

10-605/10-805: Machine Learning with Large Datasets

Fall 2022

Course Summary

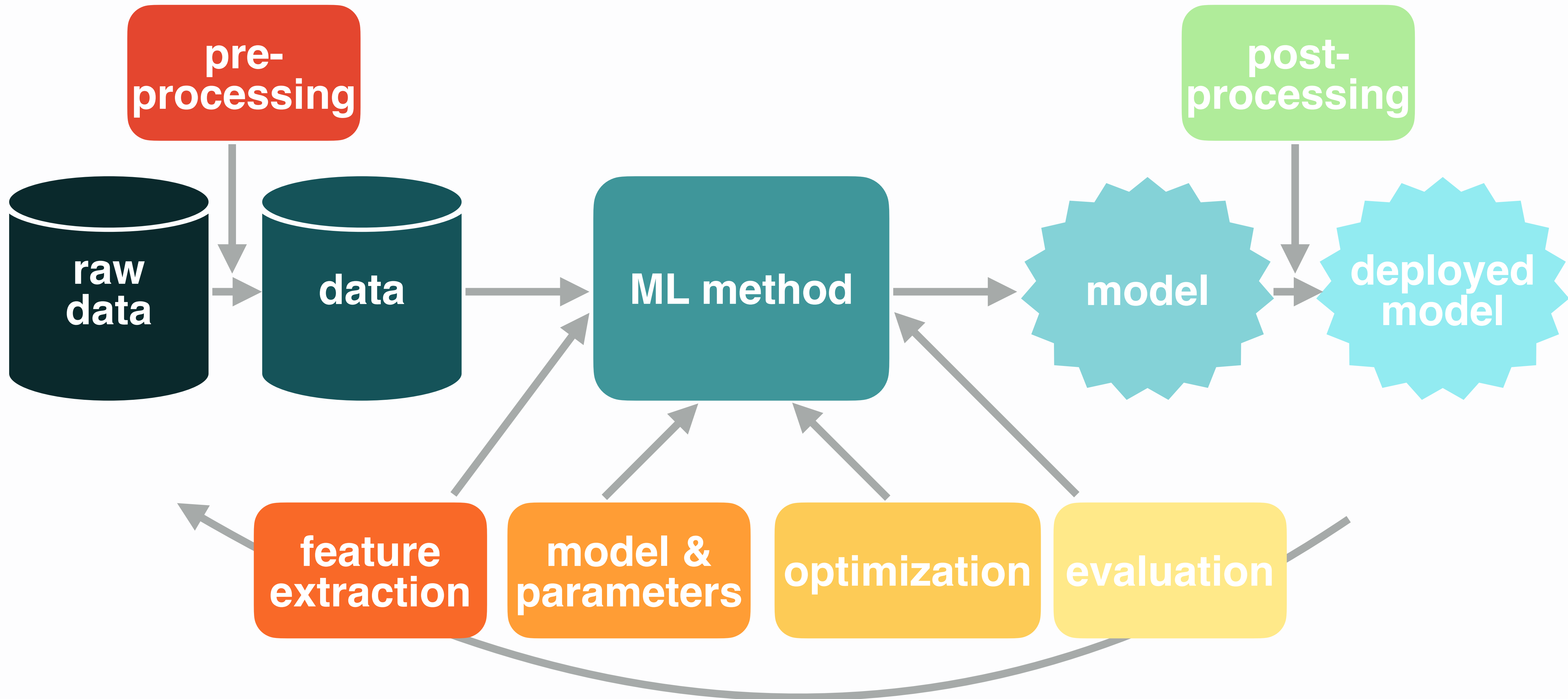
MOTIVATION:

Machine Learning with Large Datasets

machine learning is: the study of methods that
improve their performance
on some task
with experience

AN EVEN MORE REALISTIC

Machine learning pipeline



MOVING ON ...

Current concerns in ML

How can we perform the pipeline we just considered, but under some constraints?

- privacy
- fairness
- robustness
- **learning at scale**

Key course topics

Data preparation

- Data cleaning
- Data summarization
- Visualization
- Dimensionality reduction

Training

- Distributed ML
- Large-scale optimization
- Scalable deep learning
- Efficient data structures
- Hyperparameter tuning

Inference

- Hardware for ML
- Techniques for low-latency inference (compression, pruning, distillation)

Infrastructure / Frameworks

- Apache Spark
- TensorFlow
- AWS

Advanced topics

- Neural Architecture Search
- Federated learning

Approximations & Trade-Offs

- Dimensionality reduction: compactness vs representation
- Kernel approximation: rank vs reconstruction error
- Hashing: hash table size vs accuracy
- Trees: number of thresholds vs suboptimality
- Hardware: cost vs efficiency
- Optimization: per-iteration cost vs convergence rate, computation vs communication, ease of use vs accuracy
- HP tuning: budget vs covering
- Compression: size/latency vs accuracy
- Federated learning: communication vs accuracy, privacy vs utility

KEY IDEAS

ML & Systems

- Analyzing trade-offs
- Identifying key bottlenecks
- Understanding hardware / software / network constraints
- Developing abstractions
- Optimizing repeated procedures
- Co-designing methods & systems

MLSys Conference (www.mlsys.cc)

Topics of interest include, but are not limited to:

- Efficient model training, inference, and serving
- Distributed and parallel learning algorithms
- Privacy and security for ML applications
- Testing, debugging, and monitoring of ML applications
- Fairness, interpretability and explainability for ML applications
- Data preparation, feature selection, and feature extraction
- ML programming models and abstractions
- ML compilers and runtime
- Programming languages for machine learning
- Visualization of data, models, and predictions
- Specialized hardware for machine learning
- Hardware-efficient ML methods
- Machine Learning for Systems
- Systems for Machine Learning

Course Evaluations

- Course evaluations will be administered by the university registrar
- Look for an email soon with a course survey
- We'd love your feedback in order to help improve the course!

THANKS TO ...

Previous instructors, content developers

- Virginia Smith, CMU
- Heather Miller, CMU
- William Cohen, Google AI
- Anthony Joseph, UC Berkeley
- Barnabas Póczos, CMU

HUGE THANKS TO ...

Co-instructors

- Henry Chai
- Daniel Bird (Education Associate)

HUGE THANKS TO ...
TAs

- Rahul Dharani
- Kunal Dhawan
- Nikhil Gupta
- Ramya Ramanathan
- Ruben John Mampilli
- Mehak Malik
- Utsav Dutta
- Preksha Patel
- Cristian Challu

Thank you!

Exam II on Thursday

- The exam is NOT cumulative (covers everything from 10/13 and after)
- Exam will take place in-person during class time
- 20 MC, 12 SA questions
 - Note: no partial credit for MC questions

Q&A