情報提供資料

立入検査対象の選定を支援する機械学習AIアプリの概要について

東京消防庁

背景・課題

- 査察対象物は年々増加しており、現状の限られた人員で網羅的に対応することは困難である。当庁では「危険性に応じた効果的な防火安全対策の推進」を重点施策と位置づけ、立入検査を重点的・効率的に実施する体制に取り組んでいる。
- 当庁では2050東京戦略に基づき、歌舞伎町、渋谷などの歓楽街23か所を重点繁華街地域と位置づけ、査察（立入検査）により火災予防体制の強化を図っている。危険度が高い建物に優先的に立入検査を行いたいが、飲食店を含むビルに絞っても立入検査の候補は約5万棟あり、違反の可能性が高い建物を立入検査先として選定することは各署の大きな負担となっている。
- 過去に、外部委託（費用約2,000万円）でAIによる建物違反リスク予測に取り組み、有益な成果が得られたが、2つの課題が残った。
  - ① 継続するにはデータ更新のたびに多額の再委託費用が必要
  - ② 実務に活かすには職員が使えるシステム開発（業務アプリ等）が必要

プログラミング言語（Python）を活用し、機械学習AIアプリの内製開発した

機械学習AIの概要

- 機械学習AIは、総合予防情報システムの建物情報、事業所情報、過去の立入検査結果をもとに、建物ごとの違反リスクスコアを算出し、優先順位表を生成する。予測モデルは、重点繁華街地域の該否に分け、それぞれの特性に合わせた予測を行っている。



※1　建物情報及び事業所情報は年度当初のデータ
※2　違反一覧表は過去10年分（平成28年度から令和7年度）の立入検査データ

繁華街該当群（ent）

重点繁華街地域に存する建物の立入検査履歴がある建物群。
繁華街地域は重点査察の対象となりやすく、違反率自体が高い。
この運用差を反映するため、専用の予測を行っている。
既存の特徴量ベースラインで高い精度を達成している。

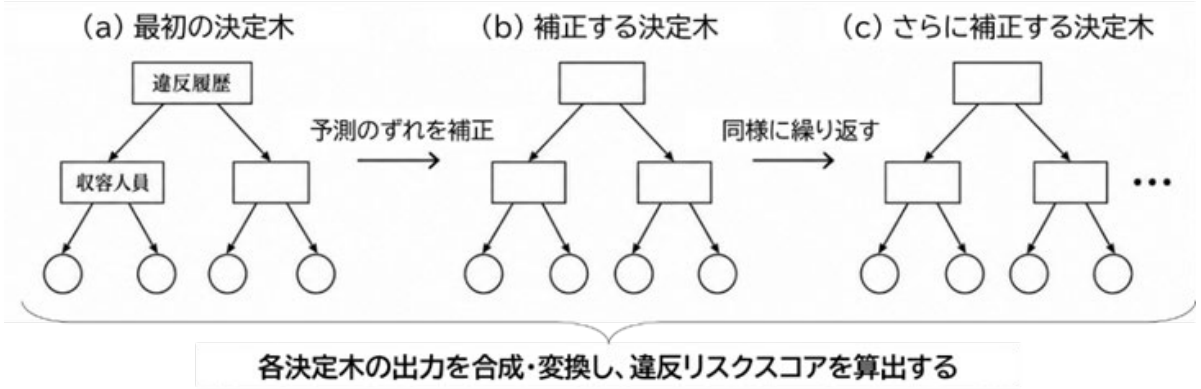
繁華街非該当群（non-ent）

重点繁華街地域に存する建物の立入検査履歴がない建物群。
違反率が低く、予測難易度が高い。単純な履歴件数だけでは拾いにくい対象を補うため、テナント構成の複雑さを表す追加指標を採用し、捕捉率を維持しつつ的中率を改善している。

機械学習AIの分析手法

- 今回の機械学習AIでは、**「勾配ブースティング決定木」**という分析手法を採用している。過去の違反の有無や収容人員の大小などの条件に応じて、枝分かれするようにデータを分けて予測し、条件を順に組み合わせることで、一つの条件だけでは分けにくい対象物にも対応できる。

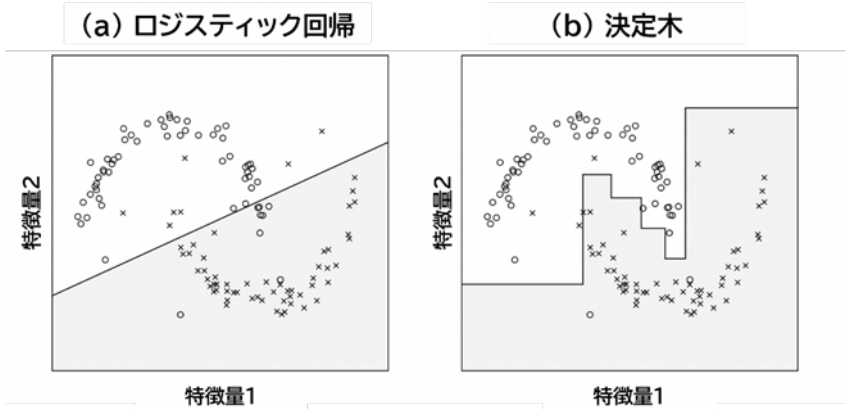
勾配ブースティング決定木のイメージ



1. 最初の決定木で予測し、その予測と実際の立入検査結果とのずれを確認する。
2. 次に、そのずれを小さくする決定木を追加する。
3. その後も、それまでの木を組み合わせた予測に残るずれを補うように、木を順次追加して学習する。
4. 学習後は、複数の決定木の出力を合成・変換し、建物ごとに0から一までの「違反リスクスコア」を算出する。

ロジスティック回帰と決定木の判定境界の比較イメージ

※ロジスティック回帰と決定木による分類の違いを、仮想データで示した模式図



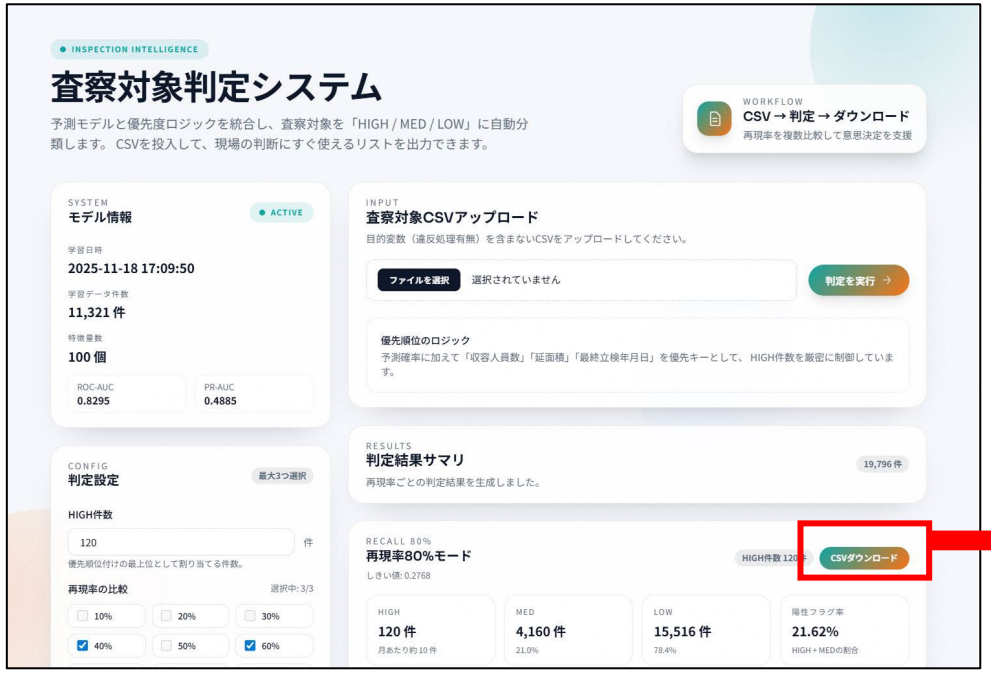
- ロジスティック回帰では、二つの特徴量をそのまま用いる場合、左図のように一本の直線が判定境界になる。
- 決定木では、特徴量ごとに条件を分岐させるため、右図のように縦・横の線を組み合わせた境界を表現できる。

  → 一本の直線では分けにくい特徴量の組合せにも対応できる。

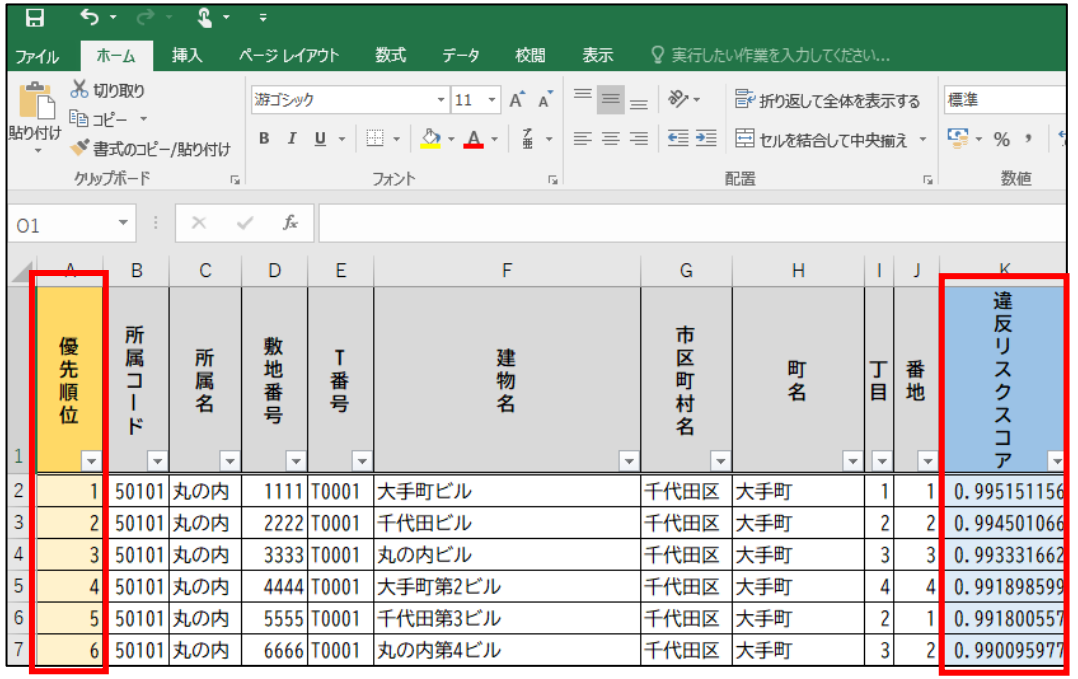
機械学習AIの使用方法

- 機械学習AIに総合予防情報システムから抽出した建物情報、事業所情報、違反一覧表のCSVデータを取り込むと、約80項目の特徴量をもとに機械学習された判定結果が表示される。

機械学習AI



違反予測リスト



	A	B	C	D	E	F	G	H	I	J	K
1	優先順位	所属コード	所属名	敷地番号	T番号	建物名	市区町村名	町名	丁目	番地	違反リスクスコア
2	1	50101	丸の内	1111	T0001	大手町ビル	千代田区	大手町	1	1	0.995151156
3	2	50101	丸の内	2222	T0001	千代田ビル	千代田区	大手町	2	2	0.994501066
4	3	50101	丸の内	3333	T0001	丸の内ビル	千代田区	大手町	3	3	0.993331662
5	4	50101	丸の内	4444	T0001	大手町第2ビル	千代田区	大手町	4	4	0.991898599
6	5	50101	丸の内	5555	T0001	千代田第3ビル	千代田区	大手町	2	1	0.991800557
7	6	50101	丸の内	6666	T0001	丸の内第4ビル	千代田区	大手町	3	2	0.990095977

機械学習AIは主管課でのみ操作可能

違反予測リストはセキュリティ上、主管課の所属デバイスに入れた機械学習AIで作成し、各消防署へ送付する。

違反リスクスコア（優先順位）の意味

違反リスクスコア（最大値「1」、最小値「0」）が高いほど、過去の立入検査結果上、避難施設関係違反が出た対象に近い傾向があることを意味する。

令和7年度の立入検査データによる検証結果

- 機械学習データ
  繁華街を擁する**主要5署（京橋、渋谷、新宿、豊島、池袋）**の査察対象物データをもとに機械学習
  ※立入検査データは**過去10年分（平成27年度から令和6年度）**
- 令和7年度の立入検査結果
  評価対象**1,413棟**のうち、実際に避難施設関係違反があった対象物は**450棟**であり、違反率は**31.8%**であった。
- 違反リスクスコアの検証結果
  違反リスクスコア上位の対象物による違反的中率は、上位20棟が**90%**、上位50棟が**86%**、上位100棟が**76%**となった。

検証対象の概要

指標	値
評価対象	1,413棟
うち避難施設関係違反あり	450棟
違反率	31.8%

違反リスクスコア上位の違反的中状況

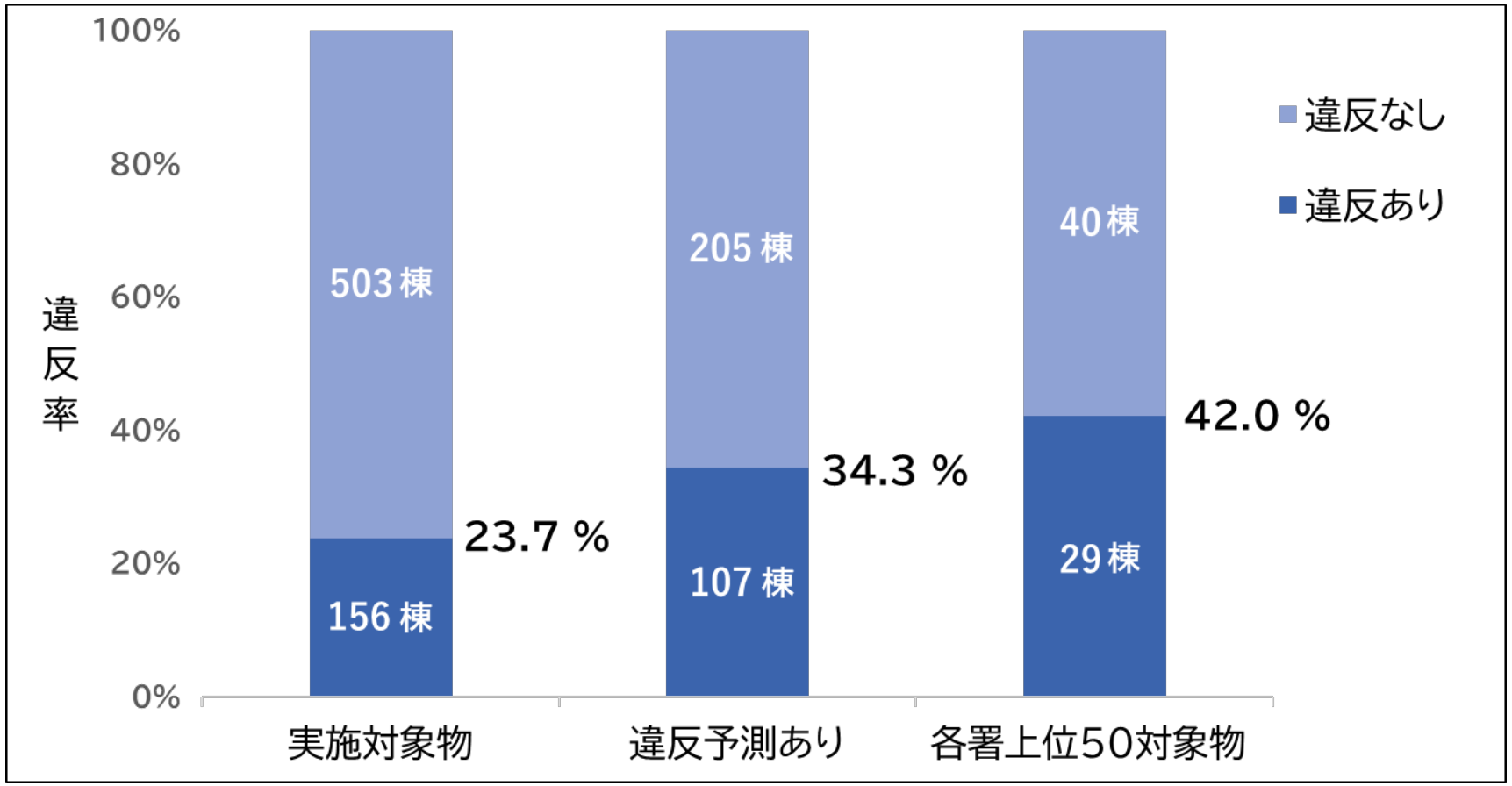
	上位20棟	上位50棟	上位100棟
違反棟数	18棟	43棟	76棟
違反的中率	90 %	86 %	76 %

違反リスクスコア上位の対象物　＝　避難施設関係違反がある可能性が高い対象物

違反リスクスコア上位の対象物による違反的中率は、実際の違反率を**大幅に上回り**、さらに違反リスクスコア上位の対象物ほど**違反的中率が高い**結果となった。

令和8年度の繁華街査察による検証結果

令和8年6月から8月の期間に9署で実施した繁華街査察の結果と、機械学習AIによる予測結果を比較した。今回の実施結果では、違反予測リストに掲載された対象物の違反率が、検査した対象物全体の違反率を上回った。違反リスクスコアを参考に対象物を選定することは、違反の可能性が高い建物への立入検査を優先する上で、有用である可能性が示唆された。



実施対象物と違反予測リストの違反率比較

実施対象物の違反率23．7%に対して、各署の違反リスクスコア上位50対象物の違反率は42．0%で、実施対象物全体の違反率を18.3ポイント上回り、約1.77倍となる。

※この比較は、同じ期間の実施対象物全体と、その中のリスト掲載対象物を比べたものである。職員による選定とAIによる選定の効果を独立した条件で比較したものではない。

予測モデルに寄与した特徴量について

- 今回の予測モデルでは、過去の避難施設違反履歴、収容人員・入居事業所数、用途・建物構造、査察履歴などを総合して、違反が出やすい対象を上位に並べている。
- あくまで「予測上の寄与」を示すもので、特徴量を改善すれば違反が減るという因果効果ではないことに注意が必要である。

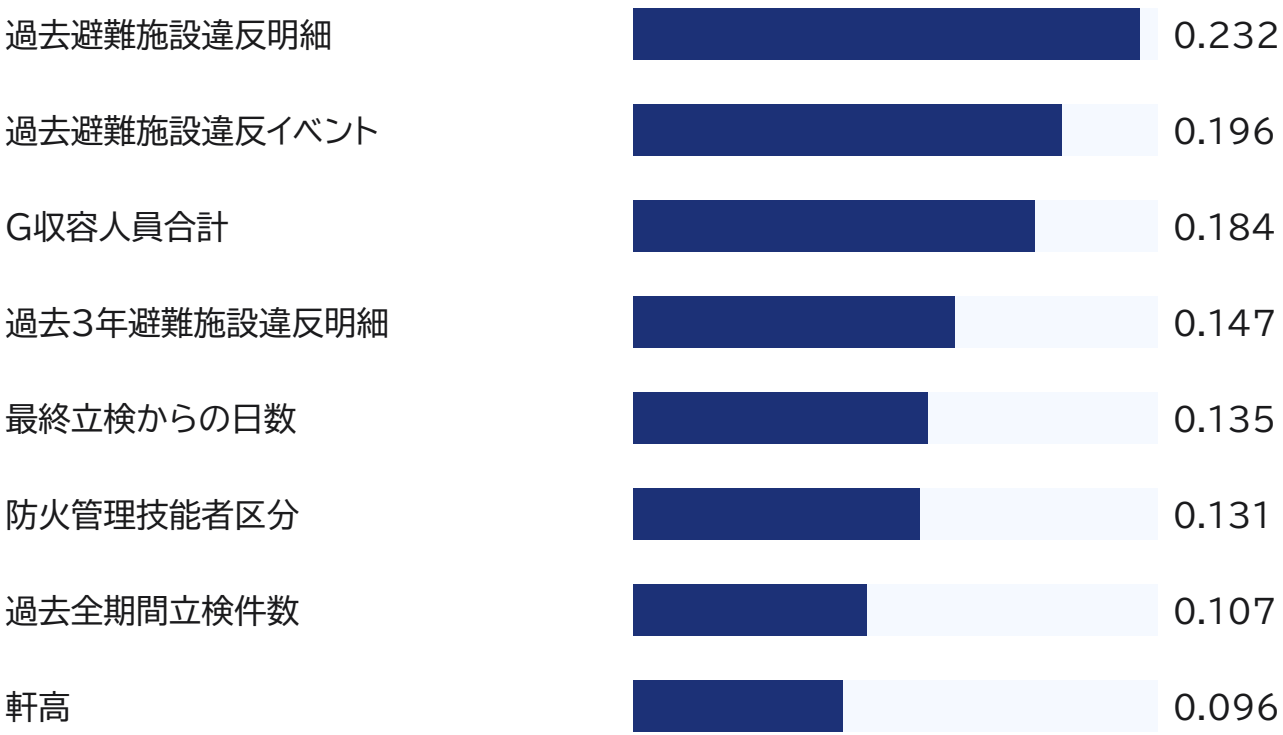
主要な特徴量

	代表特徴量	解釈
過去の避難施設違反履歴	egress_rows_all, egress_events_all, egress_rows_past_3y	過去に避難施設関係違反が多い建物は、再度違反が出やすいと学習している
G側の規模・複合度	g_capacity_sum, g_count, g_capacity_max, g_retail_ratio, g_capacity_cv	入居事業所や収容人員が多い、または用途・人員構成が複雑な建物を重視している
T側の建物属性	建築面積_num, 収容人員数_num, 建物の高さ, 軒高, 階層地下階数_num, 階層地上階数_num	建物の規模・階層・構造的複雑さを見ている
用途・地域コード	用途地域区分コード, 政令別表コード, 防火地域コード, 風俗施設有無コード	非繁華街側では、用途や地域の違いが効いている
査察・管理関連履歴	days_since_last_inspection, inspections_all, 防火管理技能者該当号数コード, g_night_guard_no_count	既知の管理リスクや査察履歴をリスクとしてみなしている

予測に寄与した特徴量（繁華街）

- entモデルは、過去の避難施設違反履歴が最上位で、再発リスクを強く重視している。G収容人員や査察履歴も併せて、既知リスクの濃い対象を上位化している。

mean |SHAP| 上位8特徴量



方向性の要約

高いほどリスク方向

過去避難施設違反明細、G収容人員合計、過去3年避難施設違反明細は、高値側で違反あり方向に寄与している。

査察履歴の解釈

最終立検年月日からの日数は値が高いほどリスクを下げる方向。近年、立入検査が実施されている対象ほど既知リスクとして上がりやすい可能性がある。

実務上の示唆

重点繁華街地域に存する建物は「過去違反があり、入居Gの収容人員が大きい対象」を重点確認候補として見る。夜間警備なしG数も管理体制の弱さを示す補助情報になり得る。

予測に寄与した特徴量（非繁華街）

- nonentモデルは、過去違反よりもG収容人員、用途地域、建築面積、収容人員、階層などの建物・用途属性を強く見ている。

mean |SHAP| 上位8特徴量



方向性の要約

収容人員は強いリスク要因

G収容人員合計と収容人員数は、高値側で違反あり方向に寄与している。人が集まる規模の大きさを強く見ている。

建築面積は単純な大規模リスクではない

建築面積は高値側でリスクを下げる方向に出ている。比較的小面積でも収容人員が多い対象、すなわち密度の高さが示唆される。

繁華街地域と比べて用途・構造が重要

用途地域区分、政令別表コード、地下階数、検査単位数などが上位。