

Introduction to Machine Learning

Ranga Raju Vatsavai, Ph.D.

Chancellors Faculty Excellence Associate Professor in Geospatial Analytics
Department of Computer Science, North Carolina State University (NCSU)

Feb. 25-27, 2019

Generative vs. Discriminative

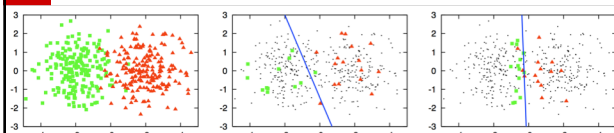
- Generative classifiers
 - Assume some functional form for $P(Y)$, $P(X|Y)$
 - Estimate parameters of $P(X|Y)$, $P(Y)$ directly from training data
 - Use Bayes rule to calculate $P(Y|X)$
 - Examples: NB, MLC/MAP, MRF, ...
- Discriminative classifiers
 - Assume some functional form for $P(Y|X)$
 - Estimate parameters of $P(Y|X)$ directly from training data
 - Examples: Logistics Regression, kNN, NN, ...

2/27/19

© Raju Vatsavai

Active Learning

- Motivation



Actual Data

Small Labeled Data
Decision Boundary is
not accurate

Actively queried for
additional labeled data.
Decision boundary is
better

How do you query for "good" examples?

2/27/19

© Raju Vatsavai

Transfer Learning

- In general, model built to learn one task do not perform on a different (or even similar) task
 - Example, take satellite image from Minnesota, build a decision tree for crop classification, and apply this decision tree to predict crop labels on North Carolina (NC) image
- Transfer learning
 - Take Minnesota (pre-trained) model and fine-tune it to work with NC crop classification

2/27/19

© Raju Vatsavai

Multiple Instance Learning

- Motivation
 - Many learning tasks involves assigning a single label to "group" of instances (e.g., down-town, NCSU campus, airport, ...)

2/27/19 © Raju Vatsavai

Deep Learning

- Motivation
 - Can you learn features directly from input data as "feature engineering" is costly

2/27/19 © Raju Vatsavai Source: Kaggle DL tutorial

Machine Learning: Big Picture

Learning Paradigms:
What data is available and when? What form of prediction?

- supervised learning
- unsupervised learning
- semi-supervised learning
- reinforcement learning
- active learning
- imitation learning
- domain adaptation
- online learning
- density estimation
- recommender systems
- feature learning
- manifold learning
- dimensionality reduction
- ensemble learning
- distant supervision
- hyperparameter optimization

Problem Formulation:
What is the structure of our output prediction?

- boolean Binary Classification
- categorical Multiclass Classification
- ordinal Ordinal Classification
- real Regression
- ordering Ranking
- multiple discrete Structured Prediction
- multiple continuous (e.g. dynamical systems)
- both discrete & cont. (e.g. mixed graphical models)

Application Areas
Key challenges?
NLP, Speech, Computer Vision, Robotics, Medicine, Search

Theoretical Foundations:
What principles guide learning?

- probabilistic
- information theoretic
- evolutionary search
- ML as optimization

Facets of Building ML Systems:
How to build systems that are robust, efficient, adaptive, effective?

1. Data prep
2. Model selection
3. Training (optimization / search)
4. Hyperparameter tuning on validation data
5. (Blind) Assessment on test data

Big Ideas in ML:
Which are the ideas driving development of the field?

- inductive bias
- generalization / overfitting
- bias-variance decomposition
- generative vs. discriminative
- deep nets, graphical models
- PAC learning
- distant rewards

2/27/19 © Raju Vatsavai Source: Prof. Matt Gormley @ CMU