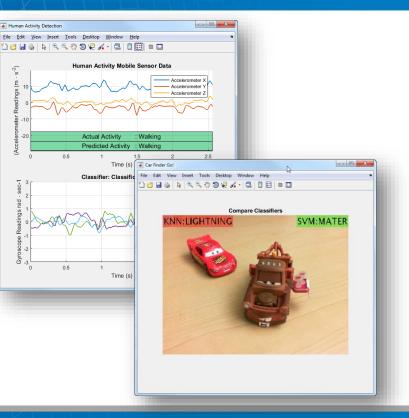


# Machine Learning Made Easy

#### Shashank Prasanna

Product Marketing Manager, Statistics and Machine Learning

shashank.prasanna@mathworks.com



© 2015 The MathWorks, Inc.



- Machine Learning
  - What is Machine Learning and why do we need it?
  - Common challenges in Machine Learning
- Example 1: Human activity learning using mobile phone data
  - Learning from sensor data
- Example 2: Real-time car identification using images
  - Learning from images
- Summary & Key Takeaways



# **Machine Learning is Everywhere**

- Image Recognition
- Speech Recognition
- Stock Prediction
- Medical Diagnosis
- Data Analytics
- Robotics
- and more...

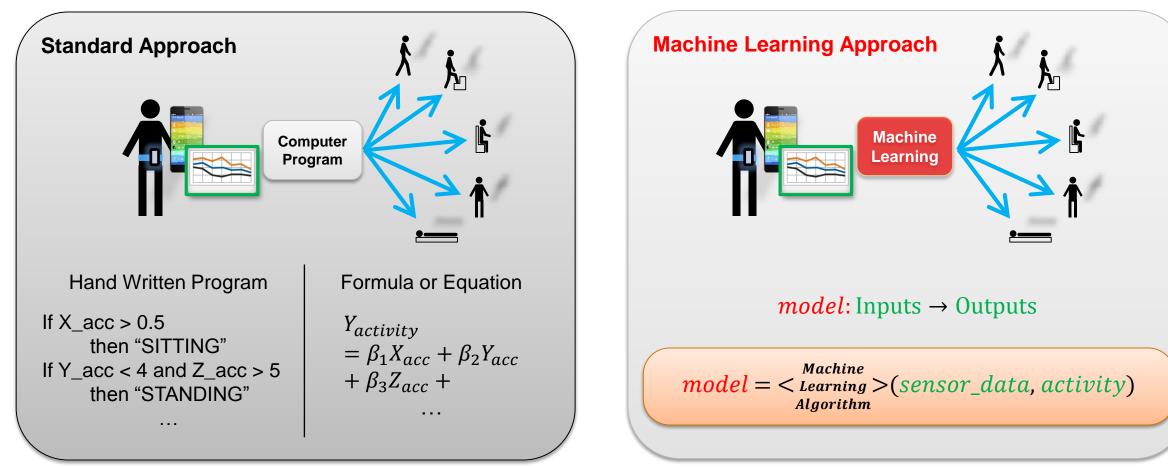




## **Machine Learning**

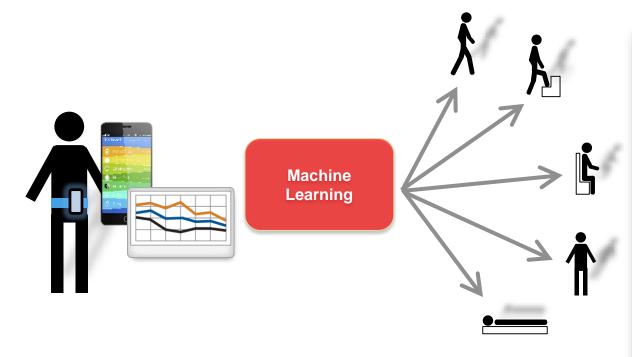
Machine learning uses data and produces a program to perform a task

#### Task: Human Activity Detection





# **Example: Human Activity Learning Using Mobile Phone Data**



### Data:

3-axial Accelerometer data
 3-axial Gyroscope data

CLASSIFICATION LEARNER		VIEW	VIEW				00 🔨	
Import Data	Feature Selection	A Boosted Trees	A Bagged Trees		Image: Subspace Subspace KNN         Image: Train Subspace Su	Advanced	Scatter Plot	Export Model
FILE	FEATURES		CLA	ASSI	IFIER	TRAINING	PLOTS	EXPORT
Data Brov	wser			)	Scatter Plot 🛛 🗶			
▼ Histor	у				Variable on X axis:		Scatter Plot	- 6 h
SVM Linear SV	М		86.4%		avg_body_gyro_x_test	0.35	Scatter Plot	ornumanA
SVM BoxConstraint = 3		87.1%		Variable on Y axis:	0.3		•	
KNN Fine KNN			94.9%		stdv_total_acc_y_test		••	
KNN		90.7%		Correctly classified Walking	test 0.25			
		94.1%	ClimbingStairs     Sitting	<ul> <li>ClimbingStairs</li> </ul>	> 0.2	×	×	
KNN NumNeighbors = 2 91		91.7%			stdv_total_acc	•	×x	
Ensemble NumLearners = 100 9.		95.9%		Misclassified - true class is: Walking	stdv 1	×		
▼ Current model			ClimbingStairs Sitting	0.1	•	••		
	Custom >					0.05		
Data Trar Status: Tr	sformation: N	one			Show Classifier Results			



# "essentially, all models are wrong, but some are useful" – George Box



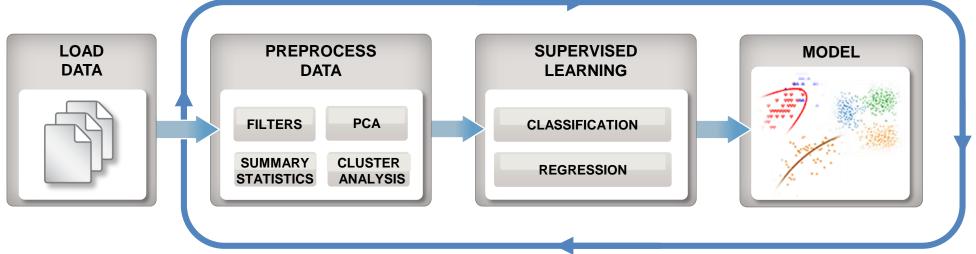
# **Challenges in Machine Learning** Hard to get started

Steps	Challenge
Access, explore and analyze data	<b>Data diversity</b> Numeric, Images, Signals, Text – not always tabular
Preprocess data	Lack of domain tools Filtering and feature extraction Feature selection and transformation
Train models	<b>Time consuming</b> Train several models to find the "best"
Assess model performance	Avoid pitfalls Over Fitting Speed-Accuracy-Complexity tradeoffs
Iterate	

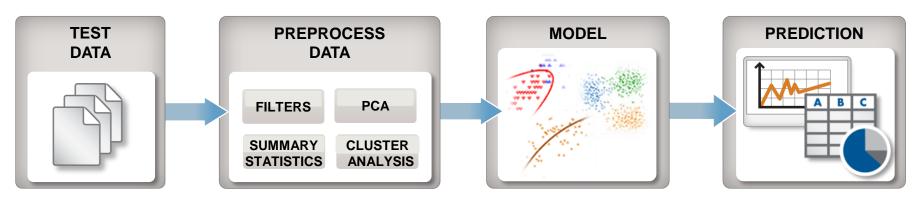


# **Machine Learning Workflow for Example 1**

#### Train: Iterate till you find the best model

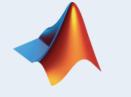


#### **Predict:** Integrate trained models into applications





- Machine Learning
  - What is Machine Learning and why do we need it?
  - Common challenges in Machine Learning
- Example 1: Human activity learning using mobile phone data
  - Learning from sensor data



- Example 2: Real-time car identification using images
  - Learning from images
- Summary & Key Takeaways



# **Example 1: Human Activity Learning Using Mobile Phone Data**

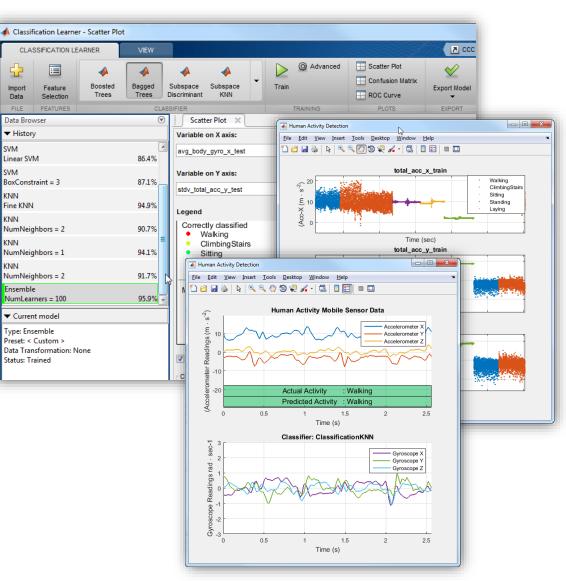
**Objective:** Train a classifier to classify human activity from sensor data

#### Data:

Predictors	3-axial Accelerometer and Gyroscope data	
Response	Activity: 🕺 🦌 🔓 🛉 🛻	

#### Approach:

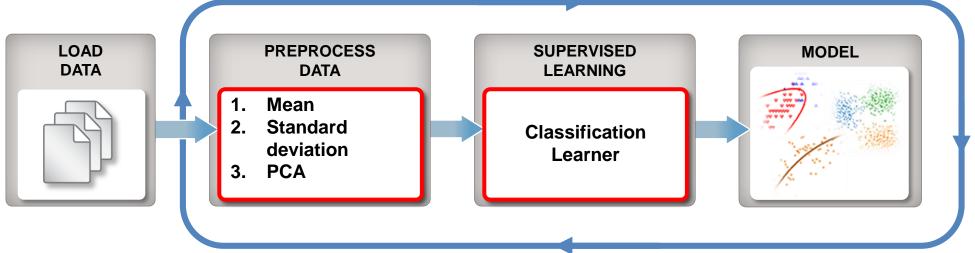
- Extract features from raw sensor signals
- Train and compare classifiers
- Test results on new sensor data



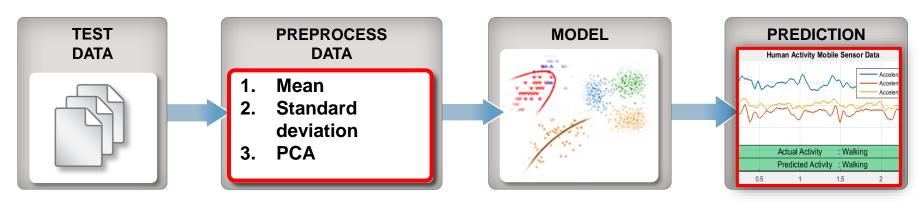


# **Machine Learning Workflow for Example 1**

#### Train: Iterate till you find the best model

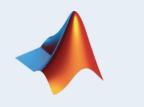


#### **Predict:** Integrate trained models into applications





- Machine Learning
  - What is Machine Learning and why do we need it?
  - Common challenges in Machine Learning
- Example 1: Human activity learning using mobile phone data
  - Learning from sensor data
- Example 2: Real-time car identification using images
   Learning from images



Summary & Key Takeaways



# **Example 2: Real-time Car Identification Using Images**

**Objective:** Train a classifier to identify car type from a webcam video

#### Data:

Predictors	Several images of cars:					
Response	NIGEL, LIGHTNING, SANDDUNE, MATER					

#### Approach:

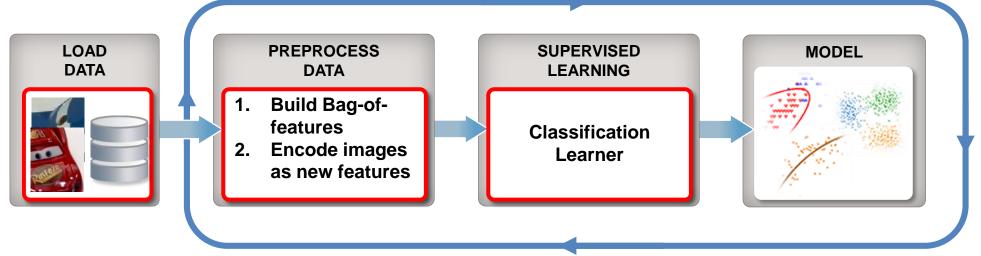
- Extract features using Bag-of-words
- Train and compare classifiers
- Classify streaming video from a webcam



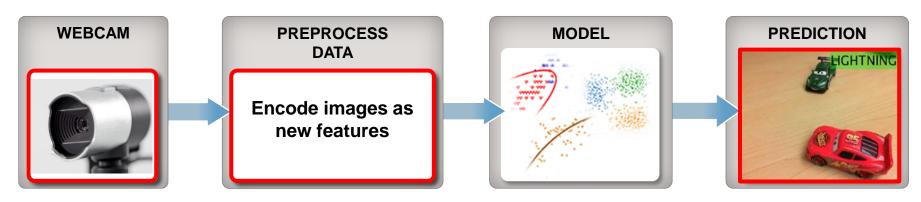


# **Machine Learning Workflow for Example 2**

#### Train: Iterate till you find the best model

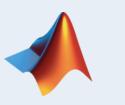


#### **Predict:** Integrate trained models into applications





- Machine Learning
  - What is Machine Learning and why do we need it?
  - Common challenges in Machine Learning
- Example 1: Human activity learning using mobile phone data
  - Learning from sensor data
- Example 2: Real-time car identification using images
   Learning from images
  - Learning from images
- Summary & Key Takeaways





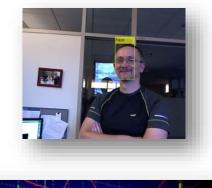
# **MATLAB Strengths for Machine Learning**

Challenge	Solution
Data diversity	Extensive data support Import and work with signal, images, financial, Textual, geospatial, and several others formats
Lack of domain tools	High-quality libraries Industry-standard algorithms for Finance, Statistics, Signal, Image processing & more
Time consuming	Interactive, app-driven workflows Focus on machine learning, not programing
Avoid pitfalls Over Fitting, Speed-Accuracy-Complexity	Integrated best practices Model validation tools built into app Rich documentation with step by step guidance
	Flexible architecture for customized workflows Complete machine learning platform



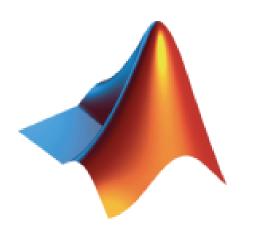
### **Key Takeaways**

- Consider Machine Learning when:
  - Hand written rules and equations are too complex
    - Face recognition, speech recognition, recognizing patterns
  - Rules of a task are constantly changing
    - Fraud detection from transactions, anomaly in sensor data
  - Nature of the data changes and the program needs to adapt
    - Automated trading, energy demand forecasting, predicting shopping trends

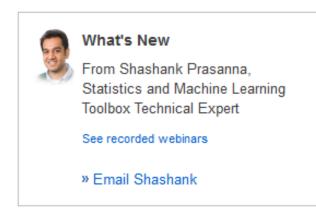




MATLAB for Machine Learning



# Email me if you have further questions





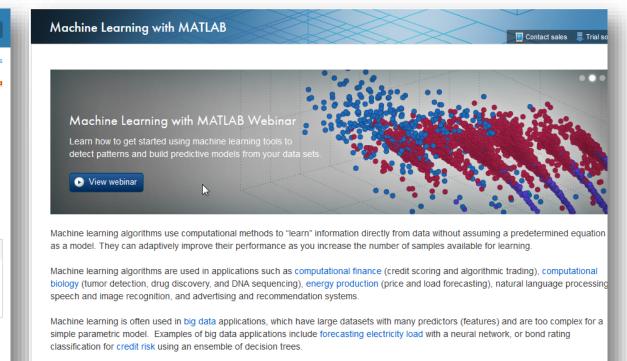
### **Additional Resources**

#### Documentation:

umentation		Search R2015a Documentation	Documentation - Q	
NTENTS Close		🖡 Trial Soft	ware 📮 Product Updates	
ucts	Machine Learning		<b>R</b> 2015a	
cs and Machine Learning	Supervised, unsupervised, and ensemble learning The aim of machine learning is to build a model that makes decisions based on evidence in the presence of uncertainty. As adaptive			
	algorithms identify patterns in data, a computer "learns" from the observations. When exposed to more observations, the computer improves its decision-making performance.			
	In supervised learning, each observation has a corresponding response or label. Classification models learn to predict a discrete class given new predictor data, and regression models learn to predict continuous responses. Applications include spam filters, stock price forecasts, advertisement recommendation systems, and image and speech recognition. The Statistics and Machine Learning Toolbox™			
ania Tanta				
SIOII allu ANOVA	supervised learning functionalities comprise a stream-lined, objective predict responses.	ct framework to train a variety of algorithms effic	ciently, assess models, and	
ine Learning	In unsupervised learning, observations are unlabeled. The goal is	s to learn the structure of the data such as reve	aling natural clusters or	
ervised Learning	variable correlations. Applications include pattern recognition in in	nages and gene expression profiles, identificat	tion of crime hot spots, and	
	microarray data reduction. The Statistics and Machine Learning To k-means clustering, and principal component analysis.	polibox unsupervised learning functionalities in	clude hierarchical and	
emble Learning	Machine Learning Basics			
iate Data Analysis				
al Statistics	Steps in Supervised Learning	What Are Linear Regression Models?		
Up Statistical Computations	Characteristics of Classification Algorithms What Are Classification Trees and Regression Trees?	Introduction to Cluster Analysis Introduction to Feature Selection		
	Supervised Learning Regression, support vector machines, parametric and nonparame	etric classification, decision trees		
		etric classification, decision trees		

Ensembles for boosting, bagging, or random subspace

#### mathworks.com/machine-learning



#### Classification

Was this topic helpful? Yes No

Build models to classify data into different categories.

boosted and bagged decision trees,

#### Regression

Build models to predict continuous data.

Algorithms: support vector machine (SVM),

Algorithms: linear model, nonlinear model, regularization, stepwise regression, boosted

Algorithms: k-means, hierarchical clustering, Gaussian mixture models, hidde

Find natural groupings and

Clustering

patterns in data.

18



# **Q & A**