

匿名化フェーズ：方針と初期データ選定

BB11と全く別のデータをCC11の初期データとして選び、スワッピング・微小ノイズで徐々に有用性を高める戦略

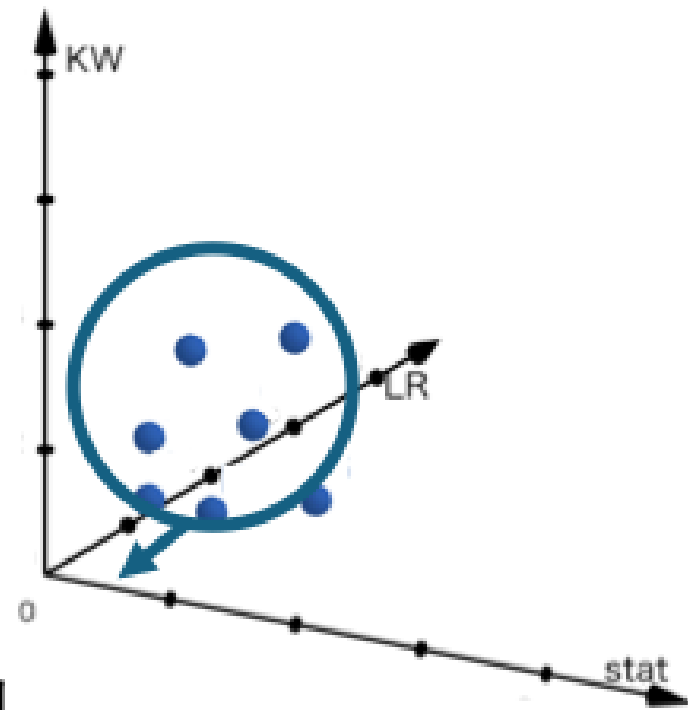
- 初期データ選定①：合成データ生成ツール（AIM, DP系モデル）
- 初期データ選定②：予備戦B11_1~3と敵チームC01~C24
→これらの中からBB11_1~3に対して有用性が高いものを選ぶ（eval_all.pyで評価）

結果 ①は有用性・攻撃耐性が低く却下。
②BB11_3に対してB11_3が最も有用性が高く、初期データとして採用

スワッピング手法

relation_swap_calibrate.pyの機能と特徴

- データの列（属性）どうしを入れ替え、有用性スコアを少しずつ改善するツール
- データの分布形状をほぼ維持したまま、相関関係やモデル整合性のみを調整



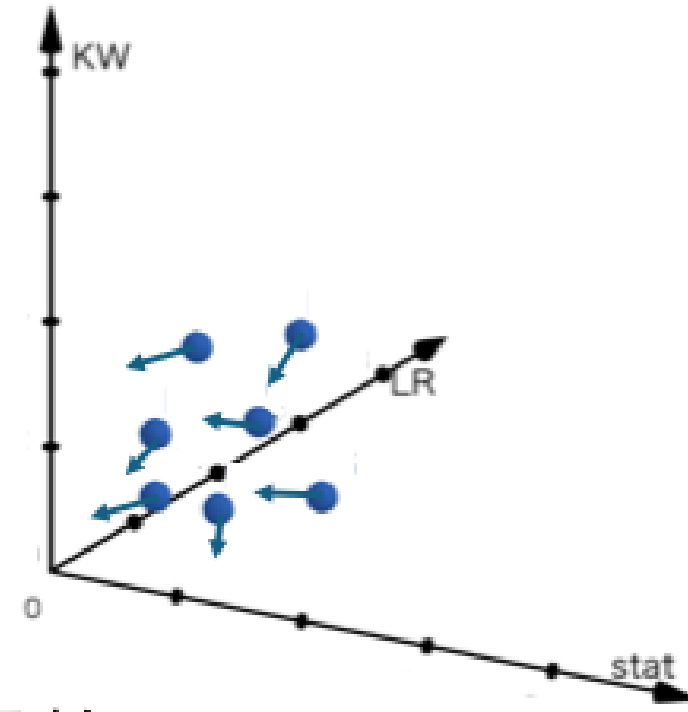
特徴：再学習不要で軽量・高速、分布破壊を抑制

効果：LR_diffの改善に特に有効

微小ノイズ手法

noise_search_calibrate.pyの機能と特徴

- データに微小なノイズを入れて評価し、スコアが改善すれば採用する山登り探索型のツール
- 値そのものをわずかに動かし、分布の形を保ちながら統計値を微調整



効果：KW_diff, stats_diffの改善に有効

※複数パラメータをChatGPTとの対話で慎重に調整

最終結果

有用性スコア

総合有用性スコア：84.54点

- utility1：37.66点
- utility2：9.77点
- utility3：19.19点
- utility4：17.91点

攻撃フェーズ：方針と手法一覧

複数の攻撃手法を実装・評価し、チームごとに最も攻撃成功数の高い手法を採用

- 提出1回目-mia2.py：距離ベース攻撃。類似度上位1万件を1として出力
- 提出2回目-mia5.py：距離＋密度(GMM)＋異常度(IF)＋自己学習で複合評価
- 提出3回目-xgbt_pred.py：学習済みモデルで予測→一致行抽出→サンプリング
- 提出4～7回目-ai_subset_search.py：統計的関係の逆追跡。ビーム探索で効率化

各手法の詳細

(提出1, 2回目はmia.pyを改良したものなので省略)

- 提出3回目：xgbt_pred.py
 - 学習済みモデルDDxx.jsonと、AAxx.csvを入力
 - 各行についてモデルの予測値を計算
 - 予測結果がAAxx.csvの実測値と一致する行のみを抽出
 - 抽出結果をAAxx.csvとして保存
 - AAxx.csvからランダムに1万件を選び攻撃に使用

各手法の詳細

- 提出4回目：ai_subset_search.py
 - AAxx.csvからランダムに1万行選ぶ
 - 近傍生成（k-swapで行入替）
 - 評価（stats, KW, LRでスコア算出）
 - スコア改善→採用
悪化→棄却
 - 上位候補のみ保持（ビーム探索）

繰り返し、最良のものを出力

最終結果

各チームに対して最も効果的だった攻撃手法

	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
mia2.py																								
mia5.py																								
AAxxランダム抽出																								
ai_subset_search.py																								

最も効果的だった手法

最終攻撃スコア：27198点

匿名化データを利用した分析手法の提案

自衛隊隊員の過負荷・オーバートレーニングの早期警戒

- 目的：部隊単位の身体/心理的ストレス傾向の可視化、対処
身体データと医療利用傾向隊員のメンタル不調の早期兆候を匿名的に検知し、カウンセリングや勤務配慮の判断材料に活用。
- 手法(特徴量作成)：心理/身体負荷の傾向および異常の検出
血圧、BMI変化率、投薬量、受診回数、受診変動率などから統合スコアを生成。「メンタル/身体不調群」の傾向を学習。
- 効果：国防力の基盤強化、他機関への応用
 - 人的リソースの持続性向上、離脱低減、部隊の即応性/稼働率維持
 - 医療、人事、指揮の連携強化→部隊の効率的マネジメント実現
 - 自治体、官公庁(消防・警察等)、災害対応などへの応用