



Dimensions of Machine Learning

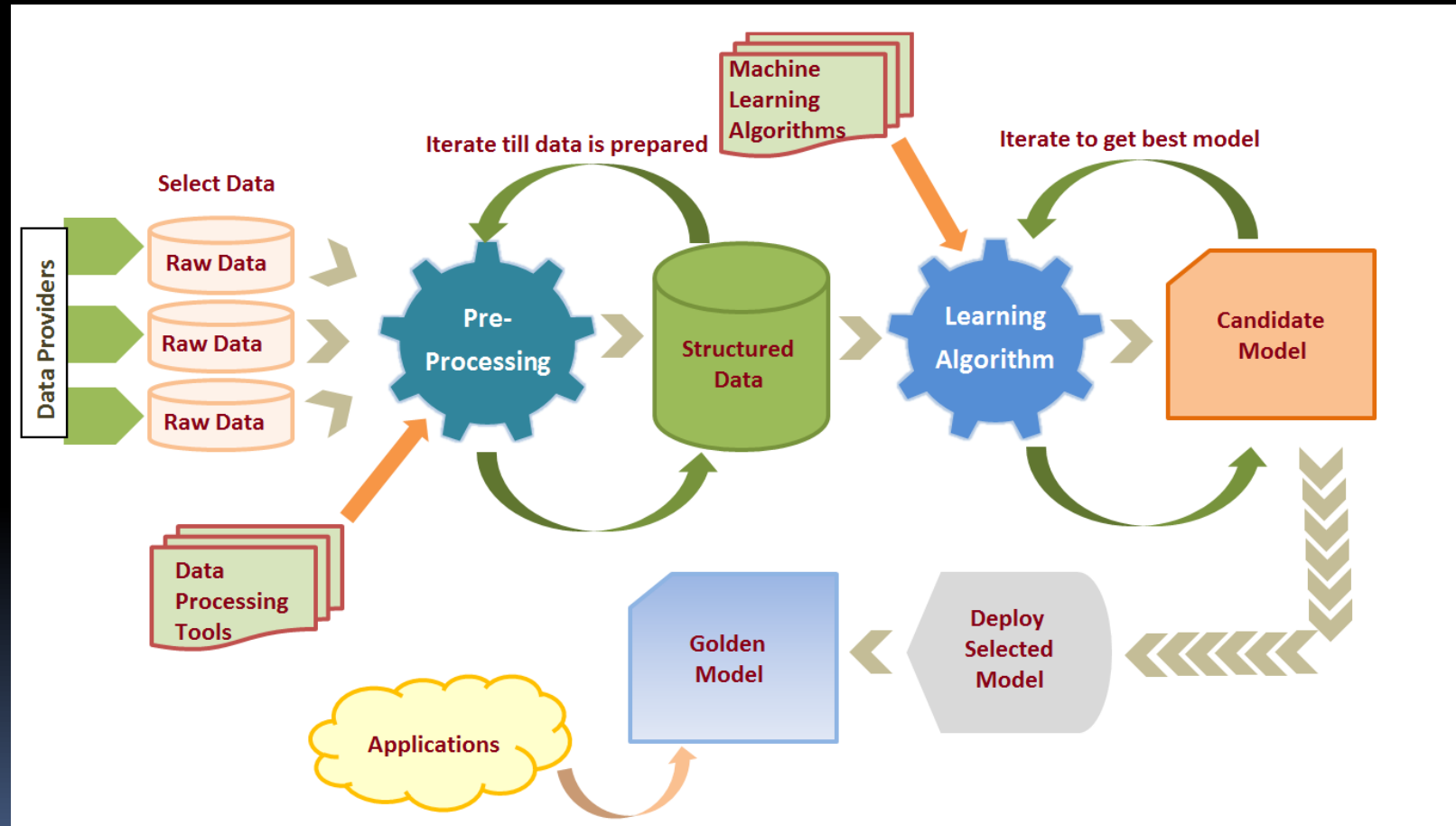
CSCI 447/547 MACHINE LEARNING



Outline

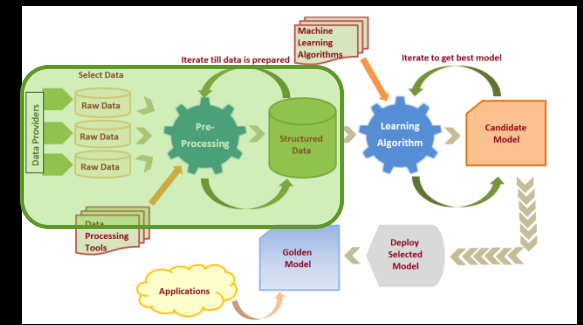
- Input
- Algorithm
- Output
- Evaluation
- Other

Generic Machine Learning Process



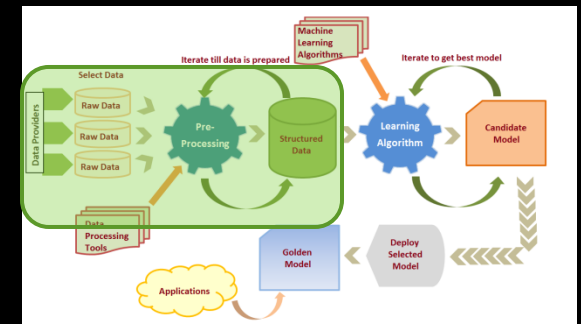
Input: Features

- Redundancy
 - Correlated Features
- Irrelevance
- Sufficiency



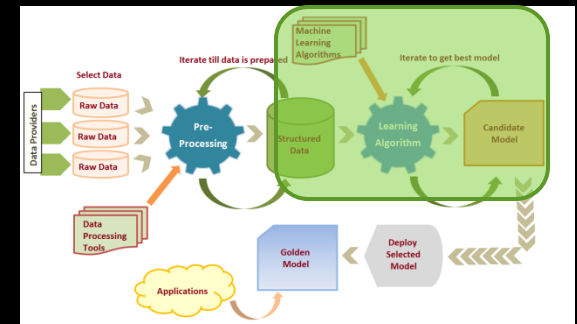
Input: Data

- Data Type
 - Numeric
 - Symbolic
 - Mixed
 - Multi-Valued
 - Unstructured
- Transformations
 - Regularization
 - Imputation of Missing Values
- Noise
- Granularity

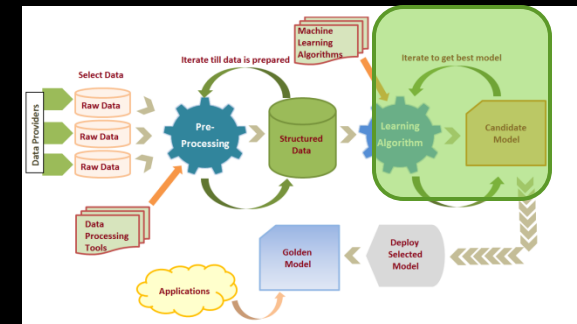


Algorithm(s)

- Supervised vs. Unsupervised
- Incremental vs. One-Shot
- Symbolic vs. Sub-symbolic
- Convergence (or not)
- Ensemble (or not)



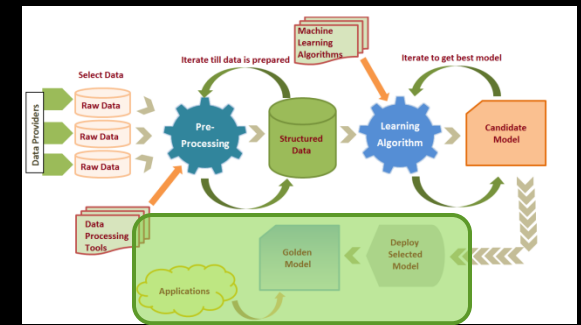
Output



- Output Transformation
- Numeric or Symbolic
- Usefulness – Substantive vs. Trivial

Other

- Concept drift
- Maintenance
- Operation Level
 - Advisor
 - Assistant
 - Actor



Summary

- Input
- Algorithm
- Output
- Evaluation
- Other

