

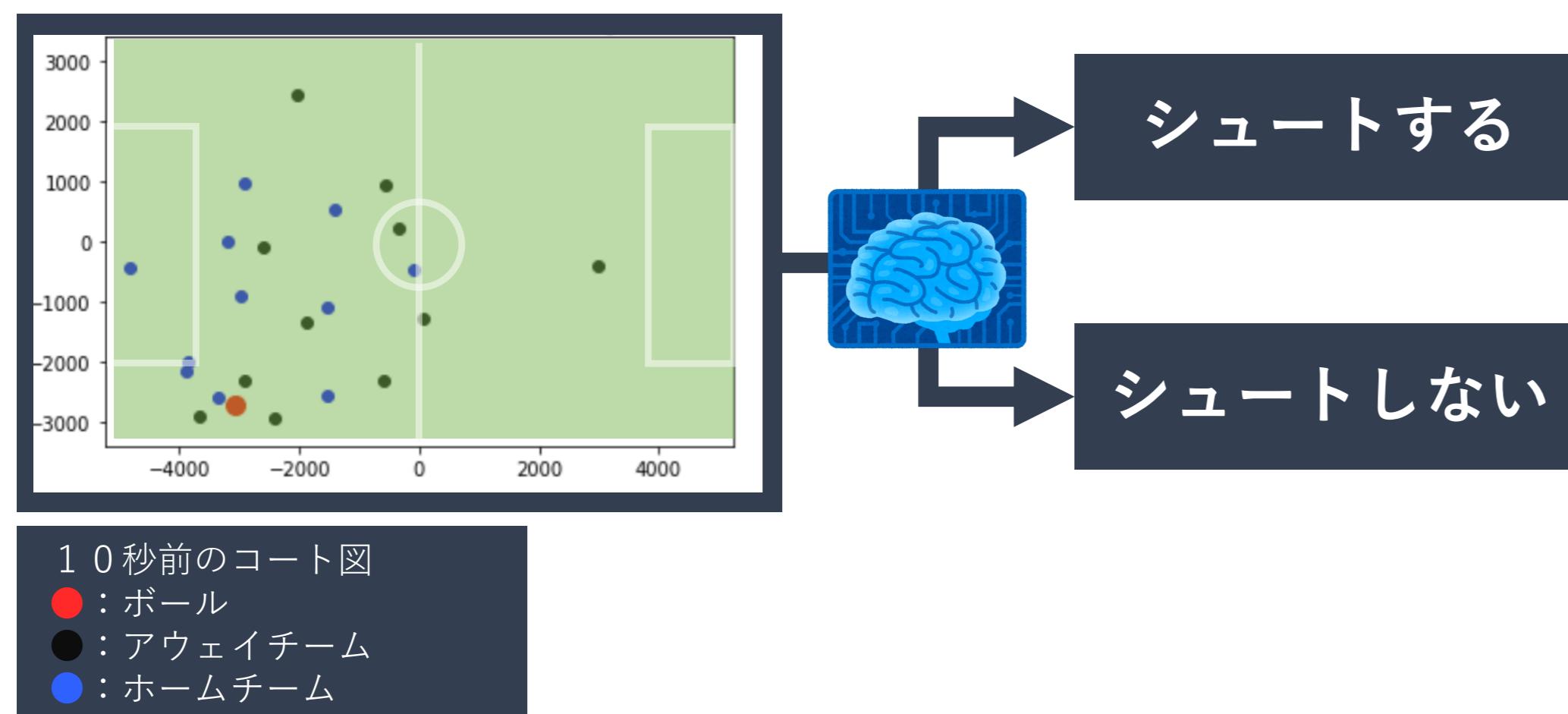
LightGBM を用いた各プレー場面におけるシュート予測とその要因分析

東瀬皓太郎 佐藤光太郎 澤田智秀 田中耀大 藤谷瑞生(東京理科大理工学部経営工学科) 安井清一(東京理科大学)

はじめに

- サッカーの試合は今のプレーがシュートにつながるかどうかが分からぬ！
- 野球などのターン制のスポーツと比べて白熱しているのか分かりづらい！

上記のことを踏まえて、本研究では、試合が白熱しているかどうかを刻一刻と移り変わる試合状況を機械学習で評価し、10秒以内にシュートが起こることを予測し可視化する。



説明変数と予測モデル

説明変数はトラッキングデータをもとに試合場面の状況を示す様々な特徴量として定義する。以下に作成したものを示す。

作成した説明変数

半径1,3,5,10m以内にいる総選手数。

半径1,3,5,10m以内にいるホーム・アウェイチームを区別した選手数

半径1,3,5,10m以内にいるホーム・アウェイチームを区別したFW, MF, DFごとの選手数

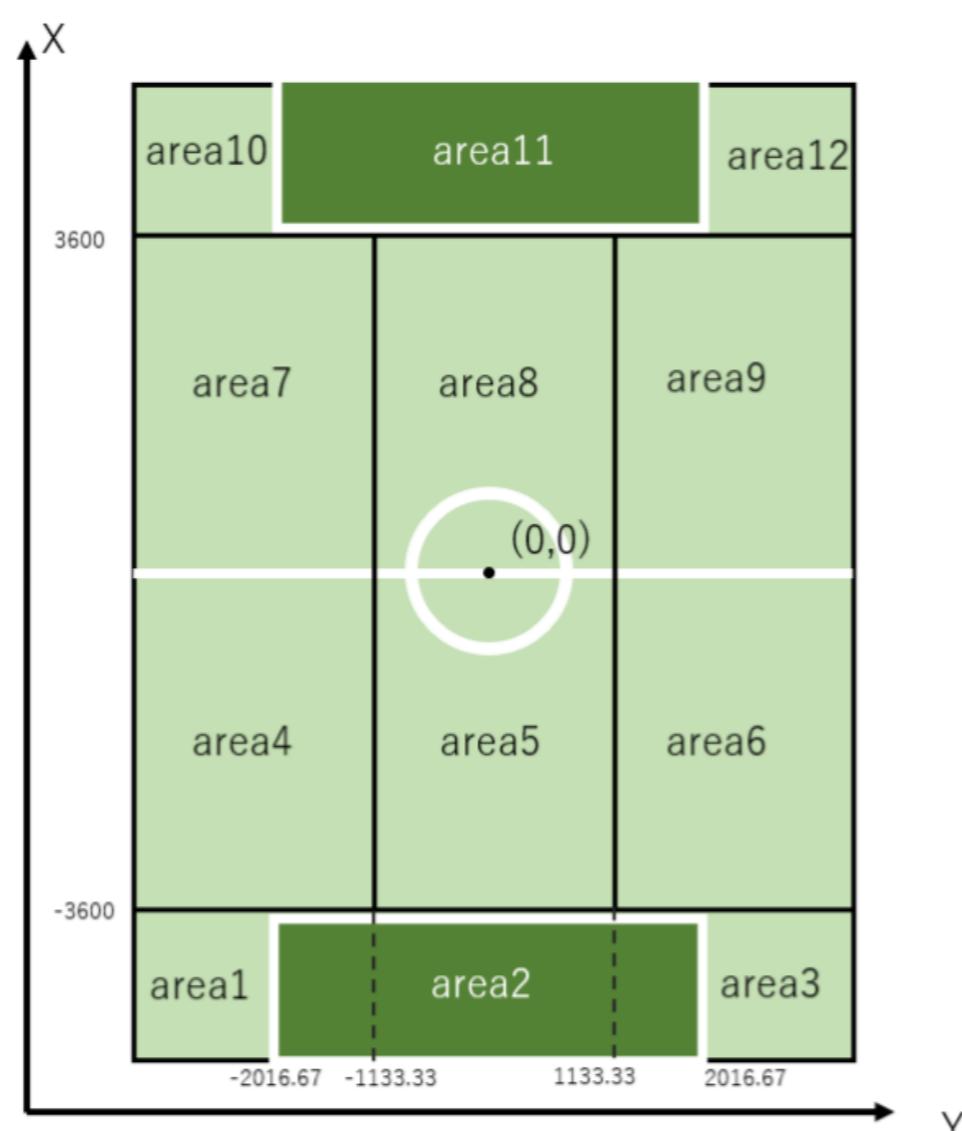
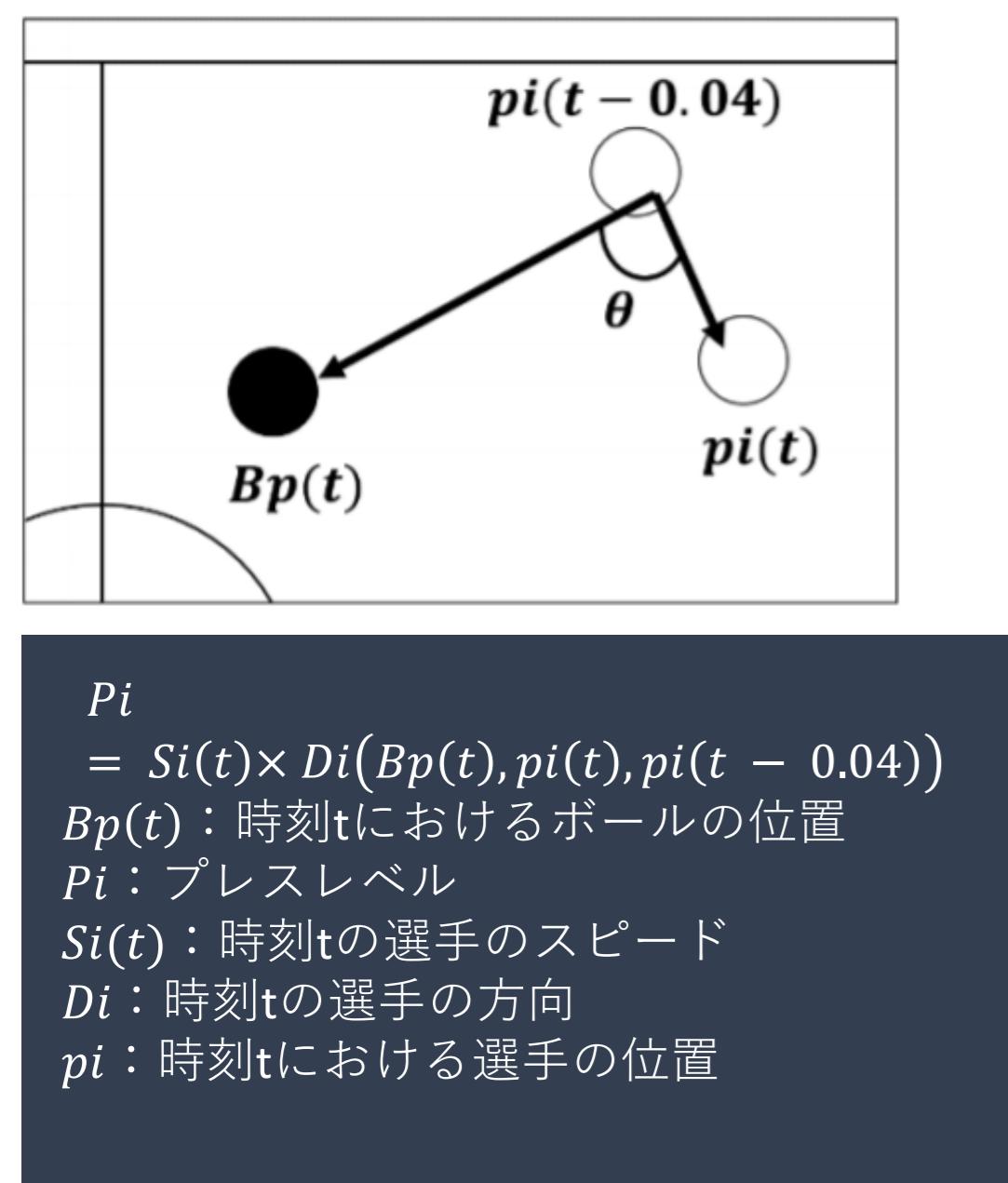
ホーム・アウェイチーム別に算出したFW, MF, DFの各ポジションにおけるx軸とy軸の長さ。

ホーム・アウェイチーム別のDFラインと、敵ゴールラインに最も近いFW選手のx座標(FWライン)

チームごとのプレスレベル[1]

下記のように12領域に分割し、各領域内に存在するチームごとの選手数

*プレスレベルは下記の式と図で定義した。



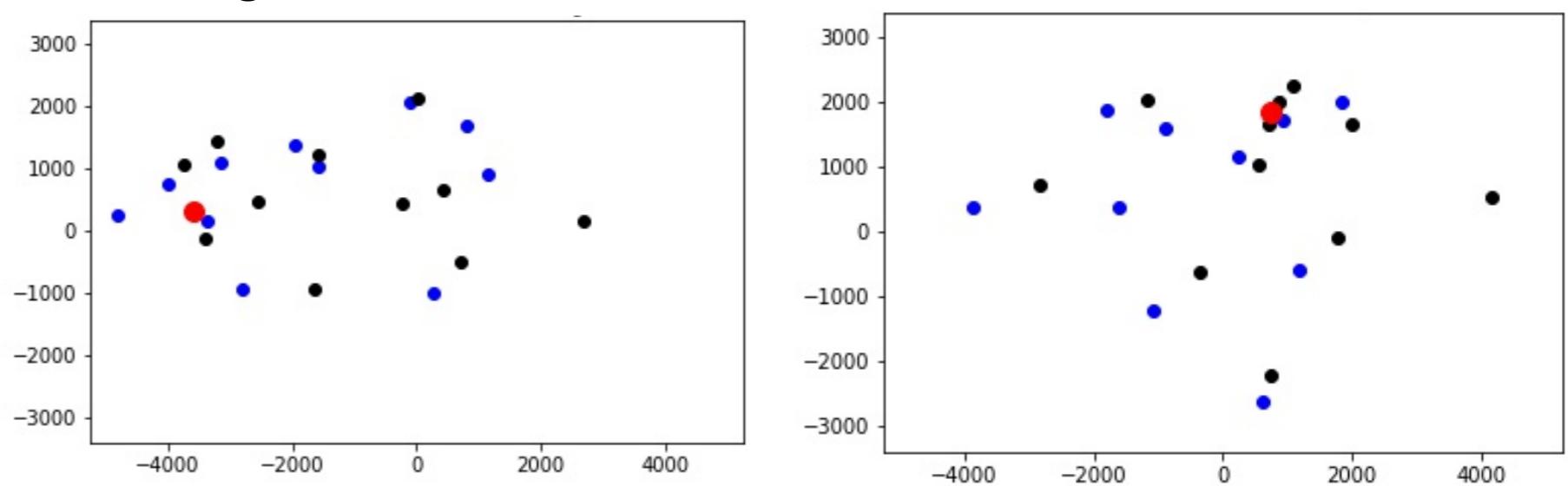
2値分類器としてMicrosoft社が提供する勾配ブースティングでの分類器であるLightGBMを用いた。学習の評価指標としてAUCを使用した。

結果

予測結果としてAccuracy 0.987(しきい値:0.5) AUCは、0.988となった。また、重要度をLightGBMのfeature importanceメソッドで算出した。重要視されていた要因は下記のとおりである。

説明変数	importance
味方 MF のyの長さ	7687
敵 FW ライン	7545
味方 MF のxの長さ	7225
味方 DF のyの長さ	7160
味方 MF のyの長さ	7104
味方 DF のxの長さ	7044
味方 DF ライン	7030
味方 FW ライン	6926
敵 MF のxの長さ	6902
敵 DF のxの長さ	6304
敵 DF ライン	5815
敵 DF のyの長さ	5815
敵 FW ラインのyの長さ	5319
味方 FW のyの長さ	5081
敵 FW のxの長さ	5026
味方 FW のxの長さ	4936
味方のプレスレベル平均	4376
敵のプレスレベル平均	4104
ボールy座標	3788
ボールx座標	3404

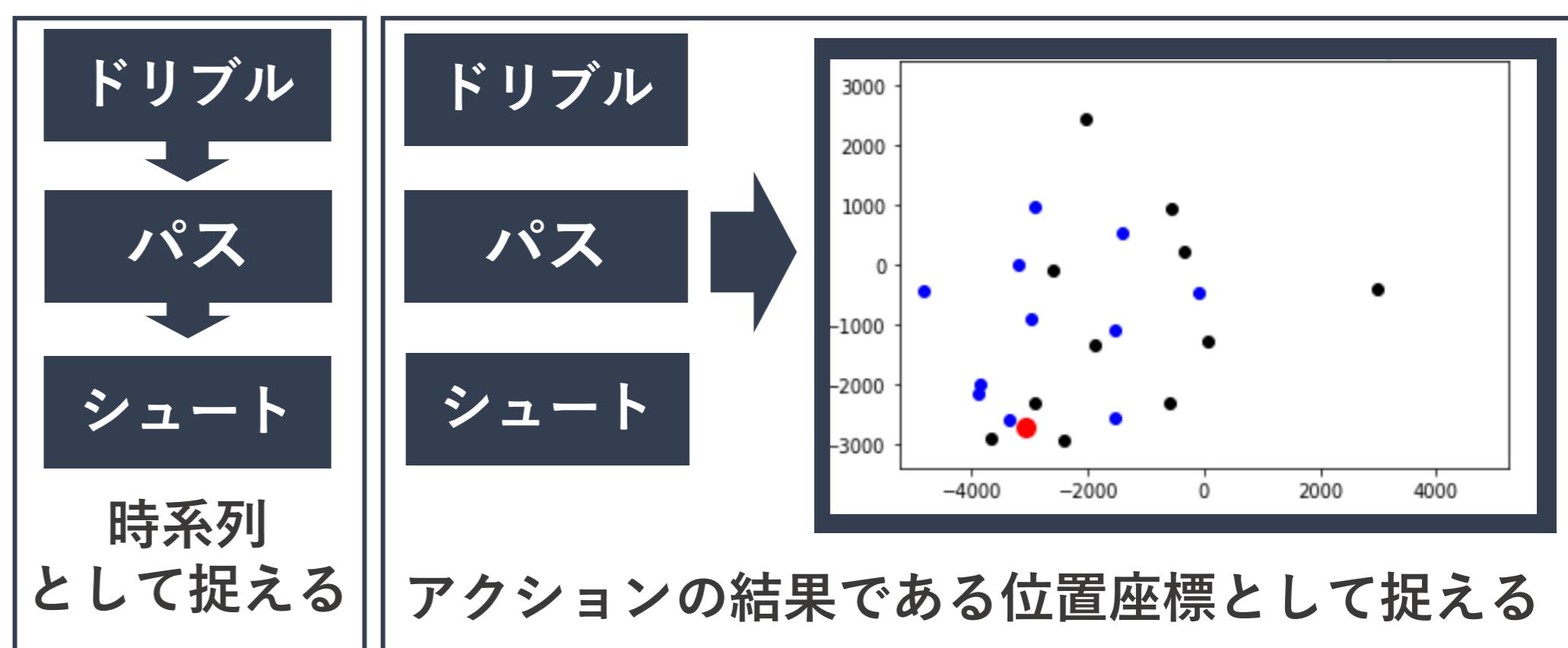
Light GBMの予測がシュートすると当てた画像の例



本予測モデルは、左図のように人間でも直感的に10秒以内にシュートすると予測できる場合だけではなく、右図のように人間には判断しづらい状況でもシュートすることを予測している。

考察 & 展望

その局面での情報はそれ以前の局面の動きの結果であり、ある時点の局面のデータというものは時系列のデータを内包していると考えられる。応用例として数値化されたシュート確率がホーム・アウェイでそれぞれ予測されるので、現在の局面においてどちらのチームが優勢かを未経験者でも一目で判断できるシステムの構築に繋げられる。



参考文献

[1] 松岡弘樹, 田原康寛, 安藤梢, 西嶋尚彦 (2019). “トラッキングデータからの戦術アクション項目の開発”。