Lecture 1: Key messages

- What is learning? "Improving skill on a task through experience" …. Task: Classification; skill: Correct classification; experience: data
- Types of learning: Supervised, Unsupervised, Semi-supervised, Reinforcement learning
- Why do machines need to learn? Deluge of data; do not know causal relationships between inputs and outputs; uncertainty; dynamic; knowledge acquisition is the bottleneck
- Applications of Machine learning: Computational Biology; Medical Imaging; multi-model image fusion; Information Retrieval/Natural Language Processing; Text/audio/image retrieval; Robotics Speech processing; Recommender Systems; Scientific Data Analysis; Financial Prediction and Automated Trading; Computer Games, Gambling,….
- AI versus ML versus DL: Computations that enable perception, reasoning and acting. Gen 1: Humans enter knowledge/rules/models and then AI perceives, reasons and acts based on that knowledge; Gen 2: Machine learns knowledge/rules/models from data (ML), but humans still provide features (data transformations); Gen 3: Deep learning: seeks to mimic human brain. Multiple layers of abstraction. Nonlinearities: $g(y)=\max(0,y)$; $\tanh(y)$, sigmoid, $\ln(1+\exp(y))$, softmax, CNN, maxpooling,…
- Why Deep Learning: Data, GPUs, New nonlinearities and signal normalization, CNN & max pooling, dropout, stochastic optimization and GANs, Lots of layers, Much better software tools, Features learned automatically
- Key tasks of ML: Classification, Regression, Clustering
- Things to remember: Make Sure Training Data is Adequate ("Data Matters More than Algorithms"), Make Sure Data is Representative ("No sampling Bias"), Clean up the Data ("Detect and Fix/Remove Outliers", "Ignore/Estimate Missing Data", "Minimize Noise Effects in the Data"), Feature Engineering ("Feature Selection", Feature Extraction", "Seek new Information"), Avoid Overfitting and Underfitting ("Kiss" Principle, "Occam's Razor", "Theory of Parsimony", "Bias-Variance Tradeoff"), Understand the difference between Training Error and Generalization (Outof-sample) Error ("Split Data into Training, Validation and Test Data Sets", "No data snooping", "Test and Validate the Model", "Cross Validation", "Bootstrap", "Model Selection", "Bayesian Model Averaging"), Exploit Domain Knowledge ("Relationships and Constraints", "No Free Lunch Theorem")
- The only way you learn is through experimentation.
- Data statistics: Count , % Missing , Cardinality, Minimum and Maximum, 1st and 3rd Quartiles, Mean and Median, Variance and Standard Deviation, Skewness , Kurtosis, Box plots, Matrix Plots
- Classification, Clustering and Regression
- What Fundamentals you need: Linear Algebra, Probability Theory, Information Theory, Optimization
- Discuss Least squares: $z=Xw+n$
  - $J=||z-Xw||^2 = z'z-2z'Xw+w'X'Xw$; $X=U\Sigma V'$
  - $w=(X'X)^{-1}X'z = V\Sigma^{\dagger}U'z$; $\dagger \Rightarrow$ pseudo inverse
  - zhat=$UU'z$ =Pz…. projection
  - z-zhat=$(I-UU')z=(I-P)z$
  - $X=QR$ …. $w=R^{-1}Q'z$….. Q via Gram-Schmidt, Householder, Givens