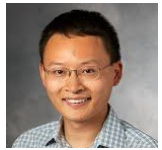
Making AI forget you: Data deletion in machine learning

TONY GINART
MELODY GUAN, GREG VALIANT, JAMES ZOU

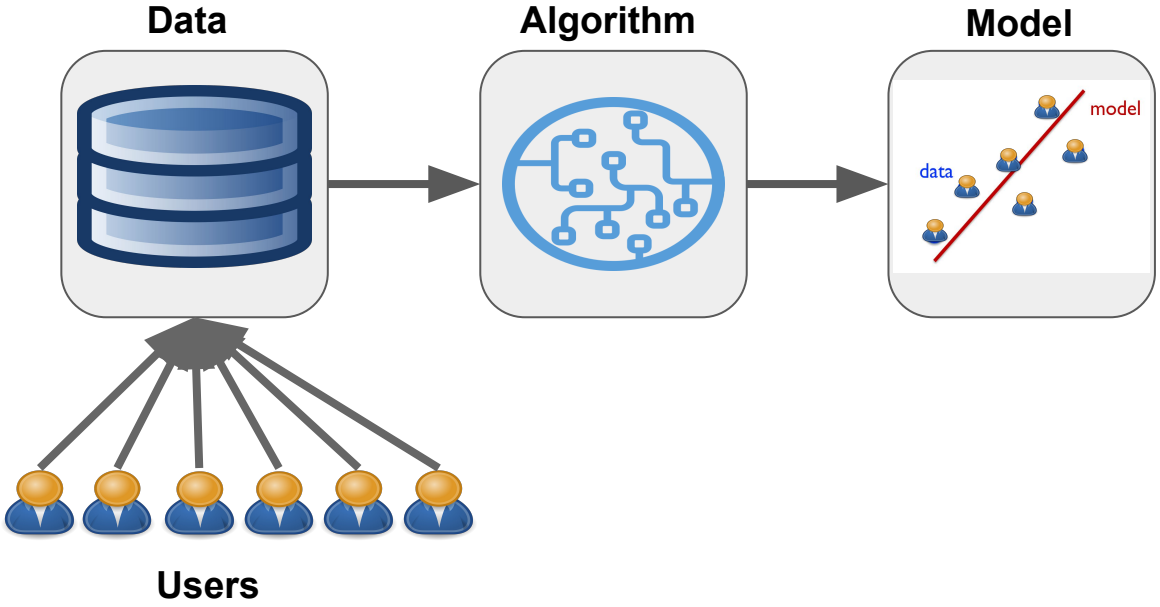




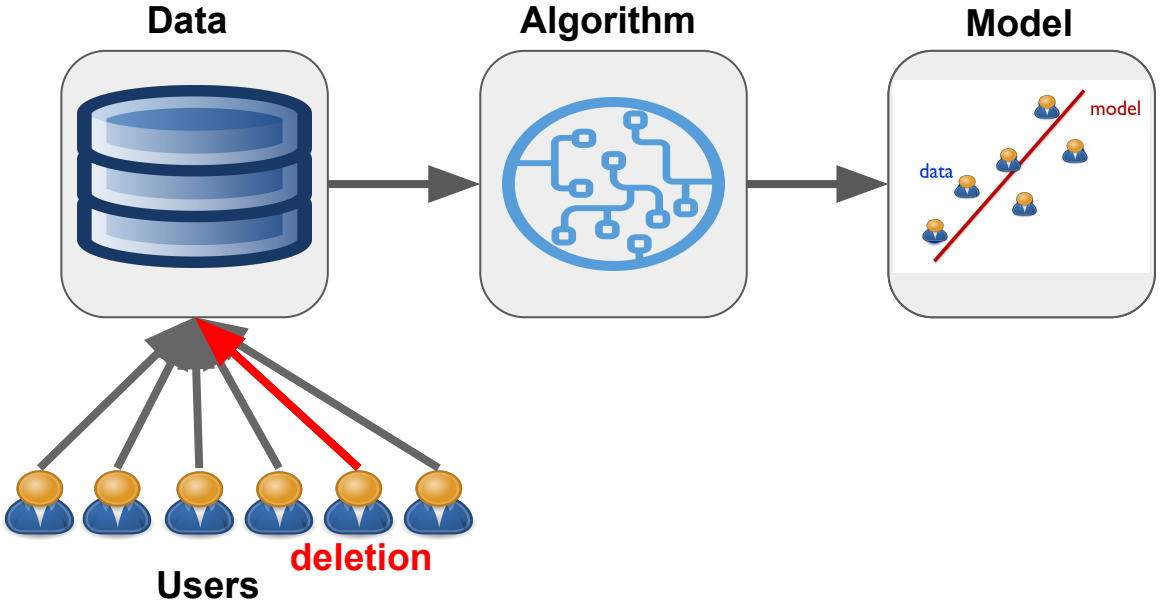
Advances in Neural Information Processing Systems

December 12, 2019

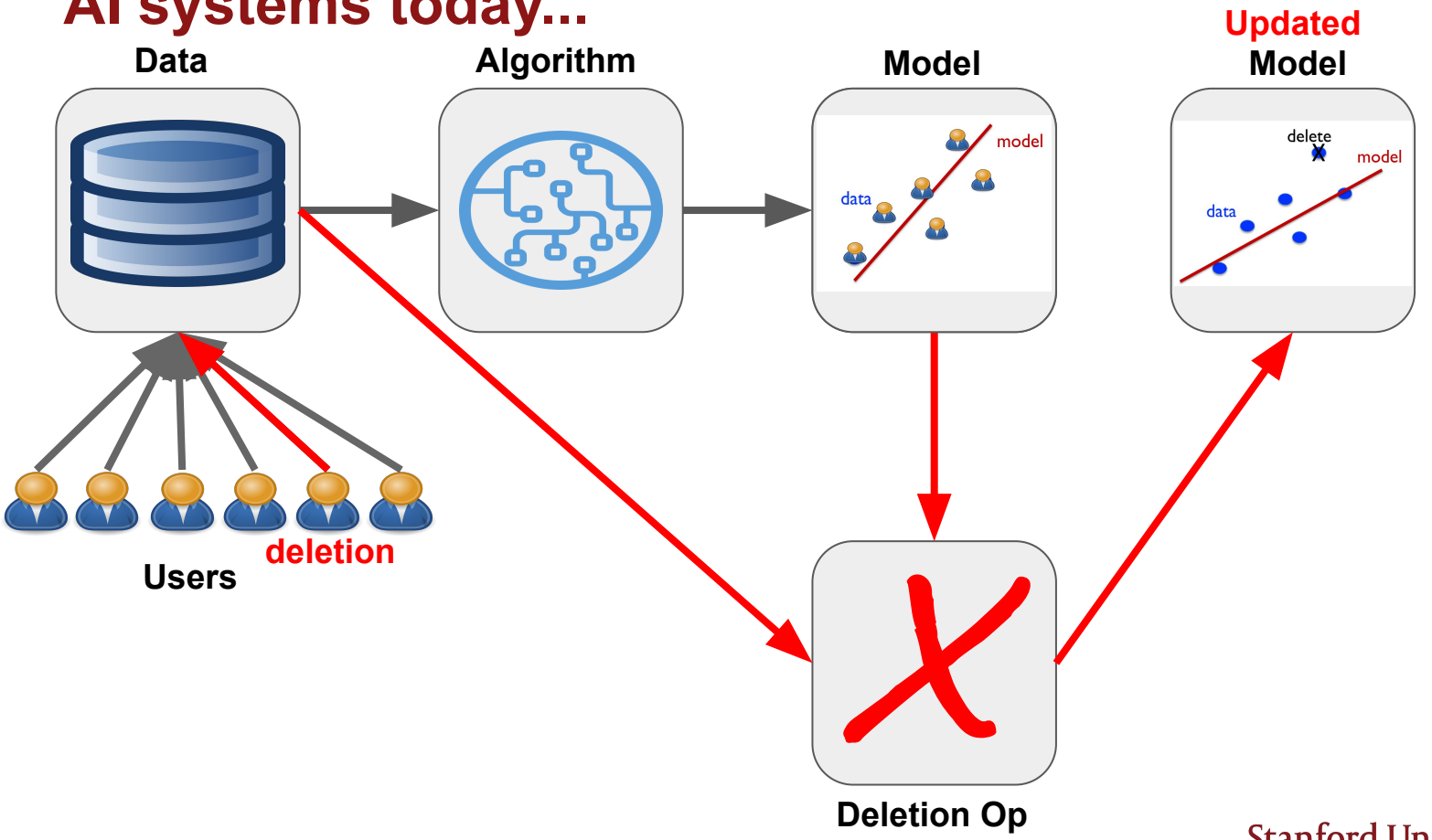
AI systems today...



AI systems today...



AI systems today...



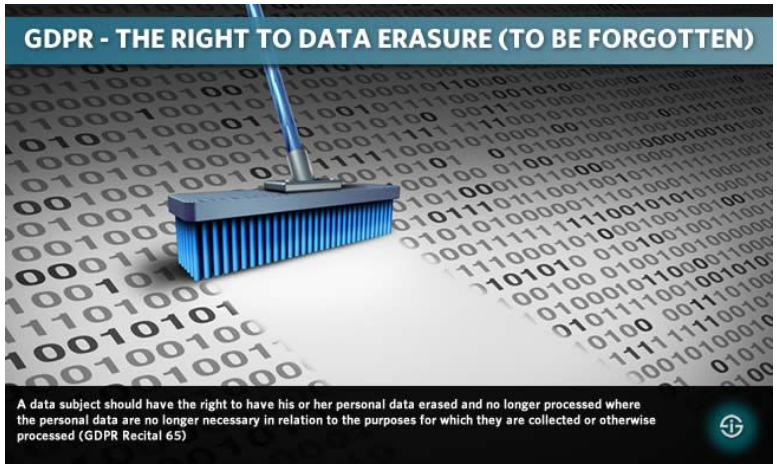
Deletion requests in the wild...

EMAIL ---- UK BIOBANK ----

Subject: UK Biobank Application [REDACTED], Participant Withdrawal Notification [REDACTED]

Dear Researcher,

As you are aware, participants are free to withdraw form the UK Biobank at any time and request that their data no longer be used. Since our last review, some participants involved with Application [REDACTED] have requested that their data should longer be used.



Contributions

1) Define deletion in ML system and notion of efficient deletion

2) Propose general principles for co-design of ML algorithms and deletion operations

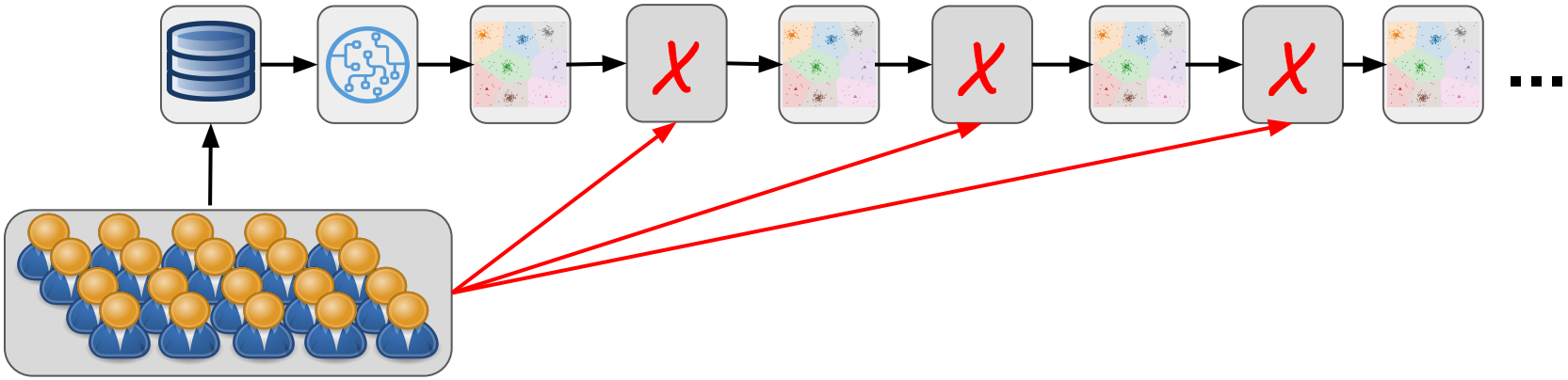
3) Introduce deletion efficient unsupervised learning

What is “data deletion” for an ML system?

Informal definition: Deleting a data point from a trained ML model means updating the model as if this point had never existed.

What is “deletion efficiency” for an ML system?

- Setting: online deletion requests from users
- Figure-of-Merit: amortized computation



Toolbox for deletion efficient ML

- **Linearity**: fast *O(1)* deletion with respect to *n* data points
- **Laziness**: E.g. nearest neighbors
- **Modularity**: Control dependency from data to parameters
- **Quantization**: Efficiently check if deletion matters

State of progress

Supervised learning:

- Linear regressions/models
- Non-parameteric (k-NN)
- Incremental SVMs

Unsupervised learning:

- 1) Quantized k-means
- 2) Divide-and-Conquer k-means

State of progress

Supervised learning:

- Linear regressions/models
- Non-parameteric (k-NN)
- Incremental SVMs

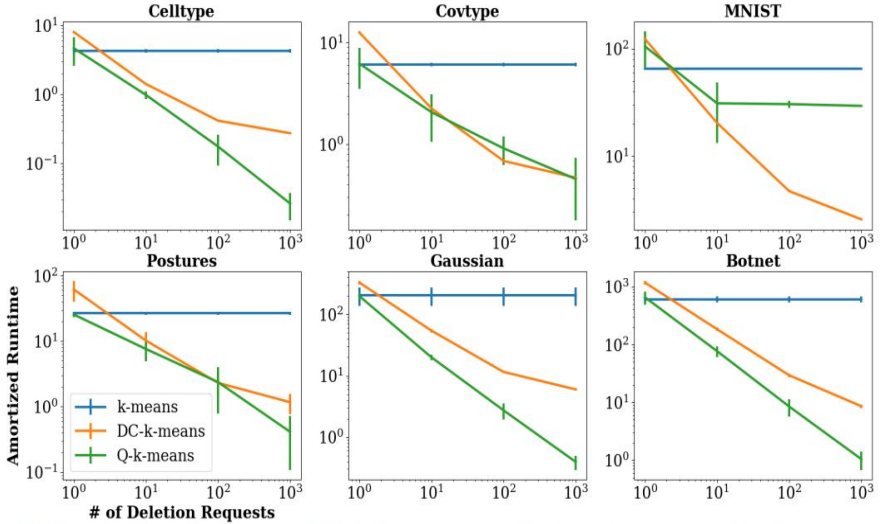


Figure 1: Online deletion efficiency: # of deletions vs. amortized runtime (secs) for 3 algorithms on 6 datasets.

Unsupervised learning:

- 1) Quantized k-means
- 2) Divide-and-Conquer k-means

100X faster deletion without loss of clustering quality

Next steps in deletion efficient ML

Models:

- Decision trees/forests
- Artificial neural networks

Settings:

- Approximate deletions
- Adversarial requests

Paradigms:

- Reinforcement learning
- Representation/embedding learning

Want to know more?

Poster session @ 5pm
#123, East Exhibition Hall B + C

Thank you!

Happy to chat more:
tginart@stanford.edu