

#### Lecture Outline

- Responsive vs. data collection applications
- Edge in data collection applications: two views
- Reducing data transmission with edge
- <u>Wednesday</u>: Edge and machine learning
  > ML training with edge involvement
  > Personalized ML models with edge

2

#### **Responsive Applications**

- Also sometimes called "interactive applications"
- Often single device
- Speed of reaction matters



#### **Data Collection Applications**

- Slower reaction time
  Speed of operation less of a concern
   No in-the-loop control
- Often look across multiple devices







- Images
- Clicks
- System performance



### Turning Data Collection Applications into Responsive Applications?

• Near-real-time reaction to traditionally long-term logged data is possible in some cases



#### Turning Data Collection Applications into Responsive Applications?

 Some decision-making and response timelines do not call for fast responses

8



- Responsive vs. data collection applications
- Edge in data collection applications

- Traditional pipeline view: reducing data transmission with edge
- Comprehensive view: edge and machine learning
  ML training with edge involvement
  Personalized ML models with edge
- 9







## Data Collection Applications: With Edge Devices and Cloudlets







- Responsive vs. data collection applications
- Edge in data collection applications
- Traditional pipeline view: reducing data transmission with edge
  - Data reduction
  - Data reduction + edge for data storage
- · Comprehensive view: edge and machine learning
  - > ML training with edge involvement
  - Personalized ML models with edge

#### Opportunities for Improvement

- Wasteful to transmit data that is ultimately not used
- Classic data reduction techniques all apply





Duke

#### Classic Data Reduction Techniques on the Edge: Filtering

- Also called selective forwarding
- Send only important data to the cloud
- E.g., :

Duke

- ➢ Of a certain type
- ➢ GEQ, LEQ, in range, …
- Supported by AWS Greengrass

19



#### Tradeoff: Data Storage Quality

- Cloud-stored data may be reduced in density and richness in the long term
   Unless the reduction techniques are fully lossless
- Lose the ability to data-mine for perpetuity

#### Classic Techniques are Important in Practice

- Reduce operating expenses
- Could be one of the main drivers for edge deployments

22

#### Tradeoff: Dealing With Changes

- Existing techniques are largely non-adaptive
  - Predefined LEQ, GEQ, ... assume that data properties are stationary
- Concept drifts need to be detected and dealt with

Detecting changes

Periodically updating rules for them

#### Lecture Outline

- Responsive vs. data collection applications
- · Edge in data collection applications
- Traditional pipeline view: reducing data transmission with edge
  - Data reduction
  - Data reduction + edge for data storage
- · Comprehensive view: edge and machine learning
  - > ML training with edge involvement
  - Personalized ML models with edge

#### Classic Techniques in Collaboration with Data Storage Policies: Video Example

- Send reduced "important data" from a video to the cloud
- Keep full video available locally for a prespecified number of days
- Send the full video to the cloud later on if required

Dukeuniversity

# <text><text><text><text>



