

Differential Privacy

Differential Privacy: Sharing Data Without Sharing Secrets

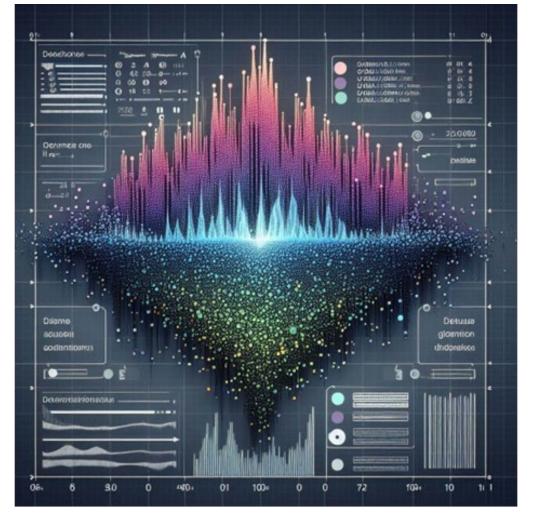
Differential Privacy

- Ensures data privacy by adding carefully calibrated noise to query responses
- Prevents the identification of individual records (anonymity)
- Still allows for accurate aggregate analysis

Goal: Enable non-technical users to use Differential Privacy with very little required technical expertise



DP suggested by Recommendation Engine with ε-value



Users will have to locally add noise to data using the **ENCYPT** interface



This noisy data will then be uploaded to the **ENCRYPT** platform



Machine Learning models can then be trained on the anonymized data

Comparison of model training with addition of noise (Differential Privacy case) and the original clean dataset showed that despite noise, data analytics is still possible.



Most accurate models are serialized (pkl format) for future use

- These models can be used for analytics on non-training datasets Code is general enough to work on any dataset
- Just need to identify the features to use and target variable
- The above is possible through the user interface





























