PWSCUP2025 Team 08

静岡大学　大木研究室
牧野由　金杰　徳増真大　濱本柊弥

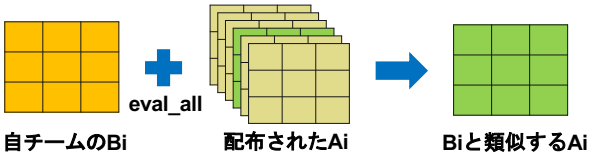
本戦匿名化フェーズ

方針

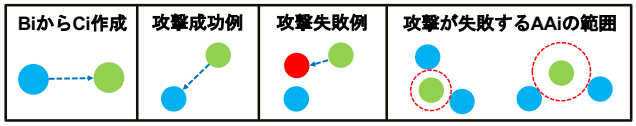
❑Syntheaデータを採用して高い有用性と匿名性を保持
❑行シャッフルを行うことでLRスコアを最大化
❑Biに対して距離が遠いデータを採用し，攻撃率を低下

匿名化データの作成

Step1：予備選終了時に配布されたAiからBiに類似するものを選定

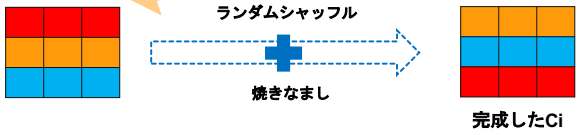


Step2：採用したAiからBiに対して距離が近い1万行を除外



Step3：Aiを基にstats+KWスコアが高くなるようCiを作成

Step4：行シャッフルによりLRスコアを最大化



機械学習モデルの作成

アイデア　Biを用いなければモデルベースの攻撃は難しい
方法　独自にSyntheaで合成データを作成して学習

- 学習/検証：Syntheaで生成した200,000件の合成データ / Bi
- Stroke_Flagに人数の偏りがあるため，F1 Scoreで最適化

総合有用性スコアとして91.18を獲得！

本戦攻撃フェーズ

攻撃手法：データ拡張

アイデア　基本統計による評価においてはピアソン相関を用いているため，共分散の値が近いほど有用性に関する貢献度が類似する

方法　各数値列の共分散が近くなるように，データを拡張する
拡張した各データ行から距離が近いデータ行を探索して10,000行を推定する

AGE	encounter_count	…	mean_weight
5	2	…	10.0
3	6	…	8.0
10	1	…	8.0
…	…	…	…
2	1	…	2.0

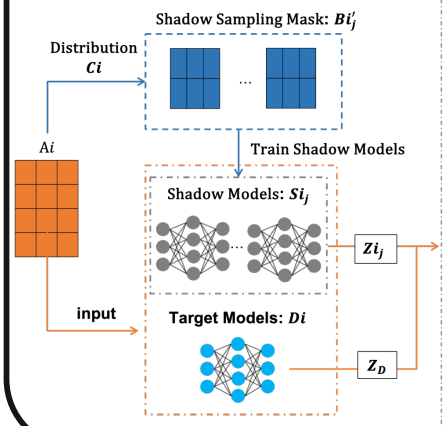
元データ（AGE列 + 数値列）

AGE_encounter_count	AGE_num_procedures	…	mean_bmi_Mean_weight
-20	30	…	20.2
15	20	…	-40.8
-5	-10	…	25.4
…	…	…	…
-10	15	…	-30.0

各数値列間の共分散

攻撃手法：Gaussian LiRA[1]

[1]: N. Carlini et.al. "Membership Inference Attacks From First Principles," IEEE S&P 2022



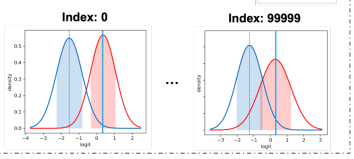
Logit 定義
Zi_j: 各シャドウモデルのロジット
Z_D :ターゲットモデルのロジット

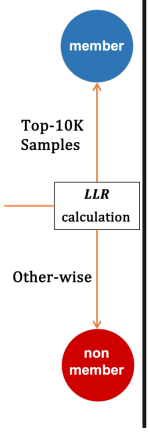
IN / OUT定義
IN:Bi_j'を学習に含むシャドウモデル群
OUT:Bi_j'を学習に含まないシャドウモデル群

Gaussian近似パラメータ
μ_{IN}, μ_{OUT}: IN/OUT 群それぞれのロジット値 $\overline{Zi_j}$
σ_{IN}, σ_{OUT} : IN/OUT 群のロジット値の標準偏差

IN / OUT 分布の対数尤度比 $LLRi$
入力データの「学習データらしさ」を示す指標

$$LLR = log\frac{\sigma_{OUT}}{\sigma_{IN}}\frac{(Z_D-\mu_{OUT})^2}{2\sigma_{OUT}^2}\frac{(Z_D-\mu_{IN})^2}{2\sigma_{IN}^2}$$





距離ベースの攻撃結果

Team 14に対して
Top-1スコアを達成

Team 14	5536

Team 9, 12, 17, 23, 24に
対して Top-3スコアを達成

Team 9	1247
Team 12	1069
Team 17	4522
Team 23	1123
Team 24	7379

モデルベースの攻撃結果

Team 1, 10, 11, 19, 20に
対してTop-1スコアを達成

Team 1	1653
Team 10	1379
Team 11	1668
Team 19	1289
Team 20	1343

全体結果

他手法と合わせて
攻撃総合スコアで**1位**を達成