DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

Differential Privacy in Medical Analysis

PhD Course Stat4Engineers

Padova, September 30 2024

Francesco L. De Faveri
Laura Menotti
Department of Information Engineering, UniPD

Scenario, Problem Definition and Methodology

Privacy Risk

When analyzing personal information, there is a Privacy risk of disclosing the sensitive information of the outliers patients.

Research Question

*"Is it possible to apply Differential Privacy during medical analysis and obtain similar results? If **yes**, at what cost?"*

Methodology

During training we add noise sampled from a Laplacian distribution parameterized by the privacy budget ε, so that:

$$X_{\text{training}} = X_{\text{training}} + \nu$$

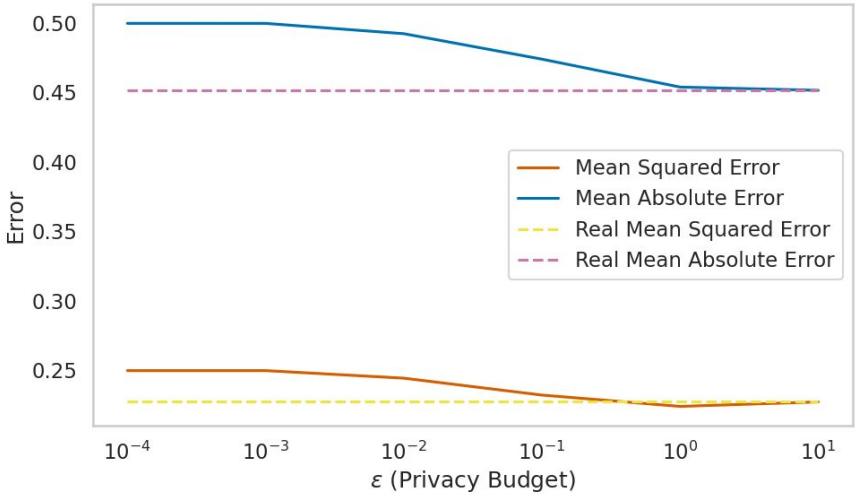
where:

$$\nu \sim Lap\left(0, \frac{1}{\varepsilon}\right)$$

Linear Regression

Varying the privacy budget ε, we performed Linear Regression on the data to predict the Status of a patient.

Our aim was to understand the differences between Mean Squared Error and Mean Absolute Error of the privacy scenarios and the real scenario.

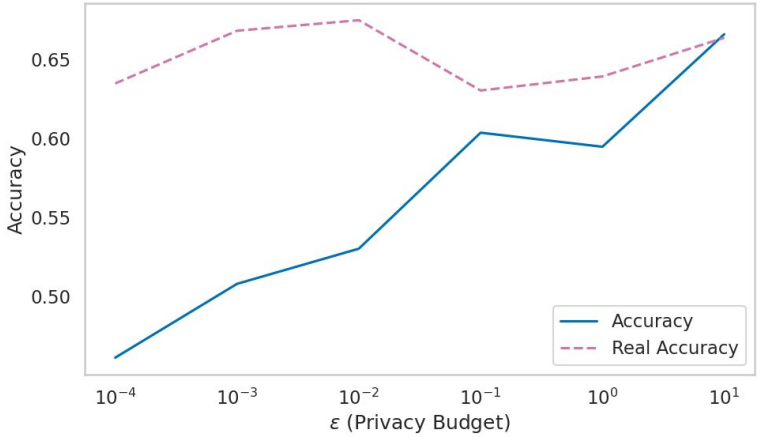


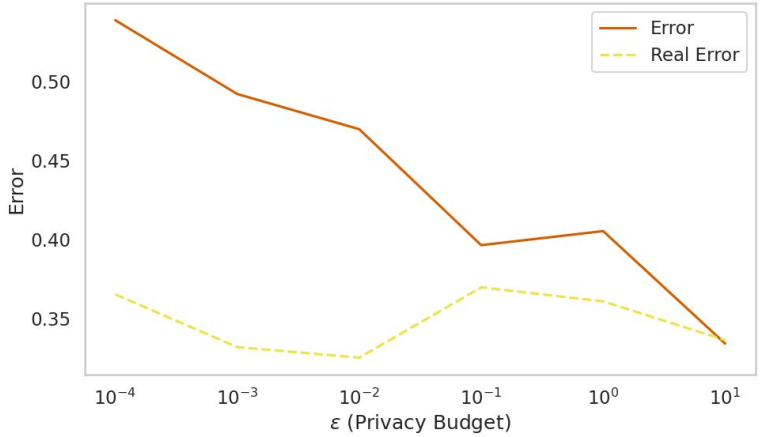
Bernard et al. Molecular International Prognostic Scoring System for Myelodysplastic Syndromes. NEJM Evid 2022;1(7). https://evidence.nejm.org/doi/full/10.1056/EVIDoa2200008

Classification using ε-DP Neural Networks

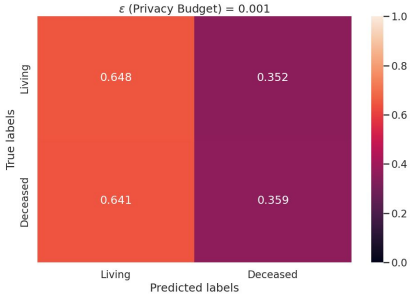
Network Configuration	
Loss Function	Cross Entropy
Optimizer	Adam
Hidden Layers	1

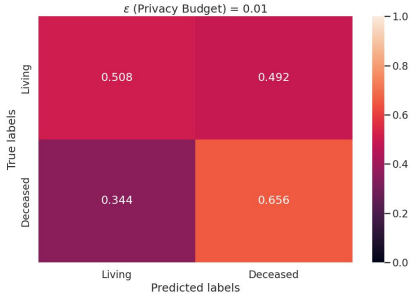
Network Configuration	
Training Epochs	100
Learning Rate	0.001
Hidden Neurons	150

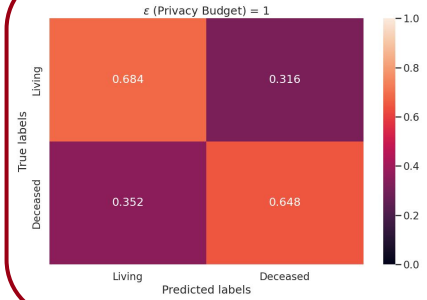


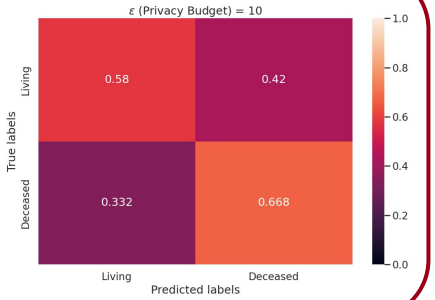


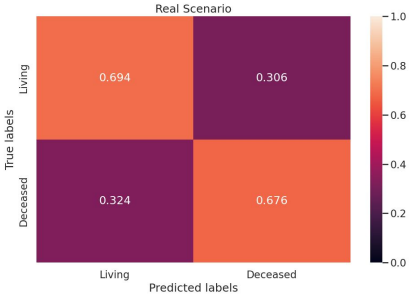
Confusion Matrices







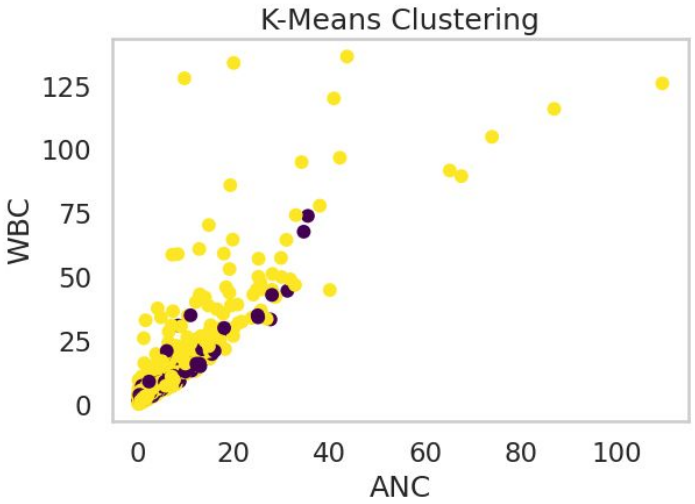


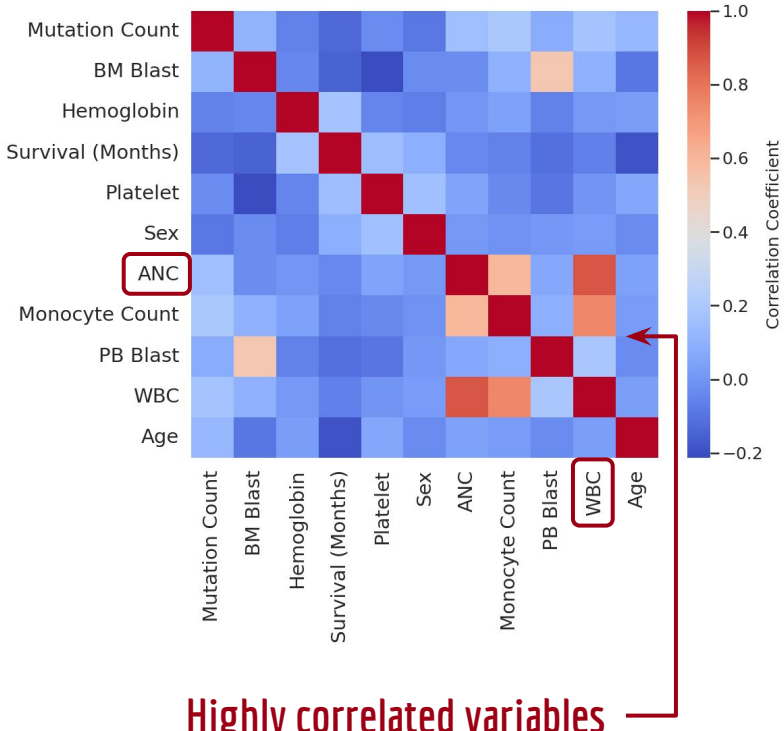


Similar pattern compared to the **Real Scenario**.
(Confirms the results of the statistical analysis)

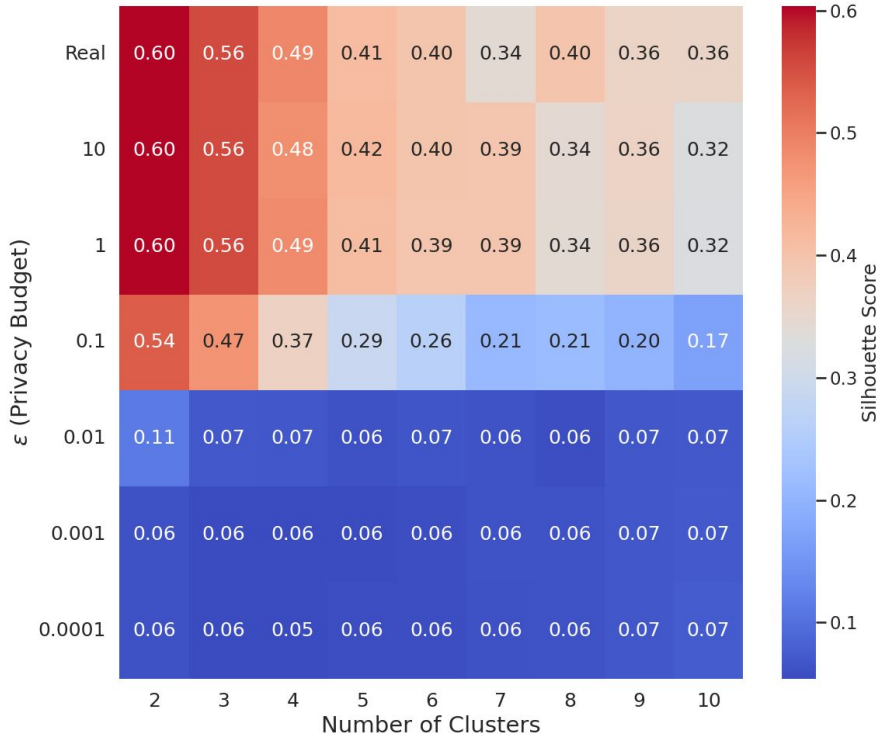
Clustering

The variables selected to visualize the clustering were chosen based on the **correlation coefficient** computed on the original values.





Clustering vs. Privacy Budget ε



DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

Differential Privacy in Medical Analysis

Thank for the attention
Question Time!

Francesco L. De Faveri
Laura Menotti
Department of Information Engineering, UniPD