PWSCUP2021 Team06 Data Protection Club

（セコム株式会社IS研究所 玉井睦 松永昌浩 香川椋平 髙山献 上井良介 川田涼平 島田陽介） 2021/10/27

匿名化手法

方針
- 可能な限りレコード削除しない。
- ノイズを入れずに再識別率(topk) 0を目指す。
- 所属推定(recall, prec)対策として囮レコードを作る。

■レコード削除

・特異レコードの削除
1. Bのカテゴリ変数列をダミー変数化、連続値列を0～1の値に変換
2. 列ごとの平均値を各列の標準値とし、各行の値との差を計算し、列の和をその行の特異性とする
3. Bのユニーク行Buに含まれる特異性の高い上位50レコードを削除しBdelを作成

・ランダム削除
1. Buから特異性の高い50行を除いたデータセットをクラスタリング
2. Bdelからuniqrt(C)<=0.5を満たす削除行数を決定
3. クラスタごとに同じ割合で削除行数を割り当て、ランダムサンプリングでレコード削除し、Ctmp, Xtmpを作成

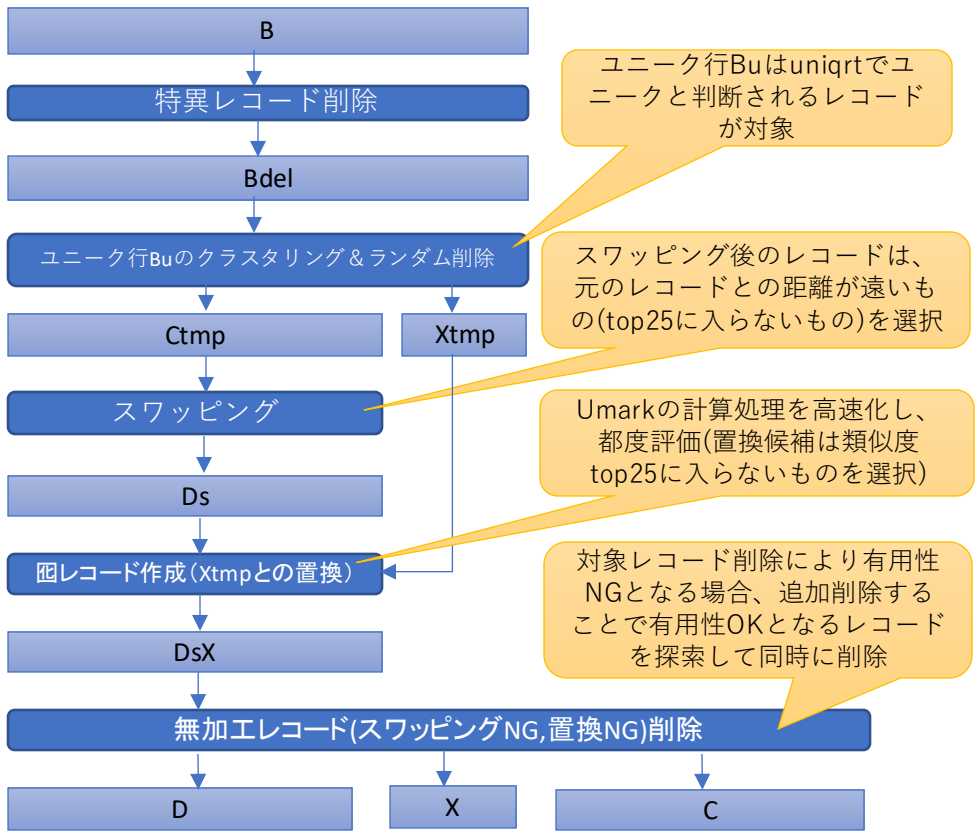
■スワッピング

・Ilossを満たす範囲でCtmp内でのスワッピングを行いDsを作成
・このとき、スワップした結果が元のレコードの類似度top25に入らないレコードを交換先に指定

■囮レコード作成

・Dsの各レコードと削除したレコードXtmpを比較し、ilossとumarkを満たす範囲で、Dsの各レコードをXtmpと置換し、DsXを作成
・このとき、置換した結果が元のレコードの類似度top25に入らないレコードを選択

■匿名化の処理フロー



攻撃手法

方針
- 原則として削除レコードの推定は行わない
- 囮レコード（距離の近い他レコードとの置換や交換）を前提とし、レコード間の距離が近いものは、あえて候補から外す。

■削除レコード推定

・原則、削除レコード推定は行わない
⇒recall=1.0, prec=0.5 が確定
・例外として、削除レコード推定を行う場合
①iloss <=4 となるレコードが存在しない場合
⇒有用性を満たしていないため削除確定
②iloss <0.5 となるレコードが複数見つかった場合
⇒距離が近すぎるレコードは囮とみなし削除判定

■再識別

・ 3<= iloss <=4 を満たすレコードからランダムサンプリング

CT	例：条件に該当するDのレコード数			削除判定
#	iloss<=4	3<=iloss<=4	iloss<0.5	
0	230	162	0	
1	84	81	1	
2	**0**	0	0	①に該当、削除
3	25	19	1	
4	30	24	0	
5	235	169	0	
…	…	…	…	
99	165	119	**5**	②に該当、削除

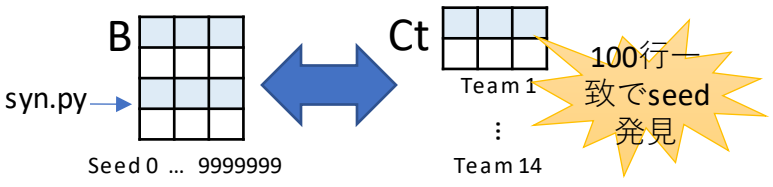
その他採用しなかった/できなかった攻撃手法

■Seed推定攻撃

・着想
・予備選ではBを活用した攻撃が強かった
・ 本戦は各チーム与えられたseedによって生成
・我々のチームに与えられたseedは1256868
・各チームのseedが近ければBが見つかるのでは？

・攻撃方法
・Syn.pyにて0～9999999でBを作成
・CTの100行と完全一致するBを探索

・攻撃結果
・7桁空間を探索したが見つからず



■オッズ比情報利用攻撃

・着想
・匿名加工フェースではオッズ比の条件が厳しい
・オッズ比はBとDのオッズ比の差の最大値
・誤差が大きくなりやすいラベルは加工が難しい

・攻撃方法
・自チーム対象の調査結果ではeduとmarが上位
・0.08以上の6ラベルすべてがeduかmar属性
・ランダムサンプリング攻撃時に両属性が異なるとilossにペナルティ

・攻撃結果
・ペナルティを与えても攻撃結果は改善されず

ペナルティ	0.5	0.25	0(baseline)	-0.25
攻撃成功率	0.038	0.042	**0.051**	0.042

自チーム本戦データの各オッズ比誤差

Label	OR
edu[T.HighSchool]	**0.094440**
mar[T.Married]	0.087104
edu[T.Graduate]	0.086284
mar[T.Widowed]	0.083402
mar[T.Parther]	0.082219
mar[T.Never]	0.080587
gen[T.Male]	0.068148