Status Report: Subsalt - MIDS

12.09.2024

Overview

Expected delivery

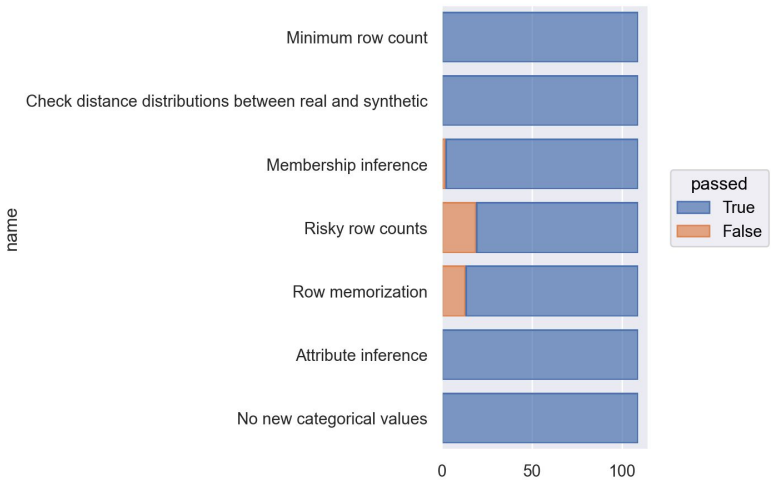
December 9, 2024

Recent progress

- 109 data entries generated
- Baseline models

Privacy Test Analysis

Overview

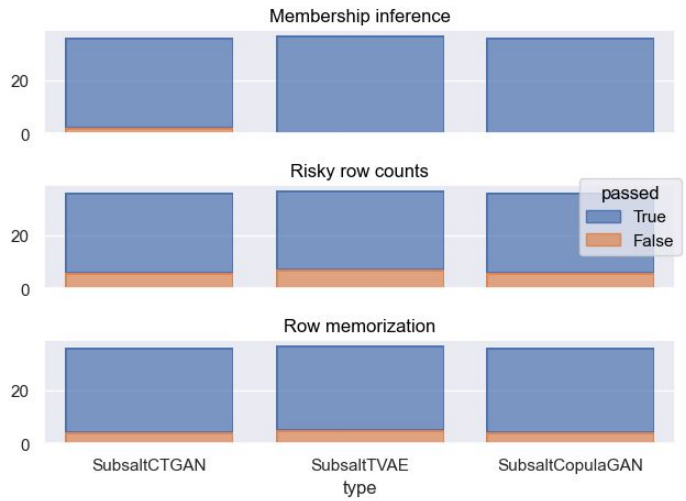


- 109 data entries
- 7 privacy tests
- 3 privacy test pass/fail
  - Membership inference 107/2
  - Risky row counts 90/19
  - Row memorization 96/13

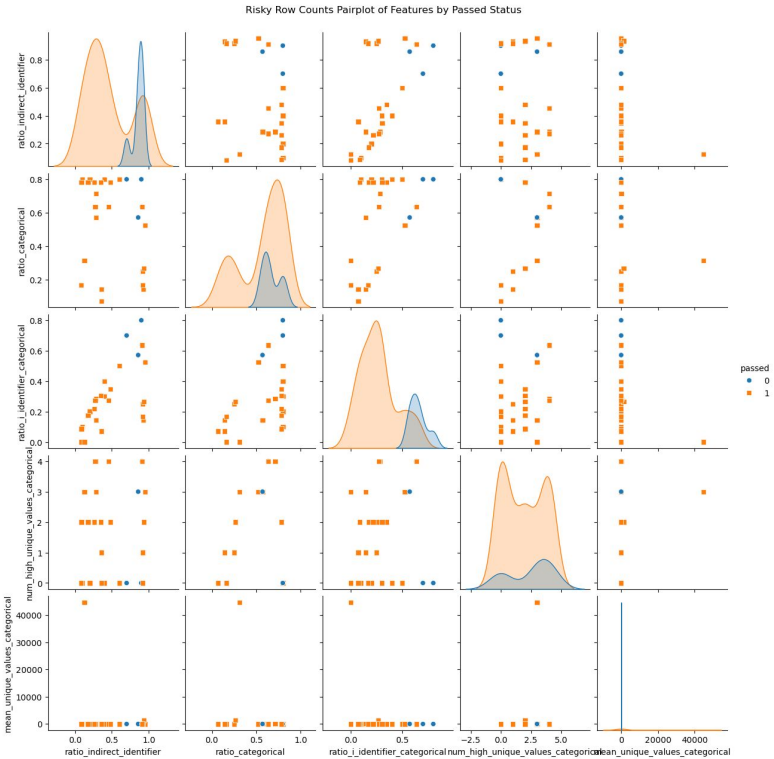
Privacy Test Analysis

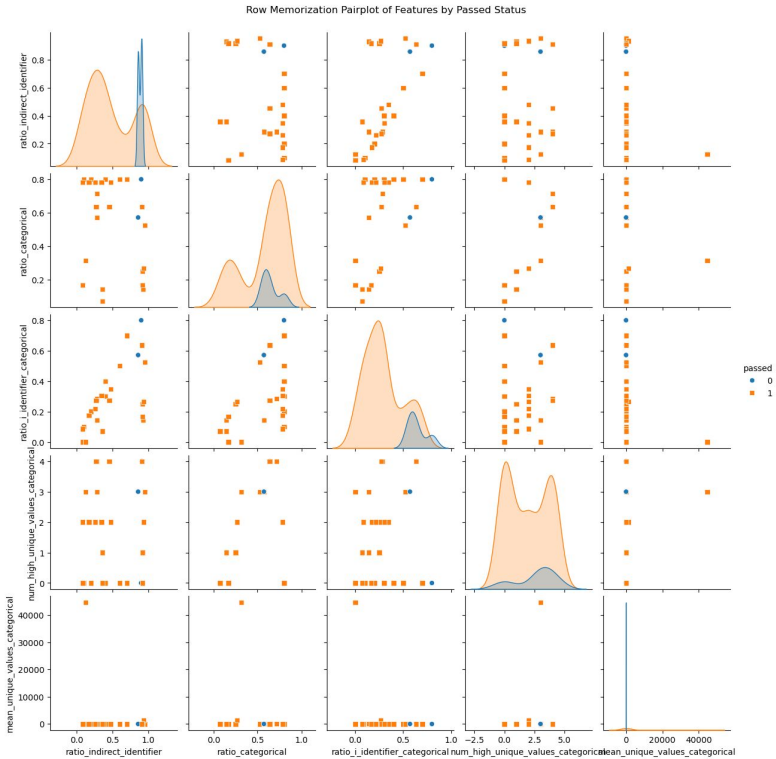
- CTGAN sensitive to membership inference
- TVAE higher chance to fail risky row counts and row memorization
- All models can fail risky row counts and row memorization

Breakdown

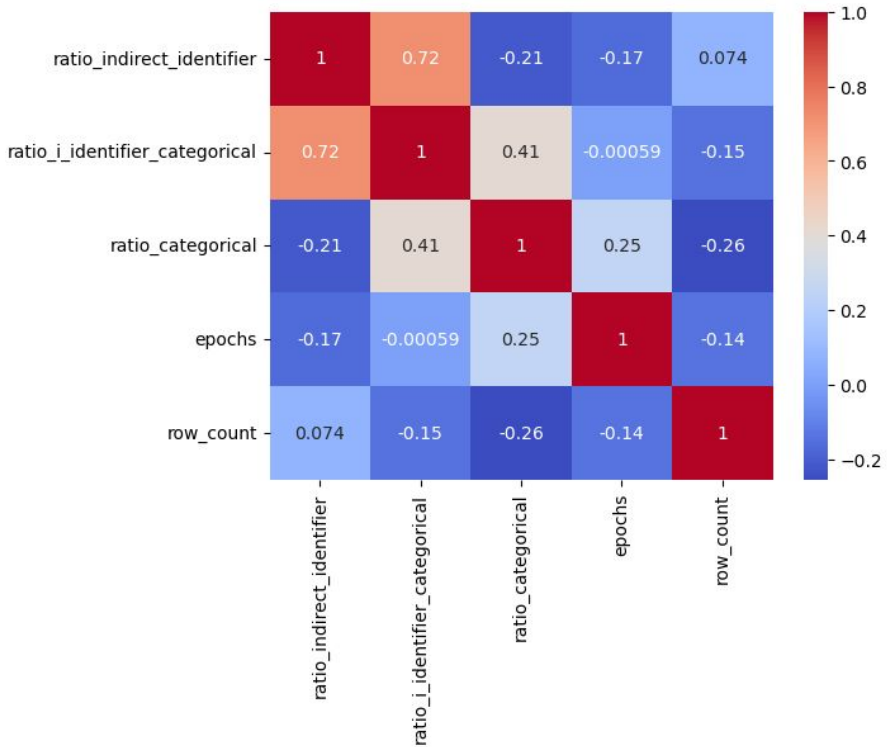


Feature Analysis





Correlation Matrix



LR Models - Risky row counts

Model 1

- Features: "ratio_indirect_identifier","ratio_categorical","epochs", "row_count","type"
- Pass/Fail: 78/16

Model 2

- Features:

  "ratio_i_identifier_categorical", "epochs", "row_count","type"
- Pass/Fail: 78/16

LR Models - Risky row counts

Model 1

Mean PR-AUC: 0.84 +/- 0.11

Mean Accuracy: 0.83 +/- 0.05

Mean F1 Score: 0.91 +/- 0.03

Model 2

Mean PR-AUC: 0.84 +/- 0.16

Mean Accuracy: 0.83 +/- 0.05

Mean F1 Score: 0.91 +/- 0.03

LR Models - Row memorization

Model 1

- Features: "ratio_indirect_identifier","ratio_categorical","epochs", "row_count","type"
- Pass/Fail: 84/10

Model 2

- Features:

  "ratio_i_identifier_categorical", "epochs", "row_count","type"
- Pass/Fail: 84/10

LR Models - Row memorization

Model 1

Mean PR-AUC: 0.89 +/- 0.11

Mean Accuracy: 0.89 +/- 0.01

Mean F1 Score: 0.94 +/- 0.00

Model 2

Mean PR-AUC: 0.90 +/- 0.05

Mean Accuracy: 0.89 +/- 0.01

Mean F1 Score: 0.94 +/- 0.00

Next steps

More privacy test fail

- Membership inference fail too less to build model
- Robust models

Different type of models

- Tree-based, Neural Networks, etc.

Feature engineering

- Large dataset (high row count)

Imbalance handling

- Improve performance of the minority class (fail)