Privacy Preserving AI Implementation

By Jordan Foss, Supervised By Dan Kim

PROJECT GOALS:

- Identify and investigate noise adding mechanisms for LDP
- Implement these mechanisms on a sample dataset
- Construct a model for comparing the effectiveness of models trained on data perturbed by these mechanisms
- Identify the strengths and weaknesses of the mechanisms on various ML algorithm

PROJECT ACHIEVEMENTS:

- Three LDP mechanisms were implemented
- Six machine learning algorithms were trained on this noisy data
- Model created capable of training and comparing different LDP mechanisms under different ML algorithms with different privacy budgets

BACKGROUND:

Data Privacy Vs Security:

Data Privacy

Compliance with data protection laws and regulations. Focus on how to collect, process, share, archive and delete the data

Data Security

Measures that an organzation is taking in order to prevent any third party from unauthorized access.

Epsilon Differential Privacy:

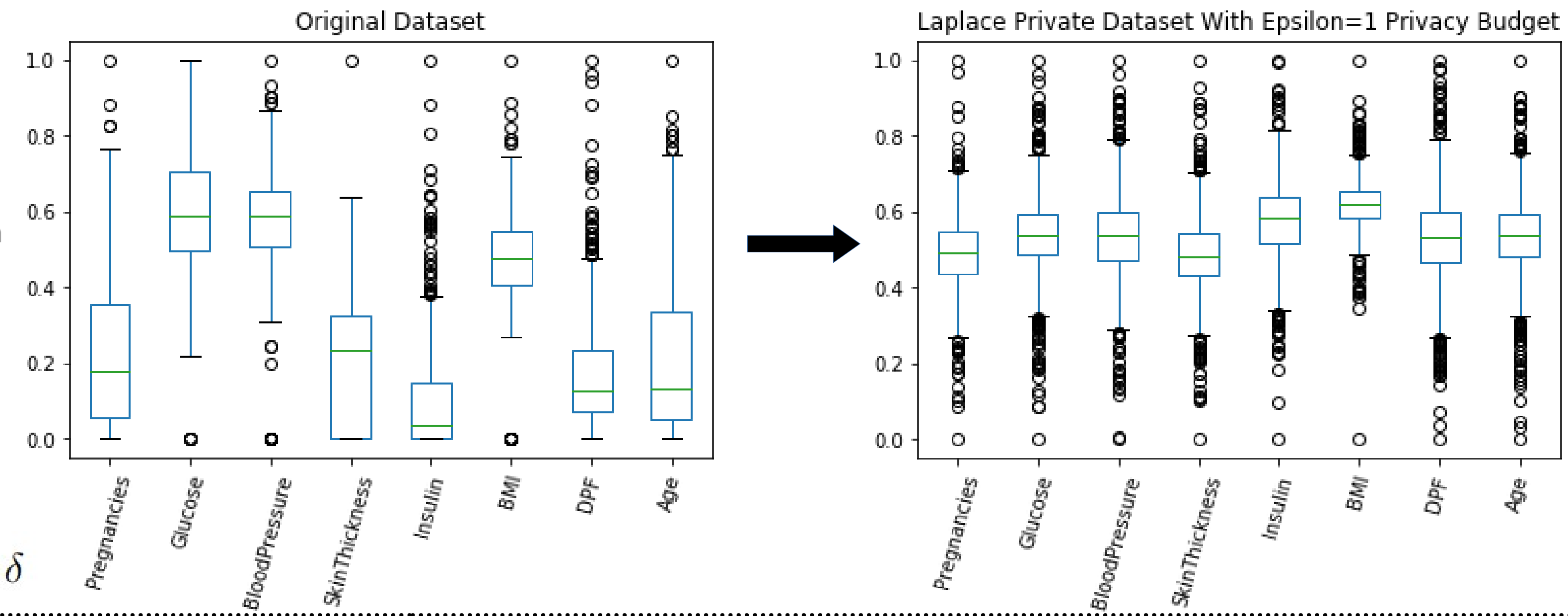
Noise must satisfy definition to be considered private

$$Pr[\mathcal{M}(x) \in \mathcal{S}] \leq \exp(\epsilon) \cdot Pr[\mathcal{M}(y) \in \mathcal{S}] + \delta$$

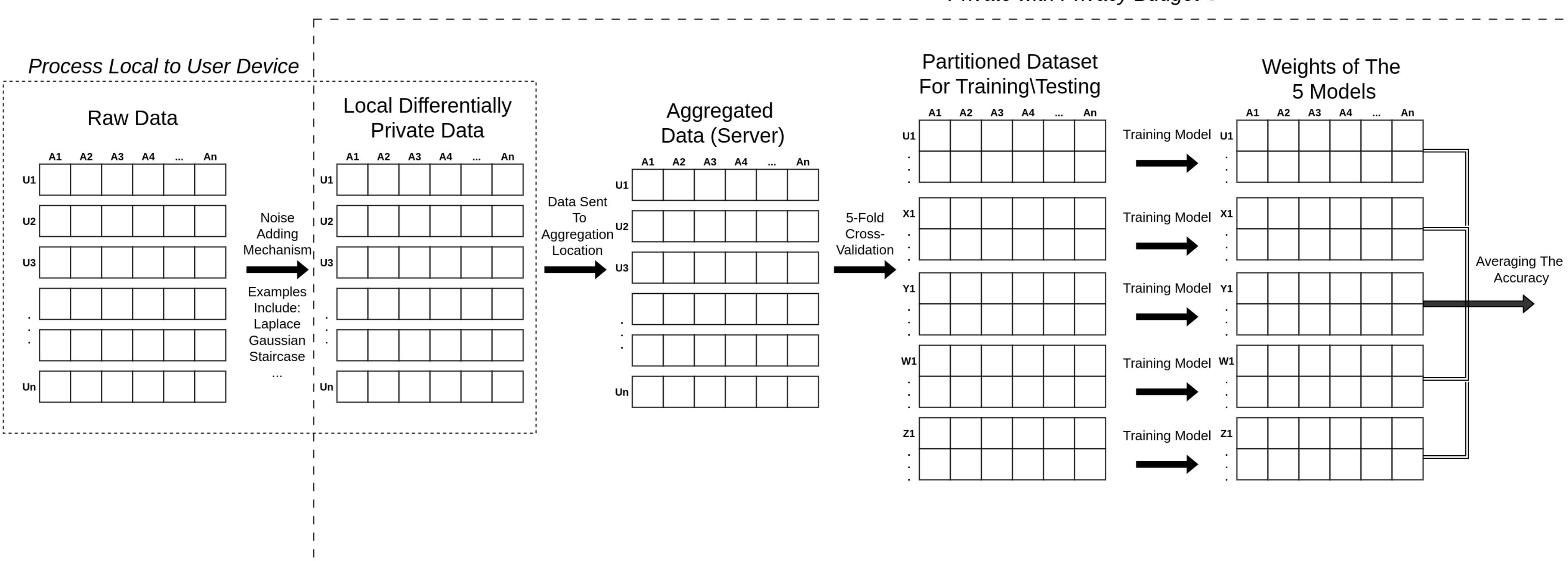
Local Differential Privacy (LDP):

Perturb the data with random noise

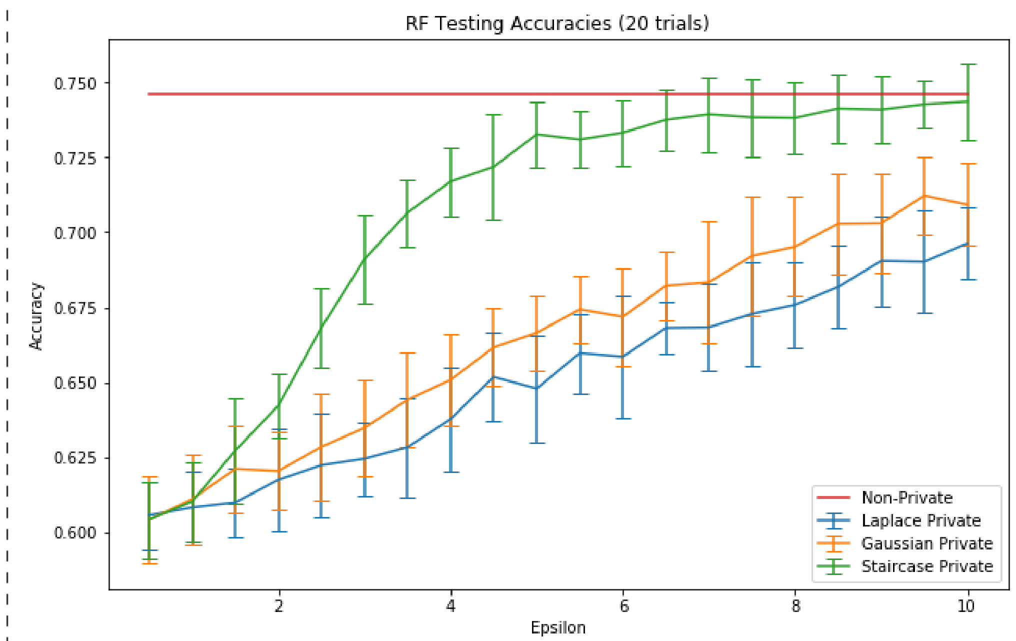
Mechanism (e.g Laplace) used to supply the random noise



SOFTWARE MODEL DEVELOPED:



Model accuracy for varying level of privacy:



The University Of Queensland Australia

School of Information Technology & Electrical Engineering

INNOVATION EXPO