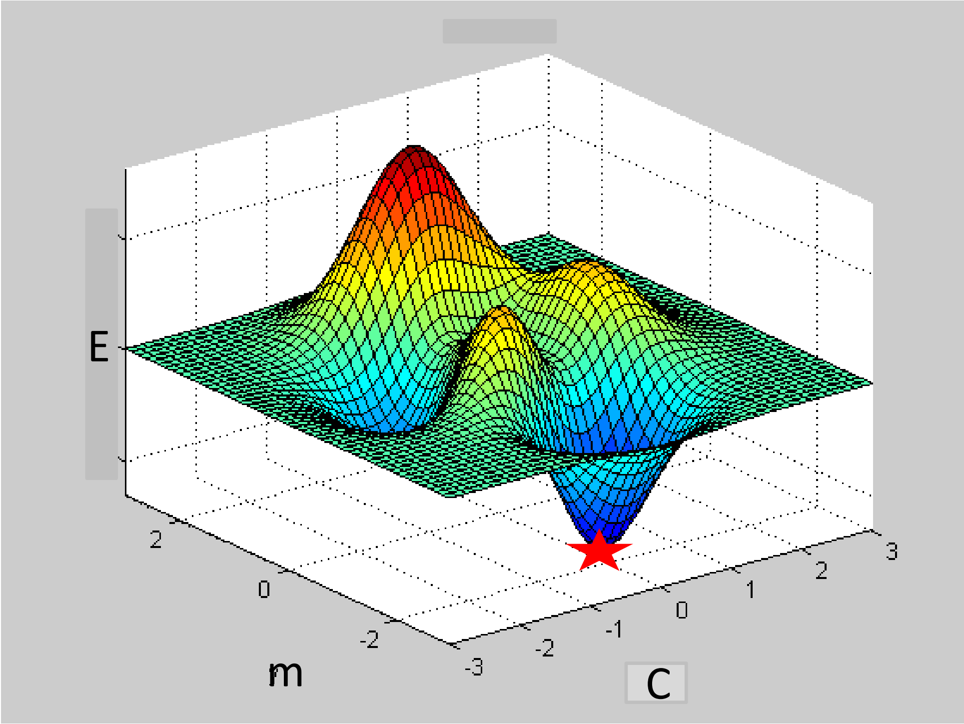
- Gradient Descent

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

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

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

Mean

Variance

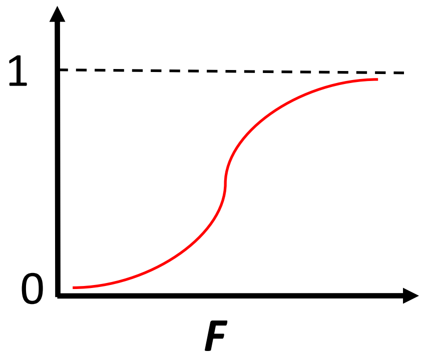
Skewness

Kurtosis

Power

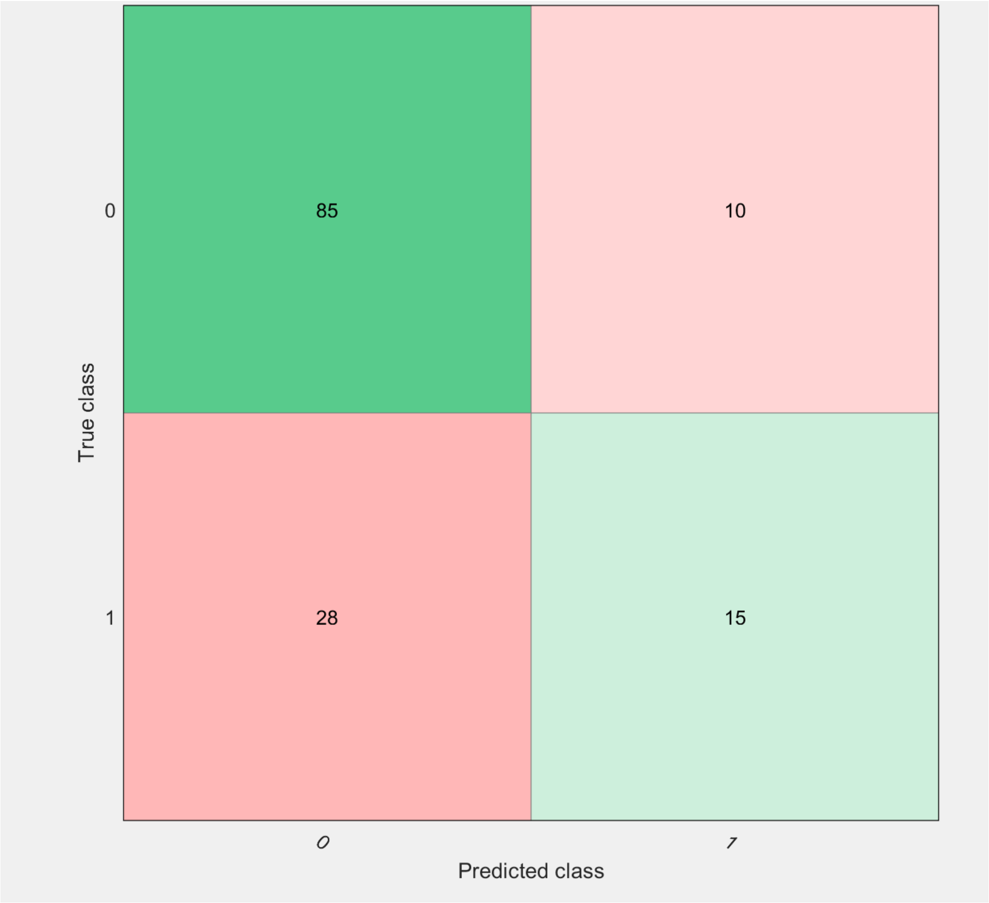
Period

Linear Regression



Logistic Regression

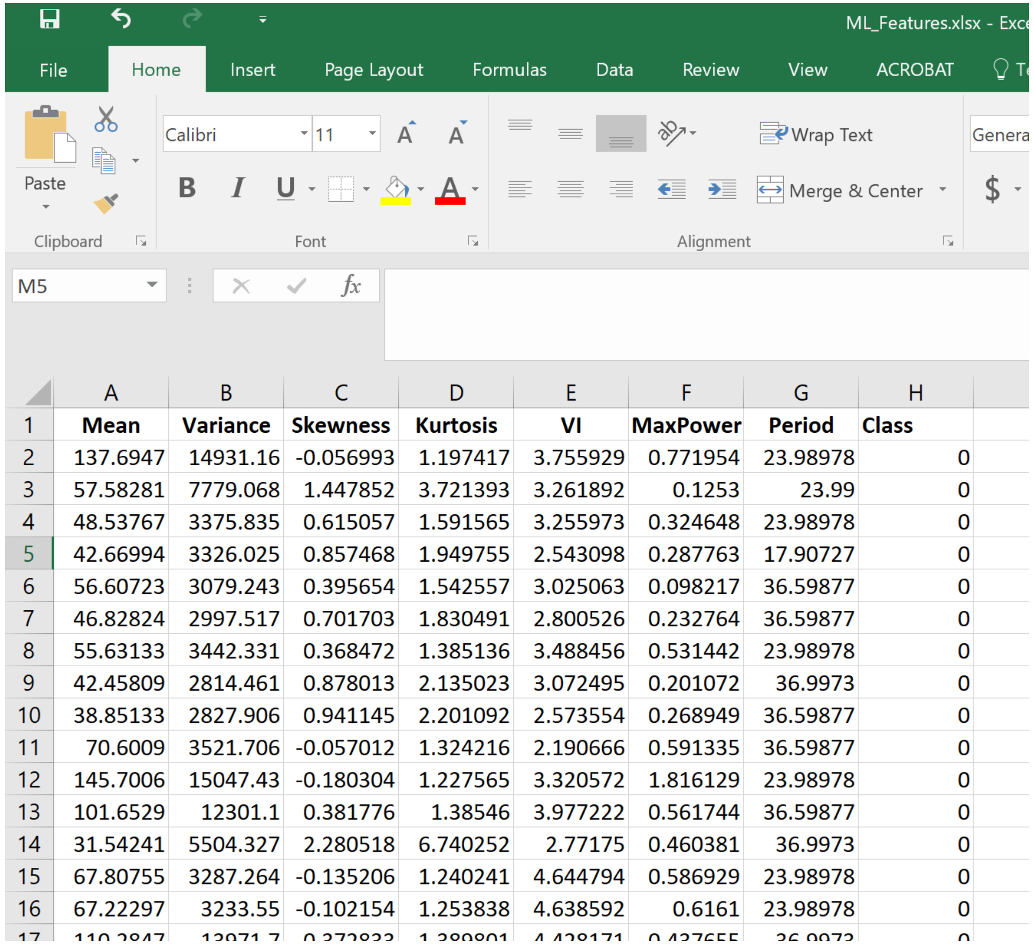
Accuracy ~73%



LEARNING APPROACHES

- ➢How to implement in MATLAB ?

<u>Step 1:</u> Create an excel sheet with features with class assignments

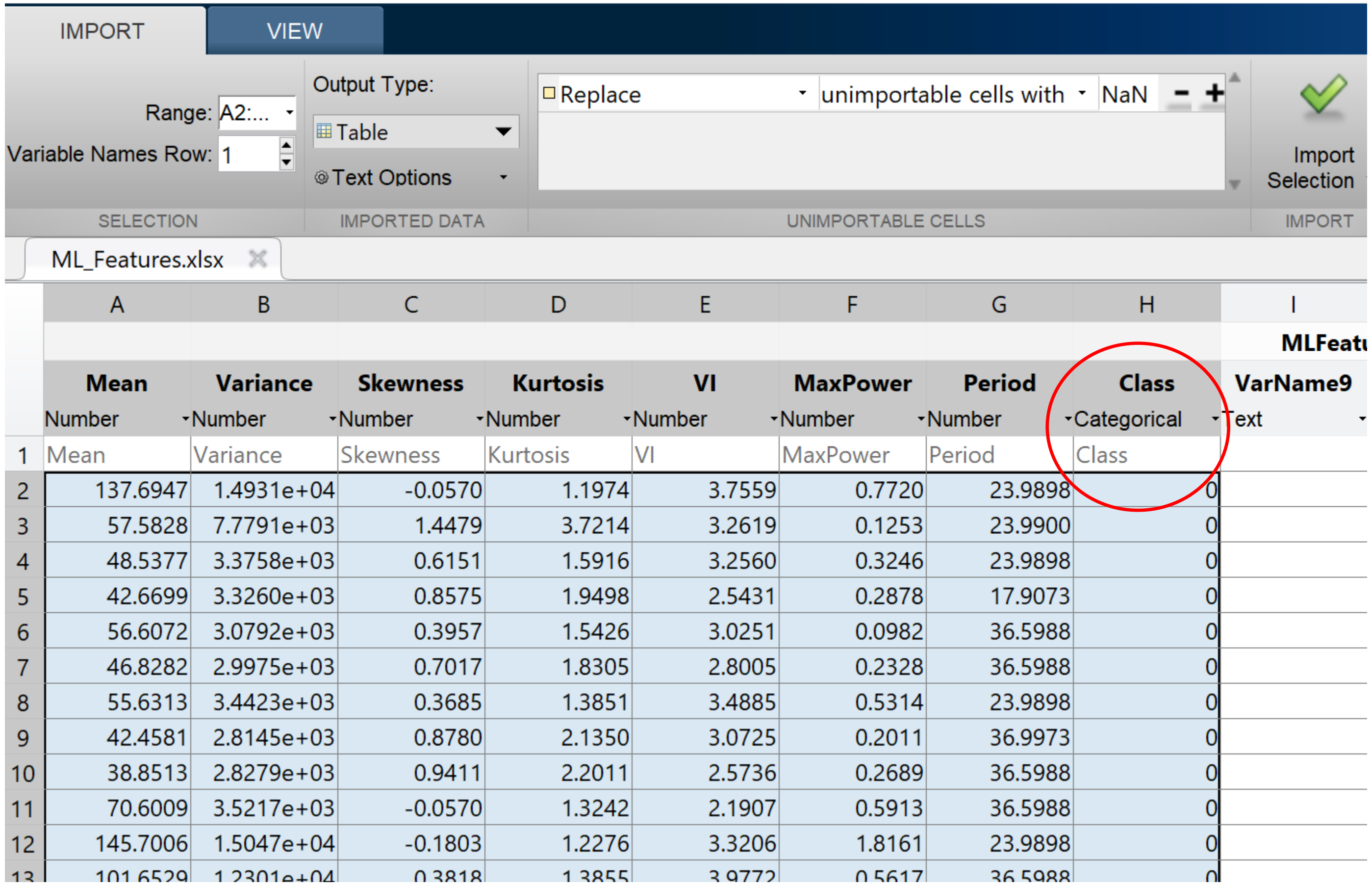


	A	B	C	D	E	F	G	H
1	Mean	Variance	Skewness	Kurtosis	VI	MaxPower	Period	Class
2	137.6947	14931.16	-0.056993	1.197417	3.755929	0.771954	23.98978	0
3	57.58281	7779.068	1.447852	3.721393	3.261892	0.1253	23.99	0
4	48.53767	3375.835	0.615057	1.591565	3.255973	0.324648	23.98978	0
5	42.66994	3326.025	0.857468	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
17	110.2847	13971.7	0.372833	1.389801	4.428171	0.437655	36.9973	0

LEARNING APPROACHES

- How to implement in MATLAB ?

<u>Step 2:</u> Open MATLAB and drag the excel file to workspace

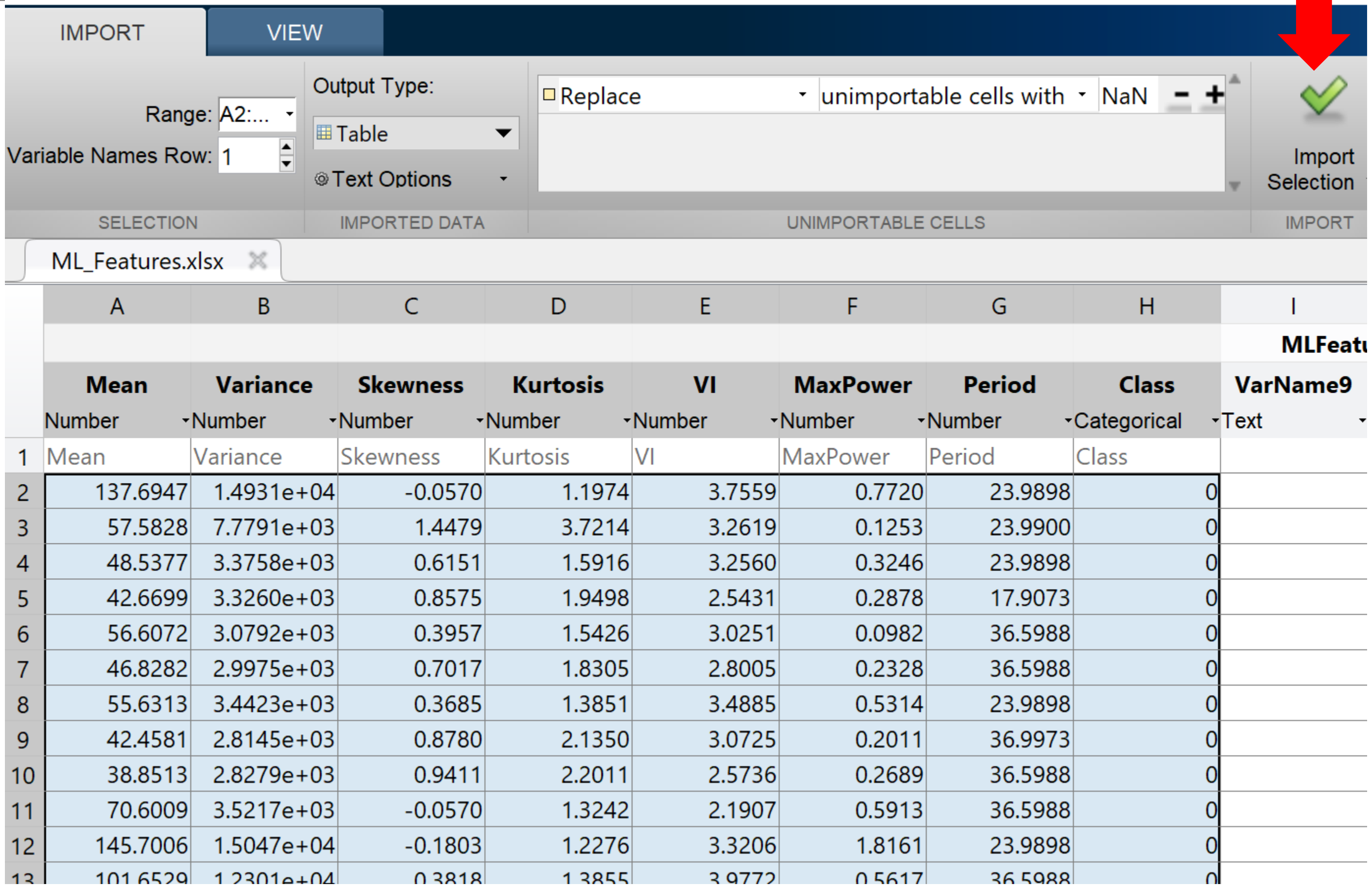


	A	B	C	D	E	F	G	H	I
	Mean	Variance	Skewness	Kurtosis	VI	MaxPower	Period	Class	VarName9
	Number	Number	Number	Number	Number	Number	Number	Categorical	Text
1	Mean	Variance	Skewness	Kurtosis	VI	MaxPower	Period	Class	
2	137.6947	1.4931e+04	-0.0570	1.1974	3.7559	0.7720	23.9898	0	
3	57.5828	7.7791e+03	1.4479	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

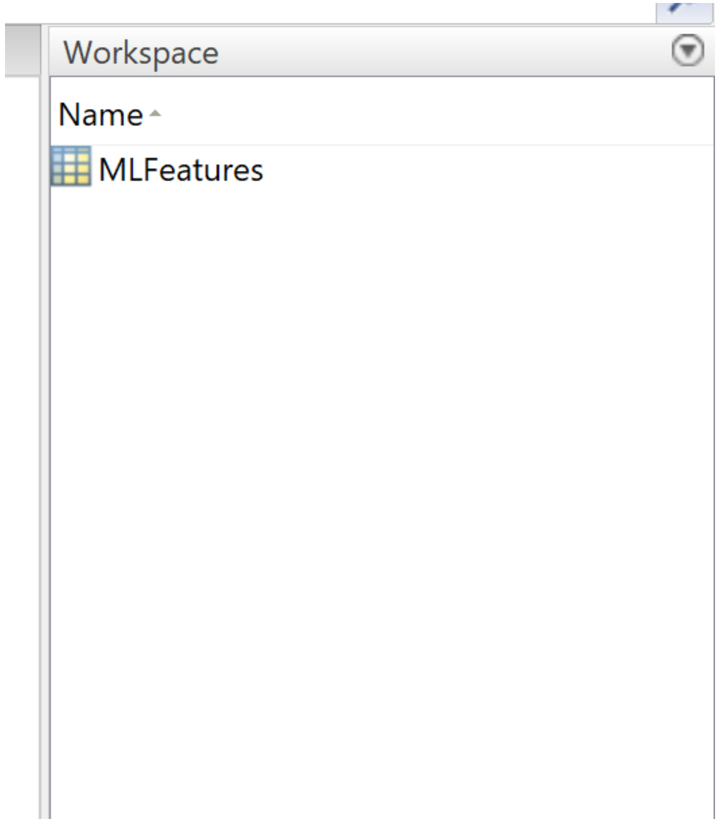


	A	B	C	D	E	F	G	H	I
	Mean	Variance	Skewness	Kurtosis	VI	MaxPower	Period	Class	VarName9
	Number	Number	Number	Number	Number	Number	Number	Categorical	Text
1	Mean	Variance	Skewness	Kurtosis	VI	MaxPower	Period	Class	
2	137.6947	1.4931e+04	-0.0570	1.1974	3.7559	0.7720	23.9898	0	
3	57.5828	7.7791e+03	1.4479	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,

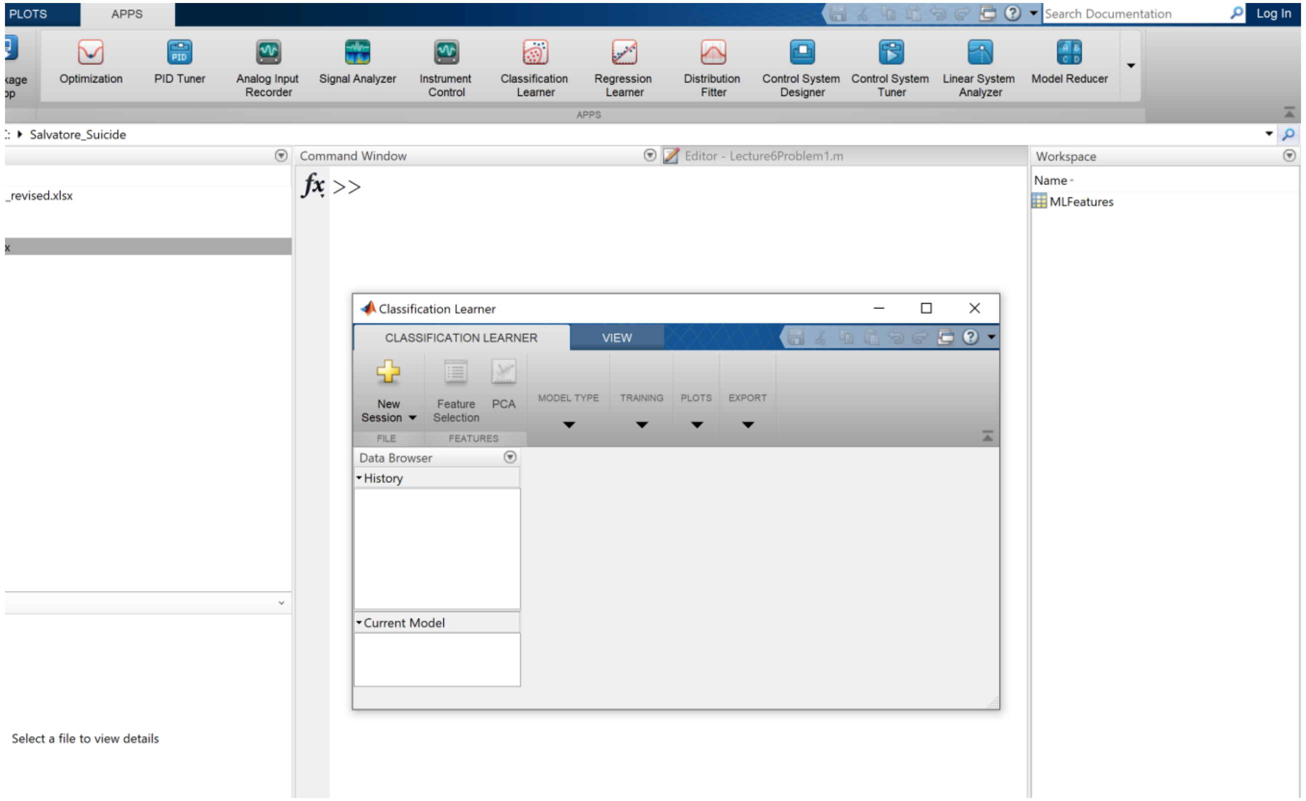
-click classification learner,

- -select Logistic Regression

  from Model Type

-click New Session,

-select from Workspace

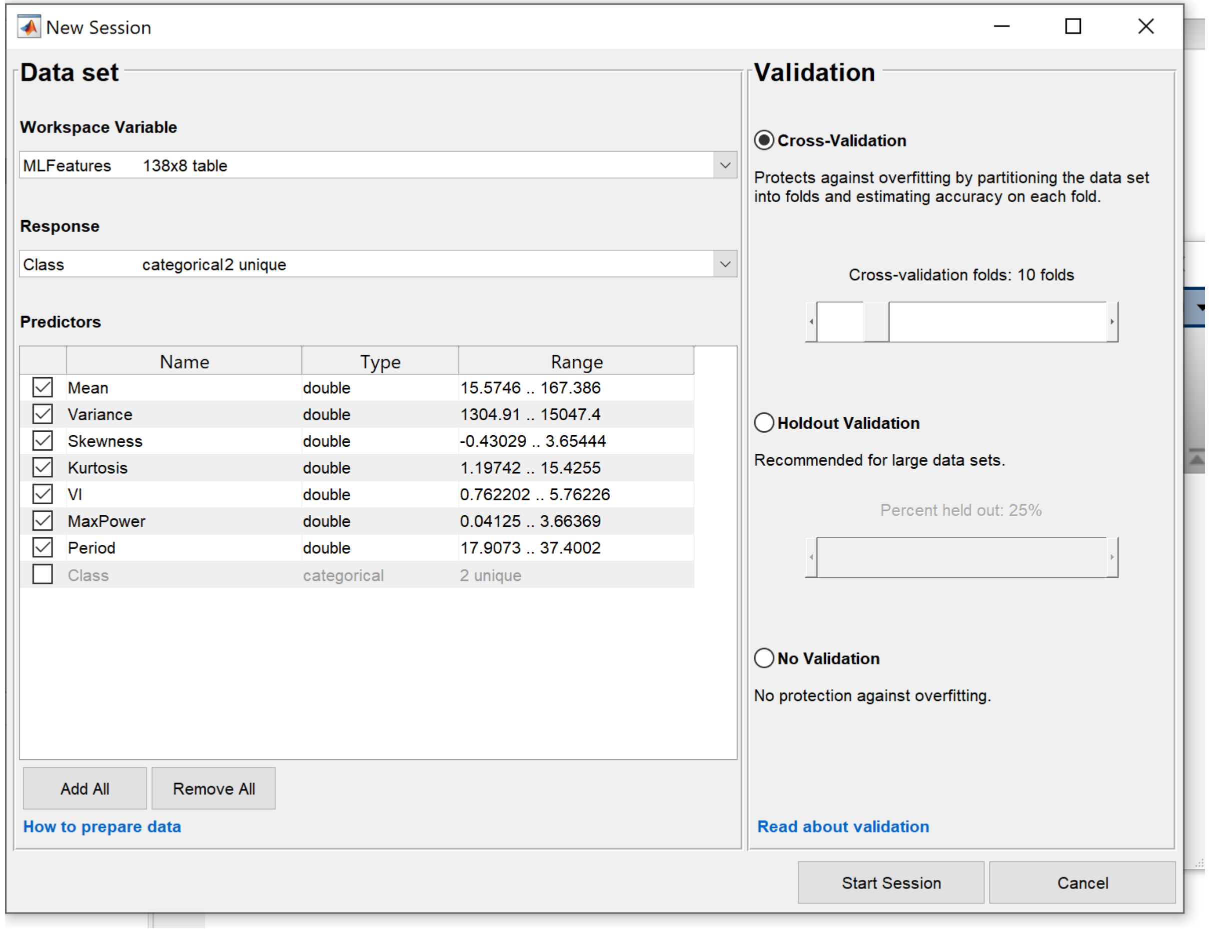


LEARNING APPROACHES

➢ How to implement in MATLAB ?

Step 6: Set 10 fold Cross validation

- Start the session



APPROACHES IN DETAIL

- SUPERVISED LEARNING

- UNSUPERVISED LEARNING

APPROACHES IN DETAIL

- SUPERVISED LEARNING
- UNSUPERVISED LEARNING

APPROACHES IN DETAIL

- SUPERVISED LEARNING (Classification / Prediction)

Provide training set with features and solutions

APPROACHES IN DETAIL

- SUPERVISED LEARNING (Classification / Prediction)

  Find the area of a rectangle

L	W	A	A1 (L+W)	A2 (L-W)	A3 (L*W)	A4 L/W
12.1	13.4	162.3	25.5	-1.3	162.14	0.90
8.6	9.7	83.4	18.3	-1.1	83.42	0.89
3.2	5.4	17.3	8.6	-2.2	17.28	0.59
6.1	10.2	62.25	16.3	-4.1	62.22	0.60
18.2	6.4	116.5	24.6	11.8	116.48	2.83
1.6	2.8	4.5	4.4	-1.2	4.48	0.57
7.7	0.6	4.7	8.3	7.1	4.62	12.83

APPROACHES IN DETAIL

- SUPERVISED LEARNING (Classification / Prediction)

  Find the area of a rectangle

L	W	A	E1 |A-A1|	E2 |A-A2|	E3 |A-A3|	E4 |A-A4|
12.1	13.4	162.3	136.8	163.6	0.16	161.40
8.6	9.7	83.4	65.1	84.5	0.02	82.51
3.2	5.4	17.3	8.7	19.5	0.02	16.71
6.1	10.2	62.25	45.95	66.35	0.03	61.65
18.2	6.4	116.5	91.90	104.70	0.02	113.66
1.6	2.8	4.5	0.1	5.7	0.02	3.93
7.7	0.6	4.7	3.6	2.4	0.08	8.13

APPROACHES IN DETAIL

- SUPERVISED LEARNING (Classification / Prediction)
  - Linear Regression
  - Logistic Regression
  - Support Vector Machines
  - k-Nearest Neighbors
  - Decision Trees and Random Forests
  - Neural Networks

APPROACHES IN DETAIL

- SUPERVISED LEARNING (Classification / Prediction)
  - Linear Regression

Given m outcomes y^i where $i = 1,2, \ldots .., m$ with each outcome depends on n features x_j where $j = 1,2, \ldots .., n$. Find the best estimate of y^i as $\hat{y}^i$ using the n features with appropriate parameters θ_j such that $J = \left\langle \left(\hat{y}^i - y^i\right)^2 \right\rangle$

$$\hat{y}^i = \theta_0^i + \theta_1^i x_1^i + \theta_2^i x_2^i + \cdots \ldots \ldots . + \theta_n^i x_n^i$$

APPROACHES IN DETAIL

- SUPERVISED LEARNING (Classification / Prediction)
    - Linear Regression

$$\hat{y}^i = \theta_0^i + \theta_1^i x_1^i + \theta_2^i x_2^i + \cdots \ldots \ldots . + \theta_n^i x_n^i$$

$$\hat{Y} = \Theta . X = h_\theta(X)$$

Cost Function to Minimize

$$J = \left\langle \left(\hat{y}^i - y^i\right)^2 \right\rangle = (\hat{Y} - Y)^T(\hat{Y} - Y)$$

APPROACHES IN DETAIL

- SUPERVISED LEARNING (Classification / Prediction)
    - Linear Regression

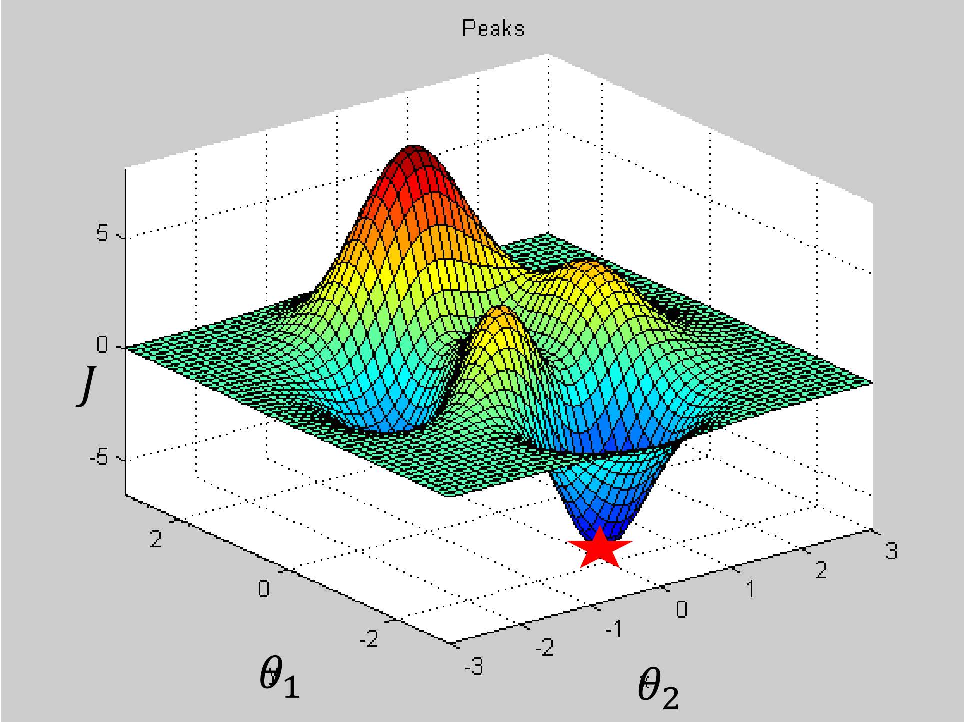
$$\hat{y}^i = \theta_0^i + \theta_1^i x_1^i + \theta_2^i x_2^i + \cdots \ldots \ldots . + \theta_n^i x_n^i$$

$$\hat{Y} = \Theta . X = h_\theta(X)$$

- Gradient Descent by **Louis Augustin Cauchy** in 1847

Cost Function to Minimize

$$J = \left\langle \left(\hat{y}^i - y^i \right)^2 \right\rangle = (\hat{Y} - Y)^T (\hat{Y} - Y)$$



APPROACHES IN DETAIL

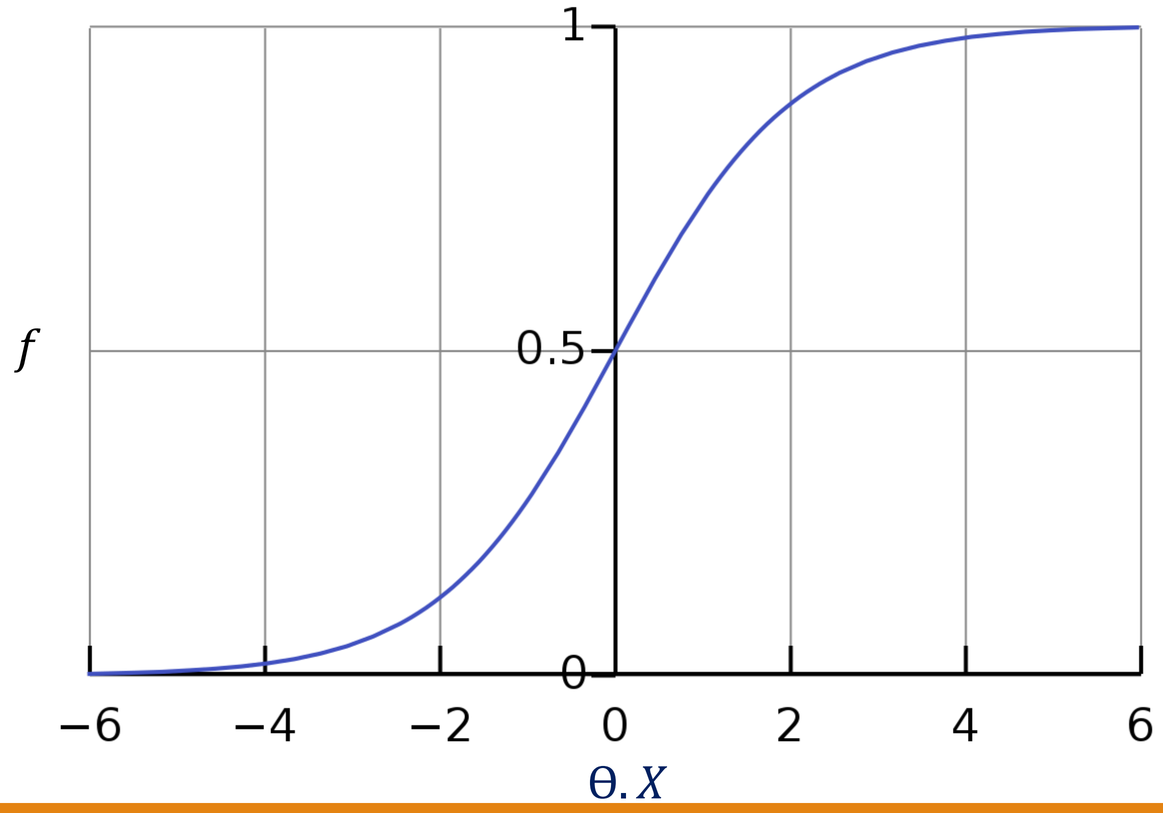
- SUPERVISED LEARNING (Classification / Prediction)
  - Logistic Regression

$$\hat{p} = f(\Theta.X) = h_\theta(X)$$

$$\hat{y} = 1 \ if \, \hat{p} < 0.5; \ \ \hat{y} = 0 \ if \, \hat{p} \geq 0.5$$

Derive Cost Function to Minimize

J



APPROACHES IN DETAIL

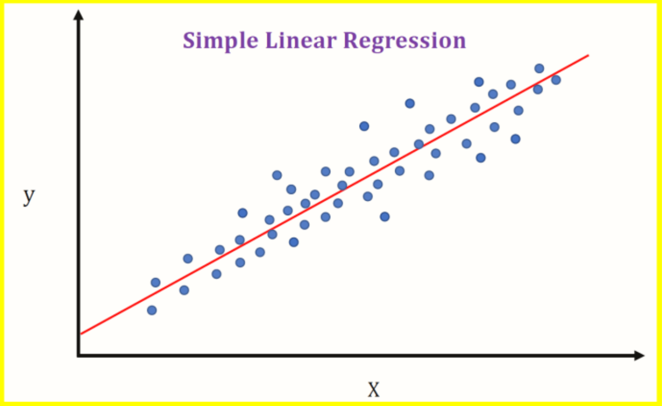
- SUPERVISED LEARNING (Classification / Prediction)
  - Linear Regression

    Mainly for regression (predicting an outcome)

  - Logistic Regression

    Mainly for classification (0 or 1)

    High Risk vs. Low Risk



https://medium.datadriveninvestor.com/machine-learning-101-part-1-24835333d38a

APPROACHES IN DETAIL

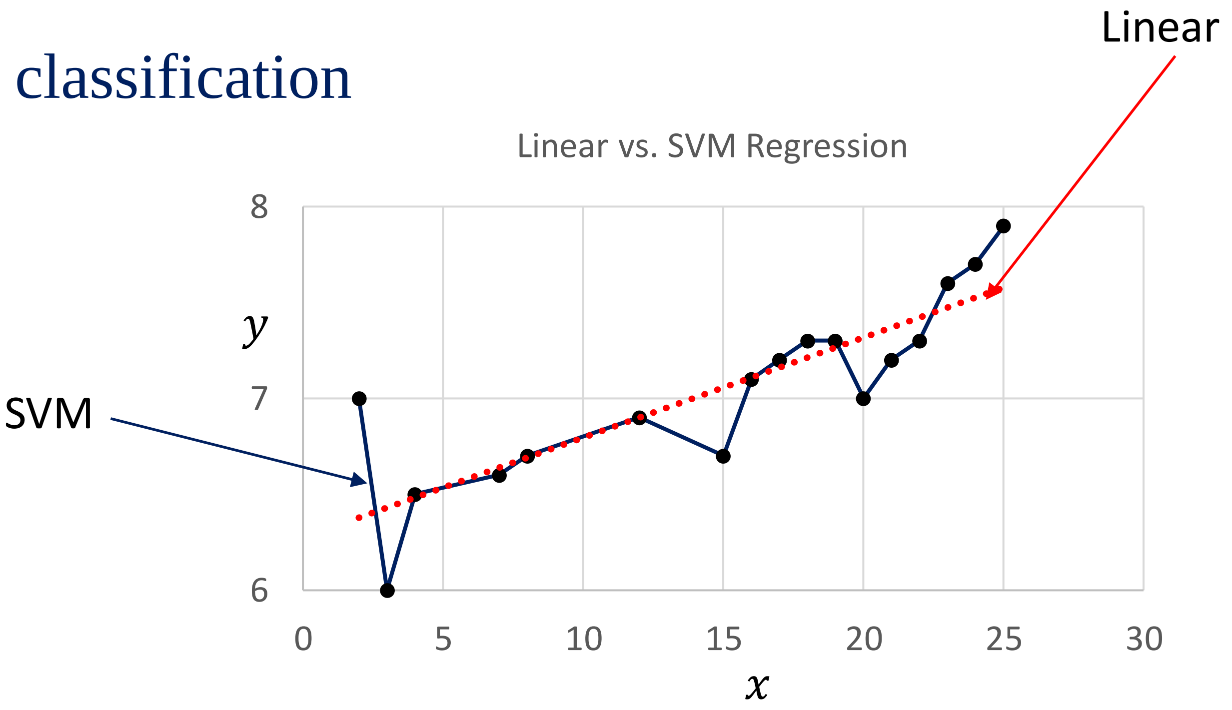
- SUPERVISED LEARNING (Classification / Prediction)
  - Support Vector Machine

Used for regression as well as classification

APPROACHES IN DETAIL

- SUPERVISED LEARNING (Classification / Prediction)
  - Support Vector Machine (SVM)

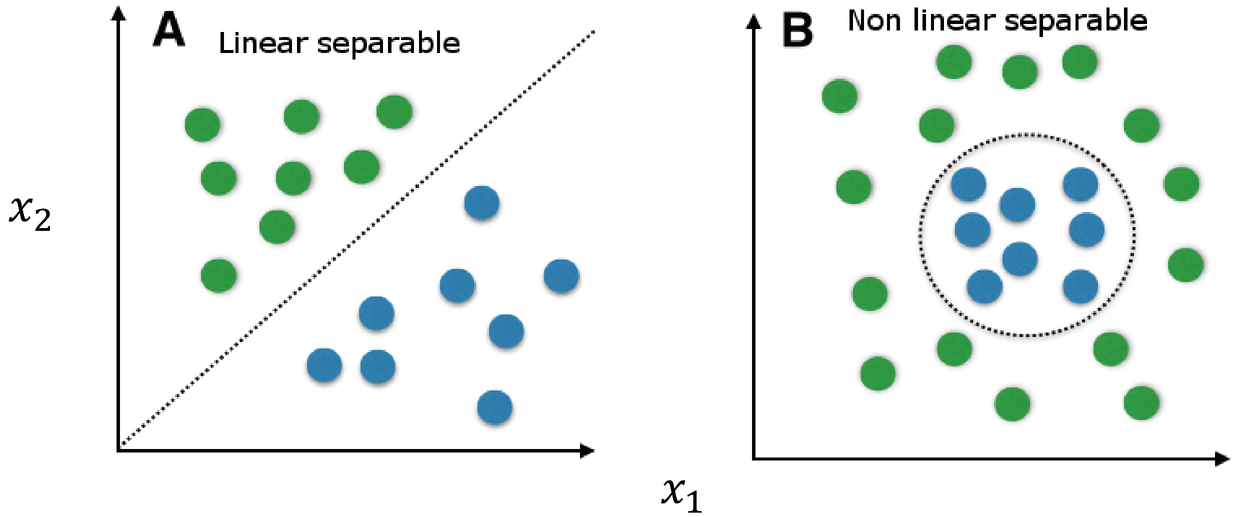
Used for regression as well as classification



APPROACHES

- SUPERVISED LEARNING (Classification / Prediction)
    - Support Vector Machine (SVM)

Used for regression as well as classification



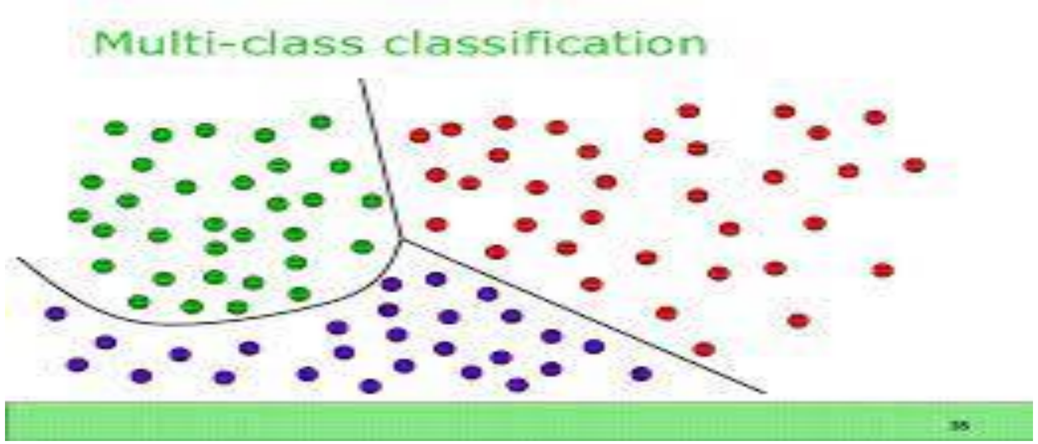
https://medium.com/@LSchultebraucks/introduction-to support-vector-machines-9f8161ae2fcb

APPROACHES IN DETAIL

- SUPERVISED LEARNING (Classification / Prediction)
  - Support Vector Machine (SVM)

Used for regression as well as classification

x_2



x_1

https://www.mathworks.com/matlabcentral/fileexchange/62061-multi-class-svm

APPROACHES IN DETAIL

- SUPERVISED LEARNING (Classification / Prediction)
  - Linear Regression
  - Logistic Regression
  - Support Vector Machines
  - k-Nearest Neighbors
  - Decision Trees and Random Forests
  - Neural Networks

Project 1

- Prediction of House Price (housing.csv) Regression Problem

longitude
latitude
housing_median_age
total_rooms
total_bedrooms
population
households median_income
median_house_value
ocean_proximity

Project 2

- Prediction of House Price (housing.csv) Classification Problem

longitude
latitude
housing_median_age
total_rooms
total_bedrooms
population
households median_income
median_house_value (High/Low) Threshold= 257500
ocean_proximity

Project 2

- Prediction of House Price (housing.csv) Classification Problem

Confusion Matrix

True Class		Predicted 1	Predicted 0	
	1	True Positive	False Negative	→ Total Positive
	0	False Positive	True Negative	→ Total Negative

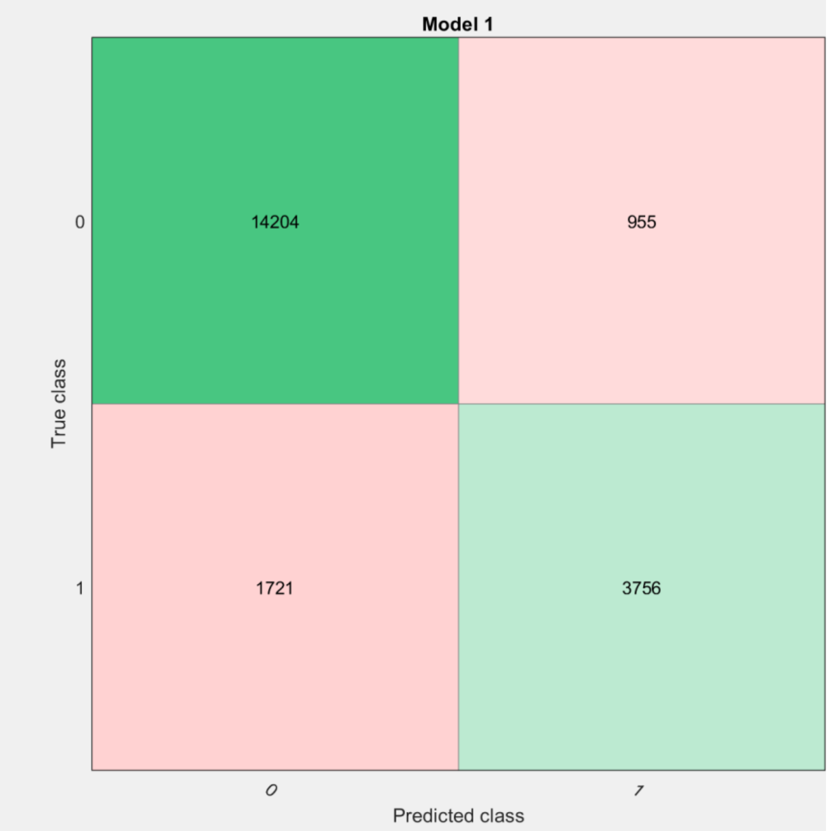
Predicted Class

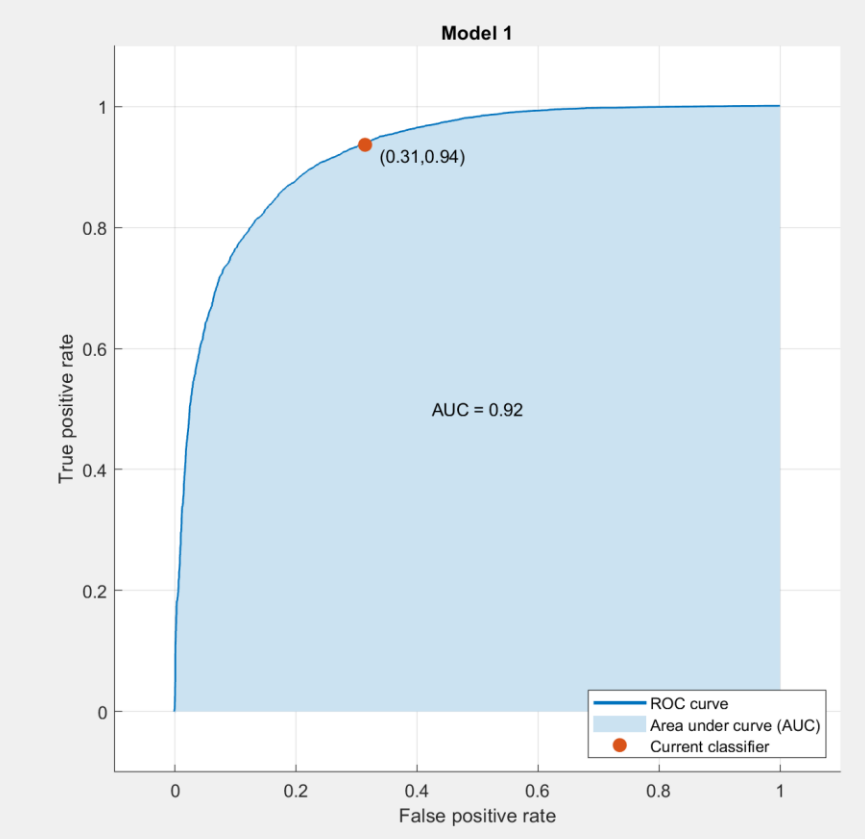
True Positive Rate = True Positive / Total Positive

True Negative Rate = True Negative / Total Negative = 1 – False Positive Rate

Project 2

- Prediction of House Price (housing.csv) Classification Problem





Project 3

- To test the hypothesis that the features of SpO_2 can detect smoker from non-smoker

Project 4

- To test the hypothesis that the features of saliva can detect COPD from other conditions

- Data set : http://archive.ics.uci.edu/ml/datasets/Exasens

https://cml.ics.uci.edu/