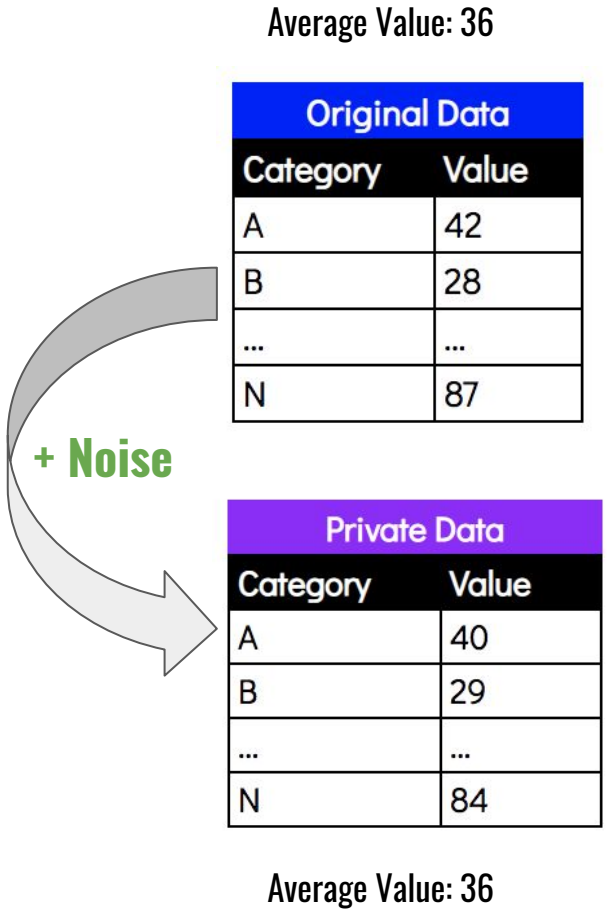
Institute for
Applied Computational Science
HARVARD SCHOOL OF ENGINEERING AND APPLIED SCIENCES

Microsoft

Fairness Impact of Privacy

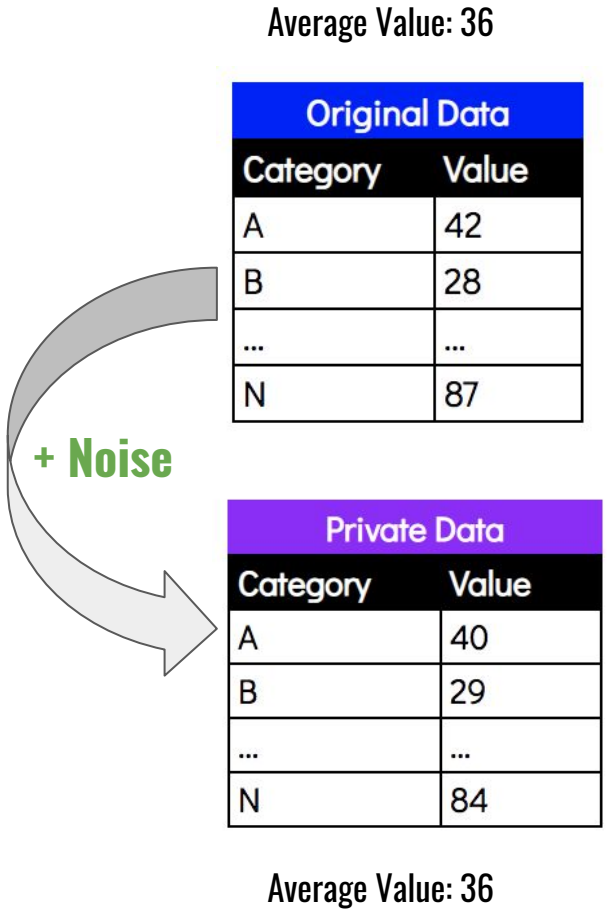
Primer on Differential Privacy

- Data inherently contains **sensitive information** about individuals
  - Even when certain primary identifiers are removed
- Differential privacy protects sensitive information by adding **noise,** while retaining as many statistical characteristics as possible



Primer on Differential Privacy

- Data inherently contains **sensitive information** about individuals
  - Even when certain primary identifiers are removed
- Differential privacy protects sensitive information through the addition of **noise,** while retaining as many statistical characteristics as possible.

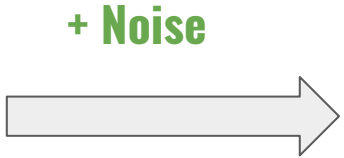


Primer on Differential Privacy

- But – mathematically impossible to preserve the full level of detail while guaranteeing privacy!

Original Data	
Age	State
23	NY
47	NE
35	NY
29	CT
...	...
52	CT

Average Age: 44
State: 0.8% NE

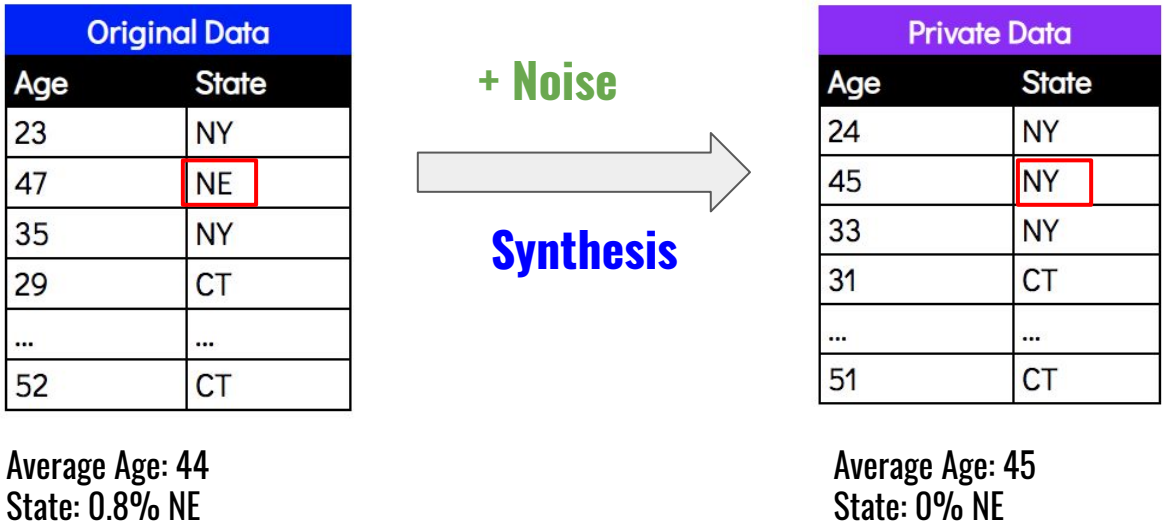


Private Data	
Age	State
24	NY
45	NY
33	NY
31	CT
...	...
51	CT

Average Age: 45
State: 0% NE

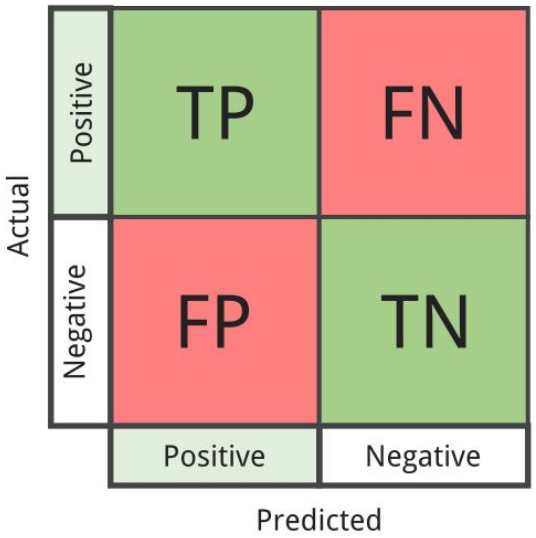
Primer on Differential Privacy

- But – mathematically impossible to preserve the full level of detail while guaranteeing privacy!



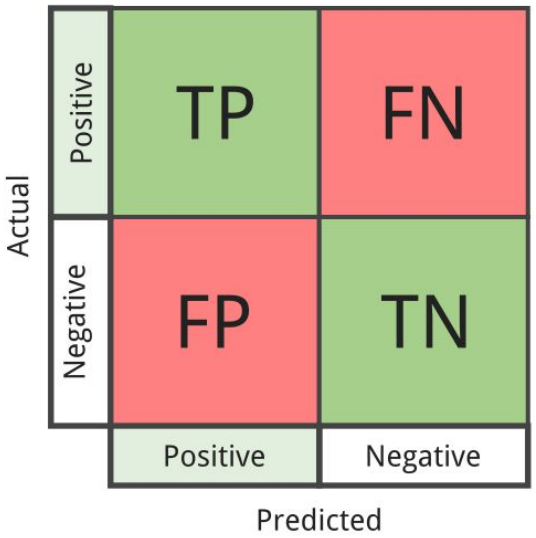
Differential Privacy and Fairness

- Differential privacy often has a **disparate impact** on model accuracy
- Can amplify existing biases and degrade **fairness**
- Tradeoffs between privacy and fairness may be necessary
- These tradeoffs are not well understood in the realm of differentially private *synthetic* data



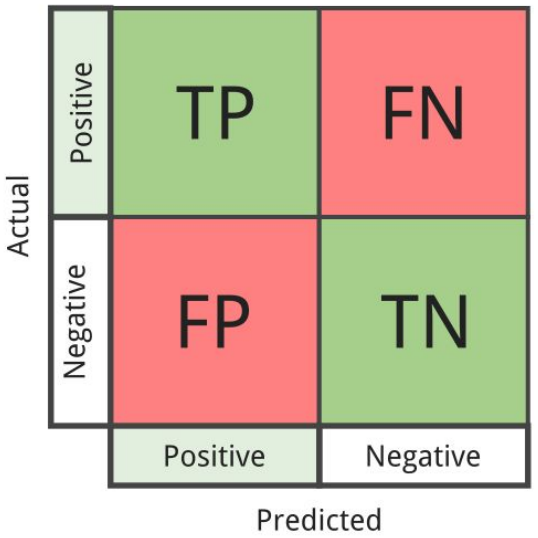
Differential Privacy and Fairness

- Differential privacy often has a **disparate impact** on model accuracy
- Can amplify existing biases and degrade **fairness**
- Tradeoffs between privacy and fairness may be necessary
- These tradeoffs are not well understood in the realm of differentially private *synthetic* data



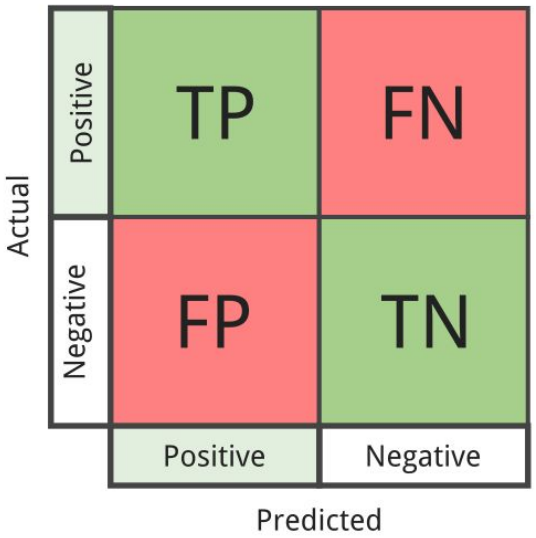
Differential Privacy and Fairness

- Differential privacy often has a **disparate impact** on model accuracy
- Can amplify existing biases and degrade **fairness**
- Tradeoffs between privacy and fairness may be necessary
- These tradeoffs are not well understood in the realm of differentially private *synthetic* data



Differential Privacy and Fairness

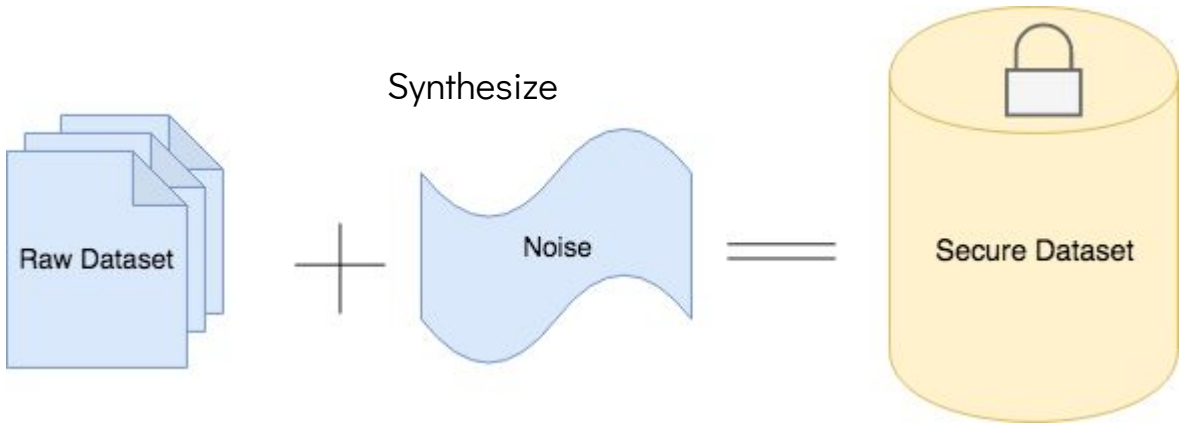
- Differential privacy often has a **disparate impact** on model accuracy
- Can amplify existing biases and degrade **fairness**
- Tradeoffs between privacy and fairness may be necessary
- These tradeoffs are not well understood in the realm of differentially private *synthetic* data



SmartNoise

Synthesizer:

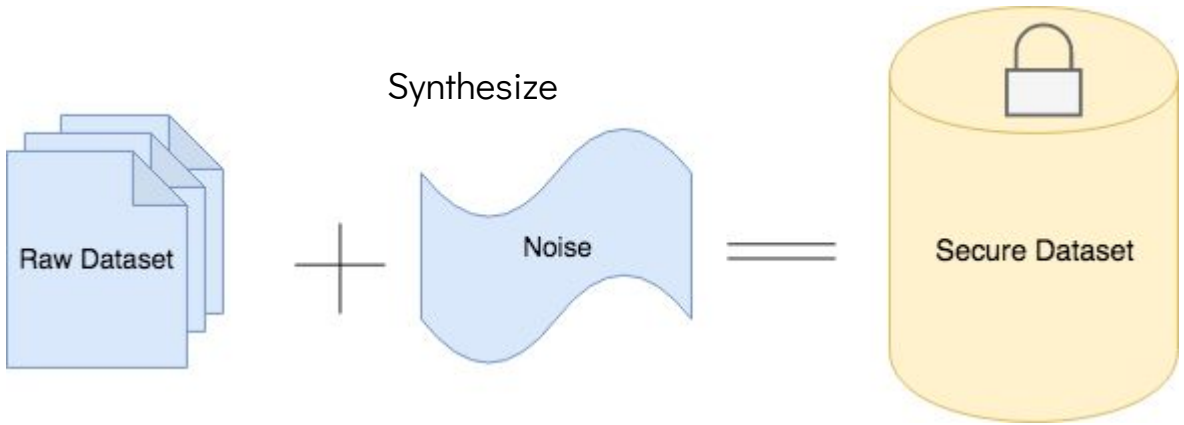
- MWEM
- DP-CTGAN,
- PATE-GAN ..



SmartNoise

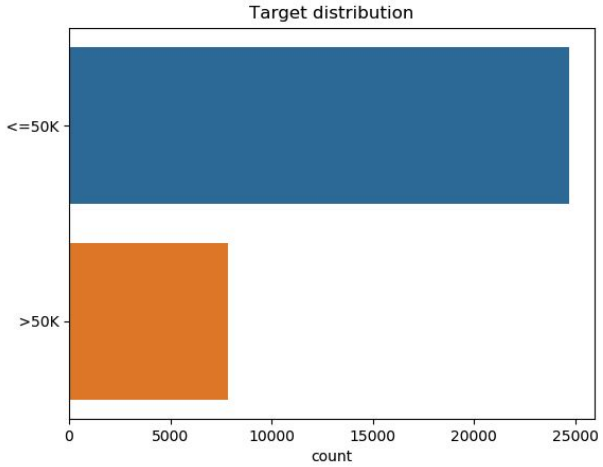
Synthesizer:

- MWEM
- DP-CTGAN,
- PATE-GAN ..



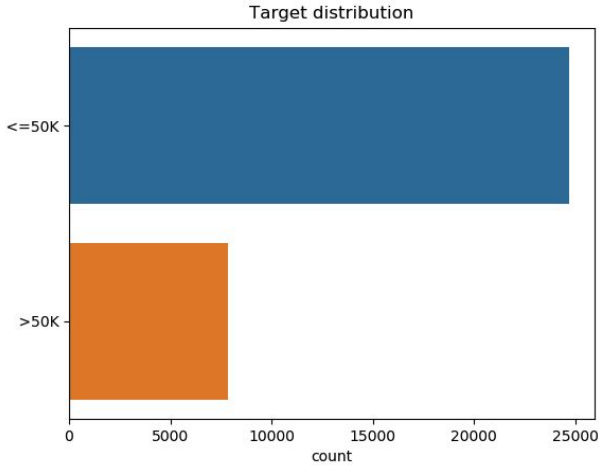
Datasets

- Adult Dataset (Census Income" dataset)
- COMPAS
- German Credit Dataset

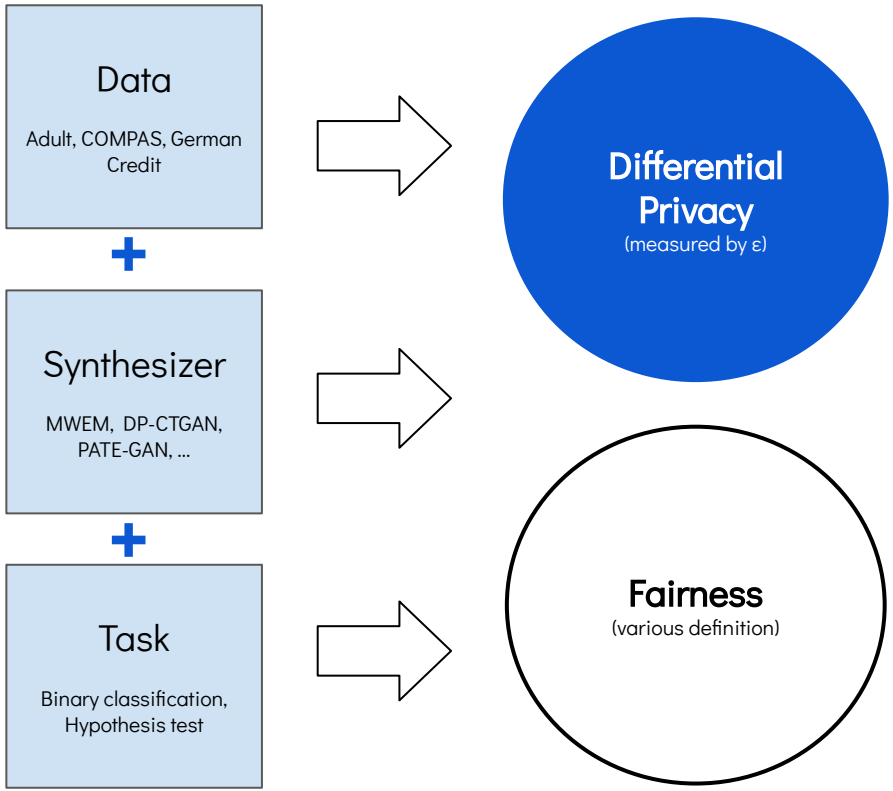


Datasets

- Adult Dataset (Census Income" dataset)
- COMPAS
- German Credit Dataset



Action Plan



Understand the tradeoff between privacy and fairness using various dataset and synthesizers through tasks like binary classification.

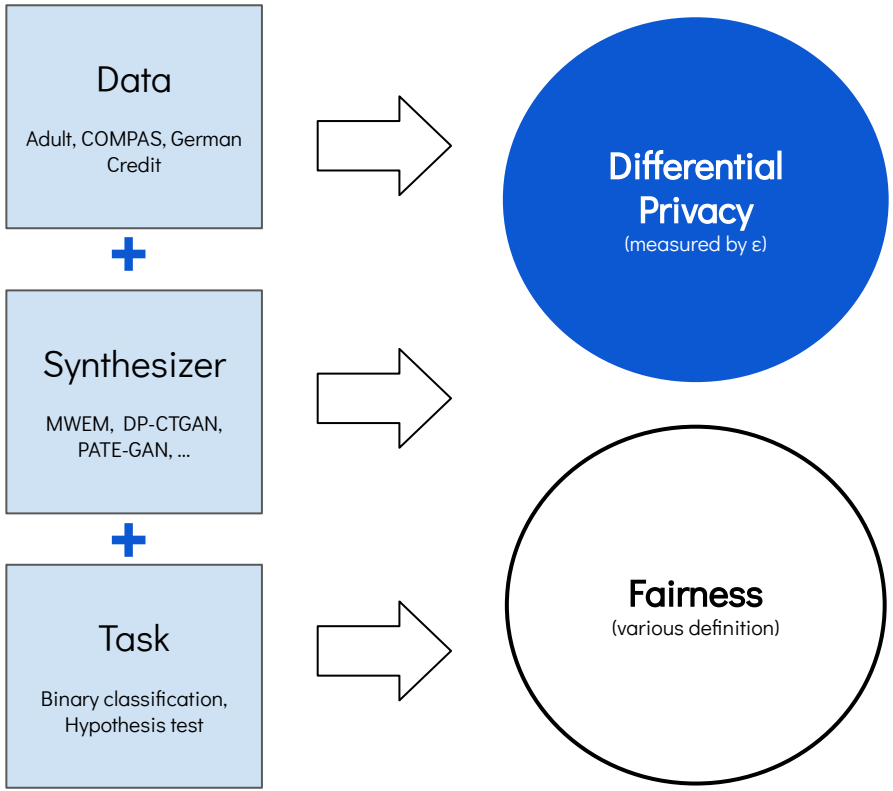
Milestone 1: Prepare fairness evaluation pipelines on data

Milestone 2: Compare and analyze results and understand the conditions that lead to good or bad results

Milestone 3: Recommend or develop pre/post-processing steps that mitigates the bias we observe

Final deliverable: Write paper and prepare the final presentation

Action Plan



Understand the tradeoff between privacy and fairness using various dataset and synthesizers through tasks like binary classification.

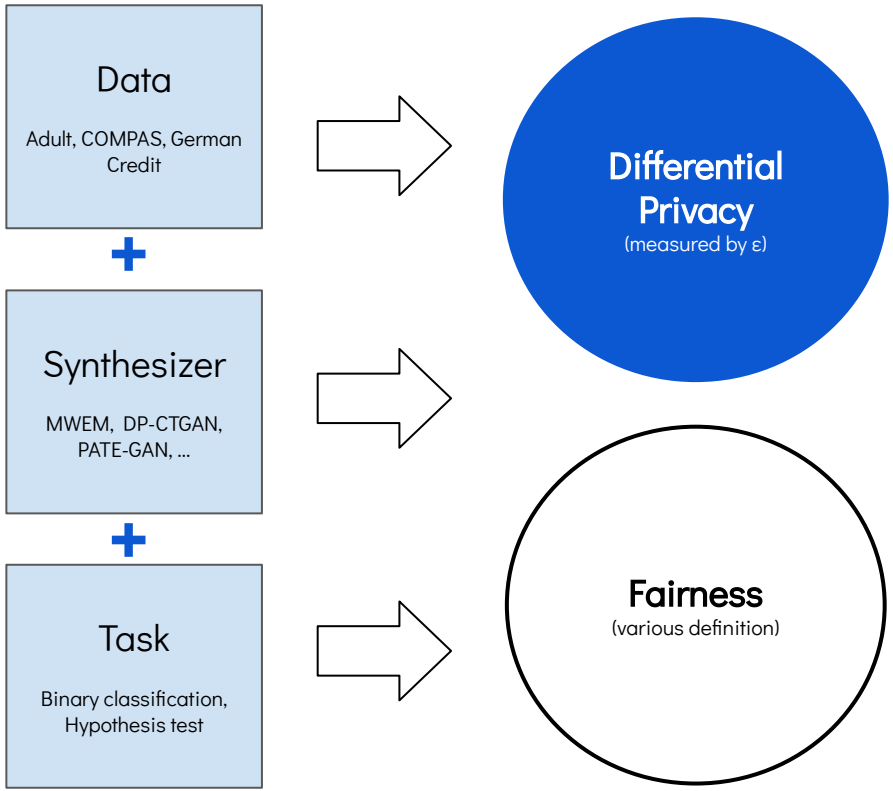
Milestone 1: Prepare fairness evaluation pipelines on data

Milestone 2: Compare and analyze results and understand the conditions that lead to good or bad results

Milestone 3: Recommend or develop pre/post-processing steps that mitigates the bias we observe

Final deliverable: Write paper and prepare the final presentation

Action Plan



Understand the tradeoff between privacy and fairness using various dataset and synthesizers through tasks like binary classification.

Milestone 1: Prepare fairness evaluation pipelines on data

Milestone 2: Compare and analyze results and understand the conditions that lead to good or bad results

Milestone 3: Recommend or develop pre/post-processing steps that mitigates the bias we observe

Final deliverable: Write paper and prepare the final presentation

Institute for
Applied Computational Science
HARVARD SCHOOL OF ENGINEERING AND APPLIED SCIENCES

Microsoft

Thank you!