Team08：Asakatsu2025

静岡大学大木研究室

牧野由*, 金杰, 徳増真大, 濱本柊弥

PWSCUP2025　2025年10月29日

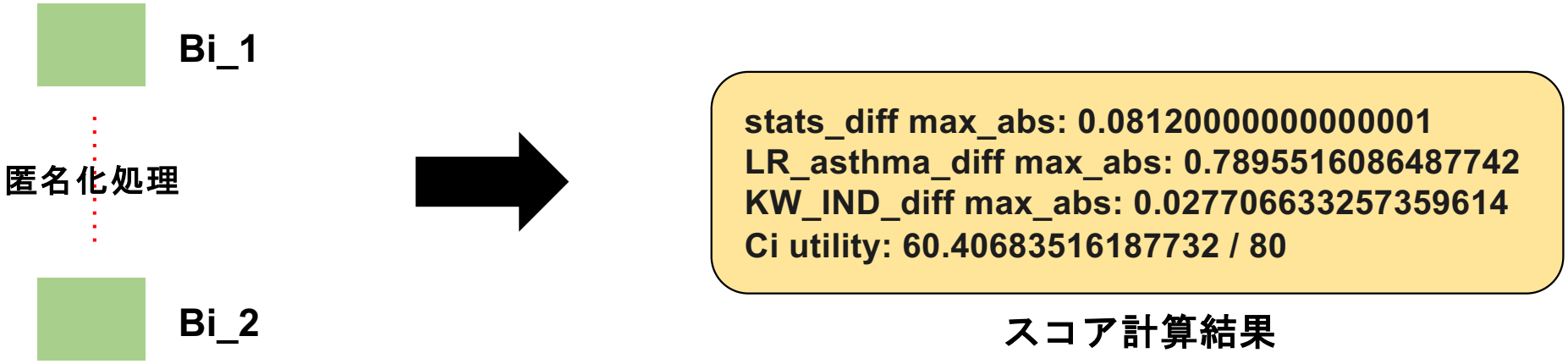
本戦

コンペの本質に気づき始めた瞬間

本戦匿名化

匿名化データ

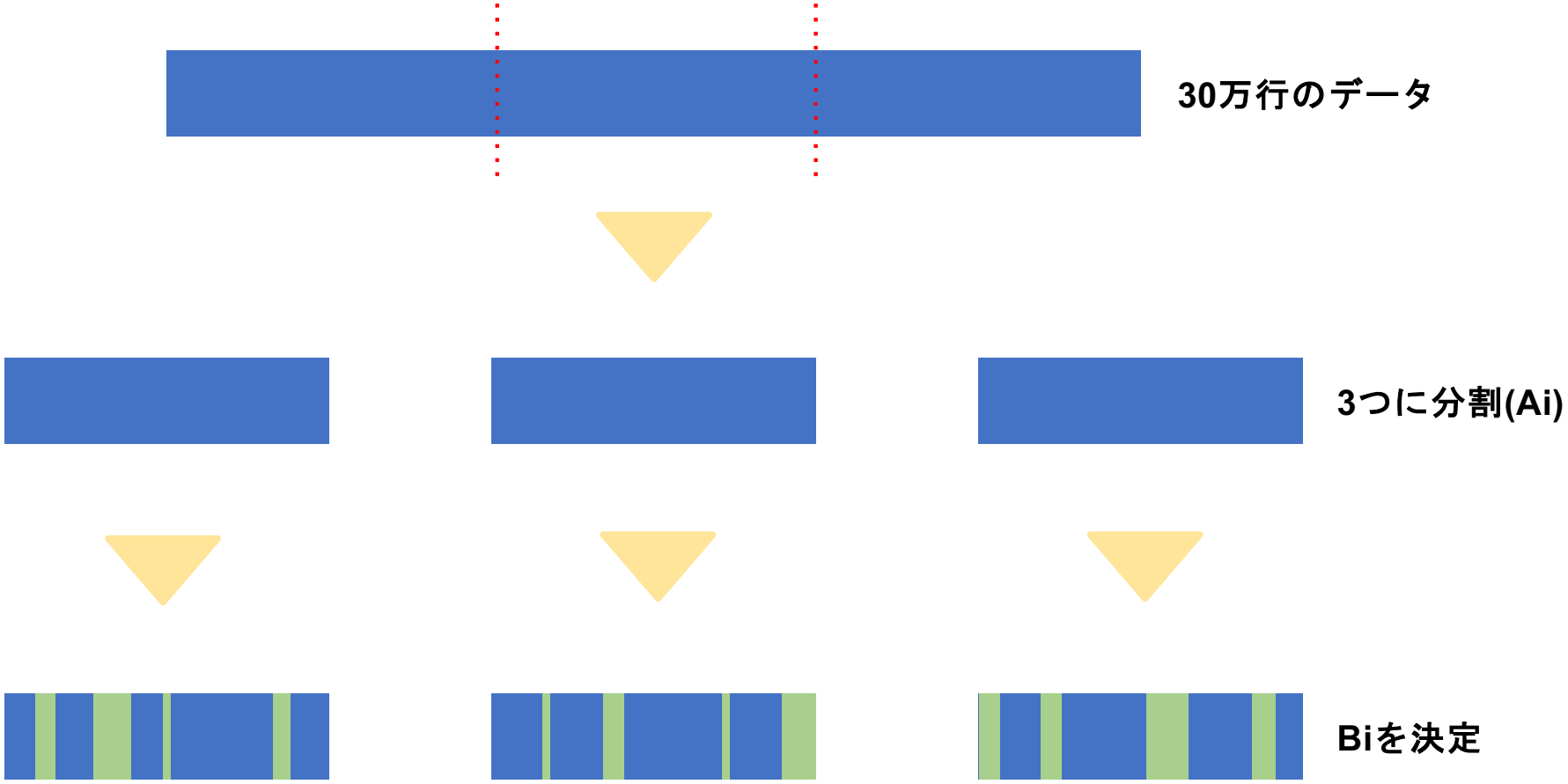
Biの匿名化データ（Ci）として別のBiを利用してみる



攻撃してみても1000件程度しか当たらない

…なんか知らんけど意外と有用性が高くて匿名性が高いデータが作成できた

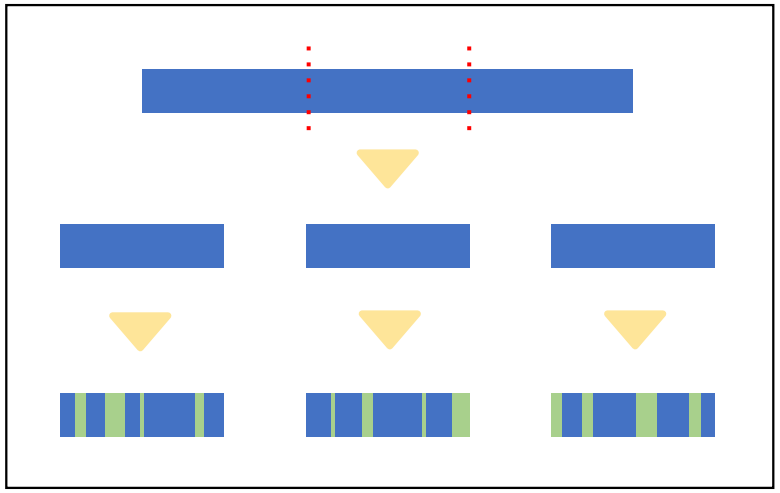
配布データ生成に関する考察



配布データ生成に関する考察

配布されたそれぞれのBiは一つのデータから抽出されたものと考えることができる

配布データを匿名データとして利用した場合，正解データと不正解データに引き寄せられる確率はランダムになる



Syntheaで匿名化データを作成しちゃえば元データを推定できないのでは？

有用性と匿名性の理論値について

有用性と匿名性がともに高いCiは作成可能か？

Ai-Biのデータ
（9万行）

1万行を抽出

stats_diff max_abs: 0.05068094212271384 6
LR_asthma_diff max_abs: 0.2095757675006184 6
KW_IND_diff max_abs: 0.02362126705166423
Ci utility: 73.30882162404579 / 80

Ai-BiからCiを作成した結果

外れデータに対してハミング距離が0のデータ（距離ベースの攻撃ができない）
を利用してCiを作成した場合でも，高い有用性スコアが得られる

有用性と匿名性の理論値について

攻撃されない匿名化データは最強なのか？

攻撃箇所をランダムに選択した場合，10分の1で攻撃できる

攻撃を行い、攻撃スコアが０の場合は残りの9万件に対してランダムに攻撃することで単純なランダムより多く攻撃されてしまう

一方で…

ランダムな攻撃が行われない場合，**攻撃方法に規則性**が生まれるため，敵チームに攻撃されてしまう箇所が集中してしまう可能性がある

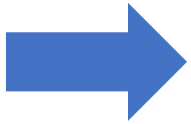
敵チームに**ランダムな攻撃を誘導できる程度の匿名性**を保つことで
より高い総合スコアが見込めると思われる

LRスコアについて

LRスコアでは学習と評価でデータが分割される
分割方法を変更するとスコアも変化

stats_diff max_abs: 0.0
LR_asthma_diff max_abs: 0.0
KW_IND_diff max_abs: 0.0
Ci utility: 80 / 80

元データ同士の比較



シャッフル

stats_diff max_abs: 6.5503158452884236e-15
LR_asthma_diff max_abs: 0.6954712266555703
KW_IND_diff max_abs: 0.0
Ci utility: 66.09057546688834 / 80

シャッフル後データとの比較

LRスコアはデータの配置を変更するだけで向上させることができる

本戦の匿名化手法

予備戦終了時に配布された全チームのAi（Syntheaデータ）を利用

▼

BBiとデータ分布が近いAiを採用，AiからBBiに対して距離が近い1万行を除外

▼

残りの9万行から有用性スコア（stats+KW）が高くなるように1万行を採用

▼

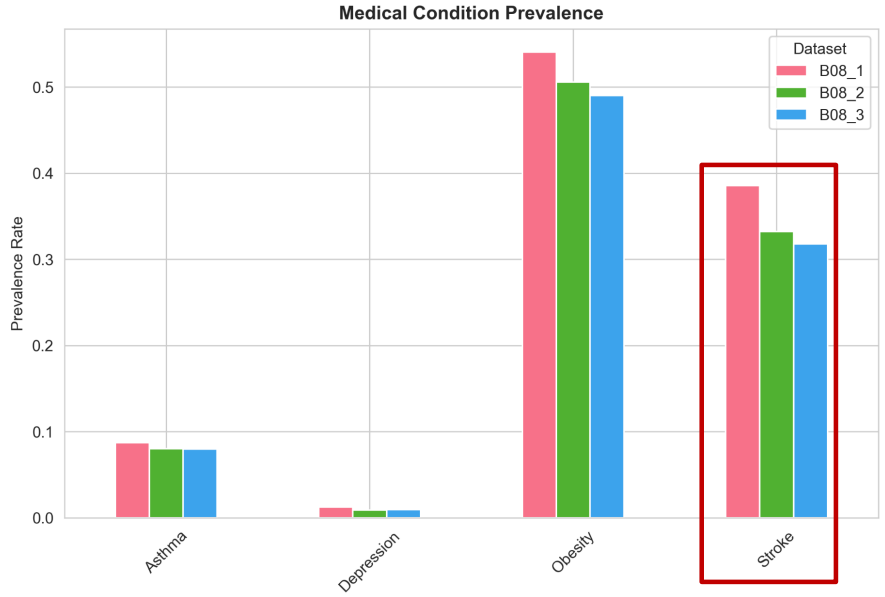
データをシャッフルし続けることでLRスコアを最大化

脳卒中リスク予測モデルの構築

Syntheaで独自に生成した200,000件の合成データで学習する

→ **Biを用いないため，モデルベースの攻撃が困難に**

Biを検証データとして最もF1 Scoreが高いモデルを採用



Stroke_flagが立つ人は全体の3割強

→ AccよりF1の方が性能指標として適する

Biによる評価

Validation Accuracy: **0.890**

Validation F1 Score: 0.848

Validation AUC: 0.954

本戦の匿名化指標においても

17.75点(Accuracy = 0.888を達成)

本戦

7手法のアンサンブル

本戦攻撃

攻撃手法

各チームの匿名化に合わせて異なる攻撃手法を採用

距離ベース

データ拡張による攻撃

チーム内ベストスコア
Team 2, 3, 5, 7, 9, 12, 14, 15, 16, 17, 18, 22, 23, 24

Team 14
に対してTop-1スコアを達成！

マハラノビス距離+ハンガリアン法

チーム内ベストスコア
Team 13

モデルベース

XGBoostの予測損失を用いた攻撃

チーム内ベストスコア
Team 4，6

Gaussian LiRA[1] の適用

チーム内ベストスコア
Team 1, 10 ,11, 19，20

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

距離ベース：データ拡張による攻撃

statsによる有用性評価ではピアソン相関を採用しているため，
共分散$(x-\bar{x})(y-\bar{y})$の値が近いほど有用性に対する貢献度が類似する

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

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

共分散に基づいて，数値列の組み合わせからデータ拡張を行う（AGEも数値として扱う）

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10000行を推定する

距離ベース：ハンガリアン法による最適割当

目的

最適割当

Aから抽出した1万件とCの全体の距離を最小化

前処理

数値列を0~1に正規化
カテゴリ列をOne-hotエンコーディングで0,1に変換

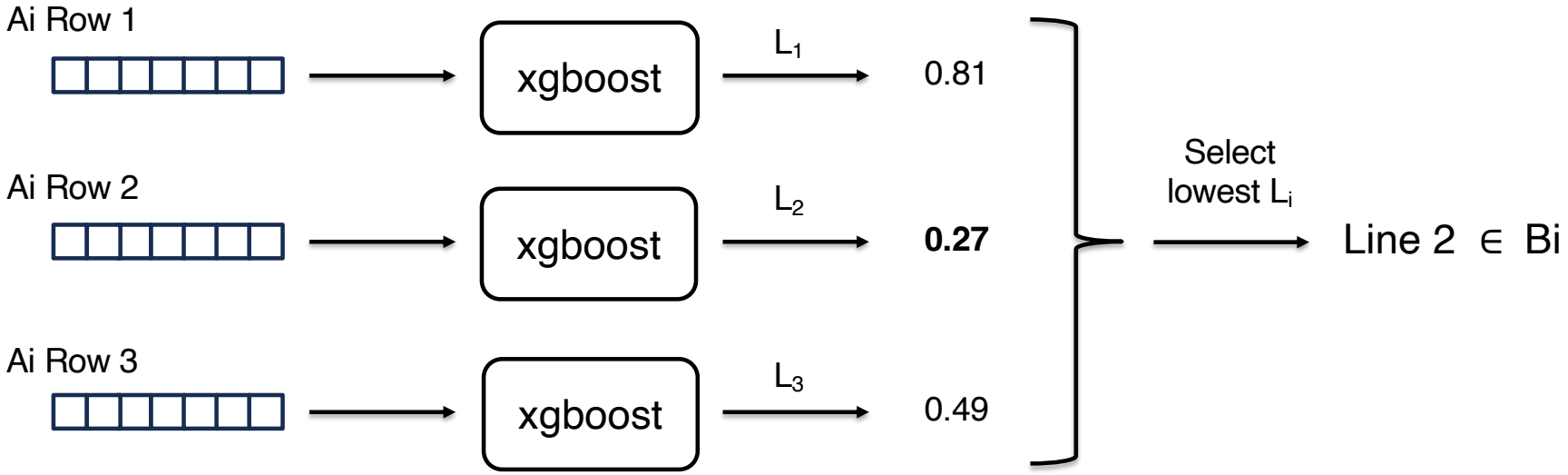
AとCの全ての組み合わせに対して距離$d(A_iC_j)$を計算

	C_1	C_2	…	C_{10000}
A_1	$d(A_1C_1)$	$d(A_1C_2)$	…	$d(A_1C_{10000})$
A_2	$d(A_2C_1)$	$d(A_2C_2)$	…	$d(A_2C_{10000})$
…	…	…	…	…
A_{100000}	$d(A_{100000}C_1)$	$d(A_{100000}C_2)$	…	$d(A_{100000}C_{10000})$

距離の合計が最小になるようにCの各レコードとAの各レコードを重複なしで割当

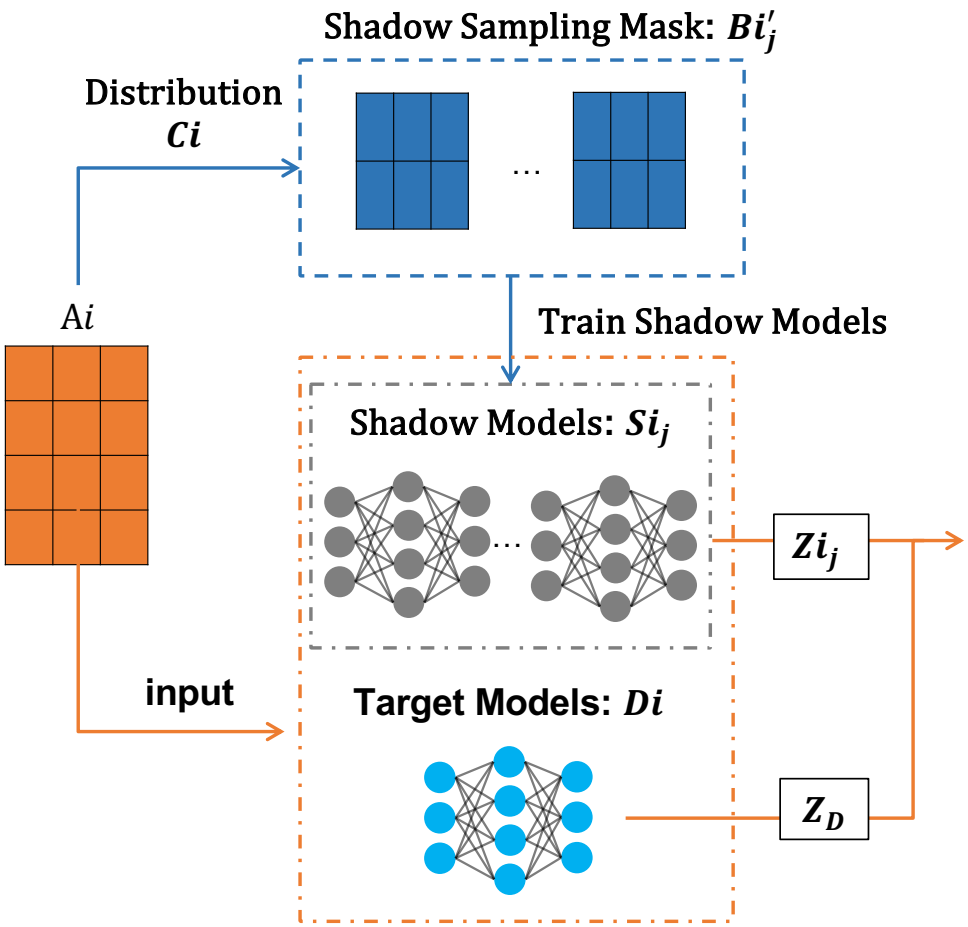
機械学習モデルベース：予測損失を用いた手法

XGBoostの学習にBiのレコードが使われていた場合，
他レコードよりも予測損失 L_i が小さくなる傾向がある可能性が高い



XGBoostモデルの予測損失が小さい**Aiの上位10,000件を選択**

XGBoost + Gaussian LiRA の理論構造



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

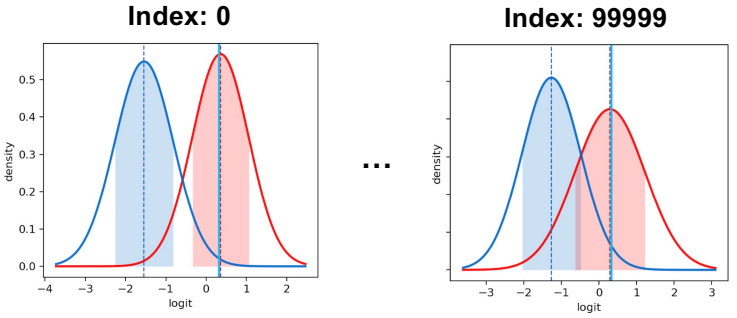
μ_{IN}, μ_{OUT}: IN/OUT 群それぞれのロジット値 $\overline{Zi_j}$

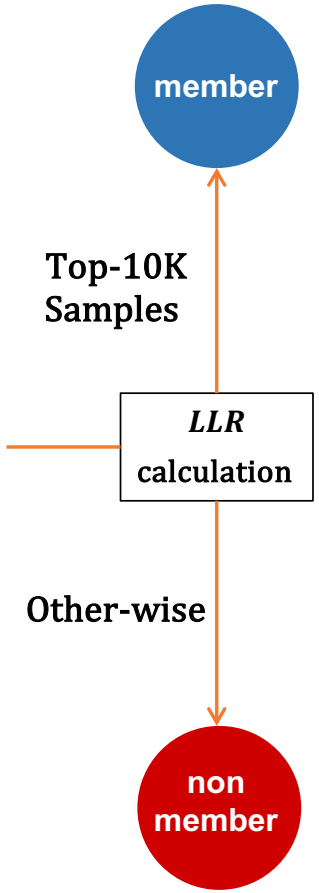
σ_{IN}, σ_{OUT} : IN/OUT 群のロジット値の標準偏差

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

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





攻撃リザルト

各チームに対する攻撃結果として

攻撃性能top1：6チーム
攻撃性能top5：17チーム

のスコアを獲得！

モデルベースによる攻撃が結果に大きく貢献し，
攻撃総合スコアで1位を達成！！！！！

チーム番号	攻撃スコア	攻撃方法	攻撃順位
チーム1	1,653	モデルベース	1位
チーム2	1,070	距離ベース	8位
チーム3	1,313	距離ベース	6位
チーム4	1,011	モデルベース	10位
チーム5	1,366	距離ベース	4位
チーム6	1,065	モデルベース	2位
チーム7	1,097	距離ベース	4位
チーム8			
チーム9	1,247	距離ベース	2位
チーム10	1,379	モデルベース	1位
チーム11	1,668	モデルベース	1位
チーム12	1,069	距離ベース	3位
チーム13	1,030	距離ベース	8位
チーム14	5,536	距離ベース	1位
チーム15	1,495	距離ベース	3位
チーム16	1,040	距離ベース	11位
チーム17	4,522	距離ベース	2位
チーム18	1,030	距離ベース	5位
チーム19	1,289	モデルベース	1位
チーム20	1,343	モデルベース	1位
チーム21			
チーム22	3,161	距離ベース	5位
チーム23	1,123	距離ベース	2位
チーム24	7,379	距離ベース	2位