人工知能を活用した犯罪・交通事故
発生予測技法の調査研究

補遺
（交通事故発生予測編）

平成 31 年 3 月

目　次

1　実証実験の実施概要

(1)　実証実験について・・・

【実証実験での交通事故及び事故類型】

本調査研究を行うに当たり、実証実験での交通事故及び事故類型は次のとおりとする。

○　交通事故

交通事故とは、道路交通法第 67 条第２項「車両等の交通による人の死傷又は物の損壊」をいう。

○　人身交通事故

交通事故統計上の人身交通事故は、「道路上における車両等及び列車の交通による人の死傷を伴う事故」とされており、自殺、確定的故意、建物等から転落し車両等との衝突等による死傷、落下物の直撃による運転者等の死傷及びその他自然災害等により運転者等の死傷した場合には人身交通事故として取り扱っていない。

○　物件交通事故

物件交通事故は、「車両等の交通による物の損壊を伴う事故」とされている。

物件交通事故は、人身交通事故とは違い、道路外で物の損壊があった場合や故意に物を損壊させた場合等であっても物件交通事故としている。

○　人対車両

人対車両については、歩行者が第１当事者又は第２当事者となった事故の類型である。

○　車両相互

車両相互については、当該事故が第１当事者及び第２当事者が共に車両等の事故である。

○　車両単独

車両単独については、第２当事者が物件や相手なし、駐車車両の事故である。

1　実証実験の実施概要

(1) 実証実験について

人工知能を用いて交通事故発生予測を行う実証実験を３回に分けて実施した。

ア　第１回実験

単一の機械学習アルゴリズム、必要最低限と考えられる特徴量、神奈川県内全域を予測対象として、簡単な実験設定で予測実験を行った。

第２回目以降に実験する人工知能による交通事故発生予測のイメージの共有及び課題の洗い出しを行うことを目的とした。

結果は委員会において交通事故発生予測のイメージを共有した。また、課題としては、特徴量不足や予測対象範囲の絞込みなどが挙がった。

イ　第２回実験

第１回実験の結果を踏まえ、オープンデータ等から入手可能な特徴量を追加するとともに、予測対象エリアを絞った。この条件の下、複数の機械学習アルゴリズム間の精度比較を行うことを目的とした。

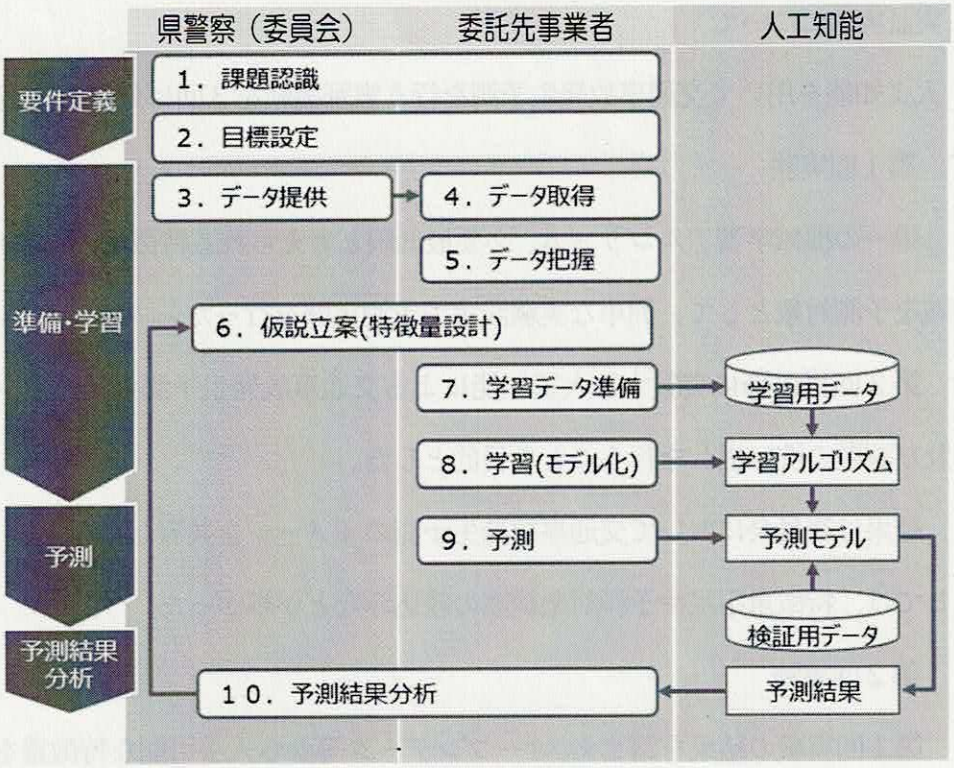
複数の機械学習アルゴリズムの精度比較の結果、予測精度の高さに基づいて機械学習アルゴリズムの絞込みを行った。

ウ　第３回実験

第２回実験で予測精度が高かったアルゴリズムで実証実験を行った。特徴量の見直し、絞り込みを行った状態でも予測精度を維持できるかを確認することを目的とした。

結果はわずかな予測精度低下で抑えられ、予測対象に関係のある特徴量の重要度の傾向が分かりやすくなった。

◆ 図表1－1　実証実験ワークフロー



No.	プロセス名	実施概要
1	課題認識	実証実験の目的に認識齟齬がないことを委員会において確認する。
2	目標設定	委員とともに課題に対する業務活用目標を設定する。 また、目標を達成するための精度目標も設定する。 その他分析対象の交通事故の選定等も行う。
3	データ提供	実証実験に必要な業務データ等を神奈川県警察から委託事業者へ提供する。
4	データ取得	提供されたデータやオープンデータ等を実証実験環境へ移行する。
5	データ把握	データ特性の把握を行う。主に基本統計分析を行い、データが人工知能の特徴量として利用可能かどうか検証する。
6	仮説立案(特徴量設計)	データ把握で取得した情報を基に業務活用目標に対し、どのようなデータを人工知能に学習させると目標を達成できるかという仮説を立案する。立案した仮説を基に特徴量の設計を行う。
7	データ準備	仮説及びデータの特性を考慮し、学習用データ(特徴量)を作成する。
8	学習(モデル化)	学習用データを用いて機械学習を行い、予測モデルを作成する。
9	予測	予測モデルの精度を評価するため、検証用データを用いて予測を行う。
10	予測結果分析	機械学習でよく使われる評価指標に基づいて、モデル別に予測結果の分析、評価を実施。 分析結果に基づいて、精度向上案を委員会において検討する。

(2) 人工知能について

本調査研究における「人工知能」とは、「機械学習を行うアルゴリズム」と定義した。

複数の機械学習アルゴリズムを用いて交通事故の発生予測の実験を行い、その結果を比較することで、交通事故予測に適した機械学習アルゴリズムを探ることが実験の目的の一つである。

本実証実験において使用する機械学習アルゴリズムは、インターネット上にライブラリとしてプログラムのソースコードが公開され、誰もが無償で利用可能なオープンソースを使用した。

オープンソースを使用する理由は

○ アルゴリズムの仕組みやプログラムが広く公開されていることから、アルゴリズムを比較評価する公平性や透明性を確保できること。

○ 世界中のユーザーによりバグ修正や改良が進められているため、動作安定性が高く、利用可能なライブラリ等も豊富であること。

○ Kaggle(1)で開催される予測コンペなどで、オープンソースソフトウエアが用いられ、高い予測性能を上げており、それらの結果を活かした精度向上の可能性を秘めていること。

などの観点からである。

実証実験で使用するオープンソースソフトウエアは、以下の４つである。

括弧内は、アルゴリズムの分類を示す。

・ロジスティック回帰（確率的線形判別）：scikit-learn[2]

・XGBoost（決定木＋アンサンブル学習）：PyXGBoost[3]

・サポートベクターマシン（カーネル法）：scikit-learn

・ディープラーニング（多層ニューラルネットワーク）：

TensorFlow[4]、Keras[5]

この４種類のアルゴリズムを選定した理由については図表１－２のとおりとする。

◆ 図表１－２　アルゴリズムの特徴と採用理由

	アルゴリズム	特徴	選択理由
1	ロジスティック回帰	•代表的な機械学習アルゴリズムの一つ •商品を購入するか否か、といったクラス分類の予測をする時に用いられる •設定するハイパーパラメータが少なく、計算も速い •複数のアルゴリズムによる精度比較の際のベースとして用いられることが多い	様々な予測問題において扱われてきた実績がある。また、シンプルな線形回帰であるため、学習したモデルの解釈を行いやすいというメリットがある。比較のためのベースとして採用する。
2	XGBoost	•決定木と呼ばれる木構造のルールをいくつも作成して予測するアルゴリズム •予測値が外れたときに、その誤差をさらに埋めるよう学習が進む •アンサンブル学習と呼ばれる方法で高い予測精度を誇るが、設定すべきハイパーパラメータが多い	世界的なデータ解析コンペティションで高い精度を出したことで有名になり、さまざまな予測に利活用されている。学習モデルが予測に際して重視している特徴量を確認することが可能である。犯罪予測にも適しているアルゴリズムと考えられるため採用する。
3	サポートベクターマシーン	•代表的な機械学習アルゴリズムの一つ •カーネルと呼ばれる方法で非線形分離が可能 •少ないデータでも高い予測精度を期待できるが、データが大量であると計算時間が膨大化する	ロジスティック回帰よりも複雑な問題を解くことに長けている。学習モデルの解釈は難しいが予想精度は一般的に高い傾向があるため採用する。
4	ディープラーニング	•人間の脳神経回路をモデルにした多層構造アルゴリズム •画像、音声、自然言語の分野で高い認識精度を誇る •一般的に大量のデータが必要である場合が多く、また学習した結果を人間が解釈することが難しいため、世界中で更なる研究が進められている	近年の人工知能ブームの中心にあり、画像や音声認識で高い予測精度を発揮していることから、犯罪発生予測に適用できるか検討するために採用する。

(3) 予測結果の評価方法について

人工知能の予測性能の評価については、様々な評価指標が使われている。

交通事故の発生予測を例として説明すると、予測結果と実際の結果は以下のような混合行列となり、これをクロス集計して評価することになる。

◆ 図表１－３　混合行列

	「発生する」と予測	「発生しない」と予測
実際に発生	真陽性（TP）	偽陰性（FN）
実際に未発生	偽陽性（FP）	真陰性（TN）

この４つの数値（TP,FP,FN,TN）を用いて、一般的に使われる評価指標を挙げる。

・正解率（Accuracy）

Accuracy=(TP+TN)/(TP+FP+FN+TN)

「発生する」「発生しない」を正しく予測した割合。

・適合率（Precision）

Precision= TP/(TP+FP)

「発生する」と予測して実際に発生した割合。

・再現率（Recall）

Recall＝TP/(TP+FN)

実際に発生した場所を「発生する」と予測できた割合。

・F 値（F-measure）

F-measure=(2×Precision×Recall)/(Precision+Recall)

適合率と再現率の調和平均。トレードオフとなる双方の評価値のバランスの良さを表す総合的評価指標といえる。

・リスク比

Risk Ratio = (TP/(TP+FP))/(FN/(TN+FN))

人工知能の能力評価指標ではないが、発生予測場所と未発生予測場所における発生率の比。

「発生する」と予測した場所が「発生しない」と予測した場所よりも、交通事故が起きるリスクが何倍高いかを表す指標であり、全くランダムに予測するモデルである場合のリスク比は１程度になる。

理論上は無限大までの数値を取ることがある。

どの評価指標を使用して予測モデルの評価をするかということが問題であるが、適用する対象やどのような予測結果を使用するかで重視する評価指標を使い分ける必要がある。

一見すると「正解率」が高い予測モデルが高性能である、という印象を受けるかも知れないが、データに偏りがある場合には注意する必要がある。

交通事故の発生は稀な事象であり、予測単位の設定にも大きく依存するが、発生データ、未発生データは通常は未発生側に大きく偏っていると考えられる。

正解率に関していえば、算出式を見ればわかるとおり、データの偏りが大きければ大きいほど、偏っている側の予測（真陰性予測）をすれば正解率が相当に高くな

ってしまうため、交通事故の予測に関していえば全て「発生しない」という予測をするだけで正解率が99%程度の数値になってしまう。

実際には何の予測もできないモデルが非常に高い正解率を出してしまうこともあるため、交通事故の発生予測の評価指標としては適さない（100%にならない限り大きな意味はない）と考えられる。

交通事故の発生予測に関していえば「発生する」という予測が

・正確な予測か

・漏れがない予測か

という部分が最も重要な要素であるといえる。

発生予測の正確性は適合率で、網羅性は再現率で判断することができるが、適合率と再現率は、片方の指標の値が上昇すると片方の指標の値が下降するトレードオフの関係であるため、どちらを重視するか、どの程度のバランスを取るかを考える必要がある。

適合率を上げることで、真陽性予測の正確性が重視されるようになるが、計算式を見てもわかるように偽陰性予測を考慮しない指標であるため、偽陰性予測が増えてしまう可能性が高まる。

一方、再現率を上げると交通事故発生の取りこぼしを減らすことが可能であるが、偽陽性予測を考慮しない指標であるため、偽陽性予測が増えてしまうことも考えられる。

この２つの評価指標の双方のバランスを重視した指標がＦ値である。

Ｆ値を用いることで、適合率と再現率というトレードオフになる２つの評価指標を、統合して一つの指標として評価できることが利点である。

アルゴリズムやモデルの比較評価を行う場合は、Ｆ値を利用することで総合的な性能評価が可能であると考えられる。

人工知能を用いた交通事故予測に関していえば、適合率と再現率、その総合的な指標であるＦ値、どの程度の数値やバランスであれば実業務に利用可能かを見極めながら評価する必要があると考えられる。

(4) 交通事故の利用データ

交通事故発生データとして

・交通事故統計原票(神奈川県警察交通事故管理システムより抽出)

を使用した。

交通事故統計原票は警察庁の統一内容のものであり、その中から発生年月日、発生時間、事故類型、緯度経度等の基本的な項目を抽出して使用した。

交通事故統計原票の情報を使用した予測事故類型は

・人対車両

・車両相互

・車両単独

である。

ア　神奈川県警察交通事故管理システム

交通事故管理システムは、交通事故統計原票関係の管理を行うものであり、人身交通事故処理を担当する警察官が神奈川県警察の基幹ネットワーク端末により人身交通事故に係る各項目を入力し、警察本部において集約した上で、交通事故統計原票に変換している。交通事故管理システムは統合地図情報管理業務（以下「統合地図システム」という。）と連動しており、この統合地図システムを活用することにより、交通事故発生場所の緯度経度の入力が可能となっている。

なお、2012 年以前の 4 年分の交通事故データについては、緯度経度が付与されていないため、神奈川県警察にて緯度経度を付与して使用することとした。

2　実証実験結果

(1) 第１回実験

神奈川県全域を対象エリアとし、人対車両の交通事故の予測を行った。

実験設定は次のとおりである。

ア　実験の設定

(ｱ) 機械学習アルゴリズム

使用する機械学習アルゴリズムはロジスティック回帰の１種類のみを用いた。

ロジスティック回帰を用いる理由としては

○　処理速度が高速である。

○　学習したモデルの解釈が行いやすい。

などの理由である。

過学習（オーバーフィッティング）対策としては交差検証（クロスバリデーション）やＬ１正則化により対応した。

(ｲ) 特徴量

交通事故等の発生に関すると考えられる、日時場所、天候や人口、主な最寄り施設までの距離等を特徴量として使用した。

使用した特徴量は、別添２「特徴量一覧」の第１回実証実験を参照。

(ｳ) 対象期間

学習及び予測の対象とした期間は、以下のとおりである。

◆ 図表２－１　第１回実験学習設定

事故類型	学習期間	正例学習件数	予測評価期間
人対車両	2017年1月～11月	4,344	2017年12月

データの最新１ヶ月分を除いた期間を学習データとして使用した機械学習アルゴリズムに学習させることで得られた予測モデルにより最新の１ヶ月分を予測させて、実際の最新１ヶ月分と比較することにより予測の評価を行った。

学習データは、学習期間内の正例（交通事故が発生したメッシュの状況）及び、負例（交通事故が発生しなかったメッシュの状況）を用いた。

交通事故に関していえば、正例及び負例の比率に大きく偏りがあり（負例が圧倒的に多い）、そのまま学習させると予測に偏りが生じる（負例が多いまま学習すると、負例を予測することで正解率が上がることから、そのような予測結果を出す方向に学習が進む。）ため、負例については負例の全体から正例よりやや多い（約 1.2 倍）件数をランダムに選択して減じた上で、学習時に正例と同数に調整する処理（アンダーサンプリング）を行い、学習させることとした。

(ｴ) 予測対象メッシュ

予測評価メッシュ数：32,041,600

内訳：神奈川県内総メッシュ数 258,400×予測対象日数 31 日×1 日当たりの時間分割数４＝32,041,600

予測の時間単位は第１回実験に限り、特徴量の設計や学習時間の関係から、６時間単位で行った。

予測の判断を分ける閾値(しきいち)については、得られた発生傾向確率が50%以上であれば「発生する」という予測結果、50%未満であれば「発生しない」という予測結果として扱い、予測結果と評価データを比較し、「予測が当たったかどうか」を示す評価指標を算出した。

イ　実験結果

予測実験を行った結果は、次のとおりである。

◆　図表２－２　第１回実験予測結果

事故類型	TP	FP	FN	TN	適合率	再現率	リスク比	F値
人対車両	537	17,710,700	5	14,330,358	0.0030%	99.08%	86.90	0.0061%

再現率が極めて高く、実際に人対車両の交通事故が発生した場所は、ほぼ網羅して予測できていた。

その一方で、再現率とトレードオフの関係となる適合率が極めて低い結果となった。

人対車両で言えば、発生すると予測した場所を全てカバーすればほぼ全ての発生を捕捉できるが、発生を予測したメッシュで実際に交通事故が発生する確率は0.0030%と低く、偽陽性メッシュが大多数を占めるという予測結果となった。

全体的に偽陽性予測(FP)に偏っているため、発生の取りこぼしは少ないものの、実務に活用しようとすると、対応に必要なリソース（警察官）が足りず、また、予測が外れて実際には発生しない場所が多いため(神奈川県全体の約５割)、警察力の効率的な配置を行うという目的には沿わない予測結果となった。

このような結果となった理由としては使用した特徴量が少なかったことが考えられる。

ロジスティック回帰は、学習で得られた予測モデル（回帰式）から予測に寄与する特徴量が分かるため、これを確認したところ

・時間帯

・土地利用区分

などのメッシュ間で差が出にくいものが重視されていた。

このことから、メッシュ間で差が出やすい特徴量が不足していることにより交通事故発生メッシュと未発生メッシュの特徴の差異を人工知能が捉えきれず、予測結果が偏る結果となったことが考えられた。

(2) 第２回実験

第１回実験の結果を踏まえ、

・複数の機械学習アルゴリズムの適用

・予測対象事故類型の追加

・特徴量の追加

・対象メッシュの絞り込み

などを行い、予測精度の変化を確認することを目的として第２回の実験を行った。

実験設定は以下のとおりである。

ア　実験の設定

(ｱ) 機械学習アルゴリズム

同一条件で前述の４種の機械学習アルゴリズムを用いて学習と予測を行い、その予測性能を比較する。

(ｲ) 特徴量

第１回実験と比較して大幅に特徴量を追加した。追加した特徴量は警察業務データ、オープンデータなどのデータから作成した。

また、交通取締り関係として、「人対車両(取締り)」「車両相互(取締り)」「車両単独(取締り)」の事故類型にのみ

・直近30日間の違反取締り件数

・直近60日間の違反取締り件数

を追加した。

使用した特徴量は、別添２「特徴量一覧」の第２回実証実験を参照。

(ｳ) 取締り件数について

直近30日・60日間の違反取締り件数は、それぞれの期間内にメッシュ内での交通取締り件数となる。

(ｴ) 一般道の交通量について

交通量はこれまで道路交通センサスを用いた研究が多いが、本調査研究では断面交通量を使用することとした。

断面交通量は、国道、県道、主要地方道等の主な幹線道路に2,375か所設置されている車両感知器から、1時間単位ごとの車両通過台数を測定したものを使用している。

そのため、車両感知器が設置されている場所の緯度経度を特定し、測定した交通量をデータ化し、対象メッシュ内に測定地点として設定した。

実証実験を行うにあたり、測定地点で測定した車両台数を有効に活用して測定地点以外に交通量の特徴量とすることとした。

(ｵ) 予測対象事故類型

予測対象事故類型については、第1回実験の「人対車両」に加え、「人対車両(取締り)」「車両相互」「車両相互(取締り)」「車両単独」「車両単独(取締り)」及び一般道と区別した「高速道路等」とした。

(カ) 対象期間

学習及び予測の対象とした期間は、以下のとおりである。

◆ 図表２－３　第２回実験学習設定

事故類型	学習期間	正例学習件数	負例学習件数	予測評価期間
人対車両	2008 年 1 月～2017 年 6 月	48,628	58,814	2017 年 7 月～2018 年 6 月
車両相互		234,014	286,337	
車両単独		6,589	8,459	
高速道路等		3,190	3,996	
人対車両(取締り)	2017 年 3 月～2018 年 5 月	4,991	7,736	2018 年 6 月
車両相互(取締り)		21,476	38,287	
車両単独(取締り)		697	1,098	

それぞれのデータの初期１年分及び最新１年分を除いた期間を学習データとして使用し（初期１年分については特徴量として利用）、最新１年分から各月ごとランダムに３日又は４日を選択した計 40 日を予測評価期間として使用した。

１年間からランダムに 40 日を抽出している理由は、交通事故発生には地域性や季節性があると考えられていたためである。

季節性については、車両相互事故を市郡に分け地域性と季節性についてトレンドの分析を行ったところ、鎌倉市が他の市郡と比較して夏季に事故が多発する特異な季節性を持つことが西[1]らにより報告されている。

それに関しては、観光地としての特性から特異な交通量の季節的な変動があるとし、また、箱根町を含む足柄下郡についても秋季から冬季にかけて交通事故が多発す傾向があり、これも観光地としての特性を考えれば合理的であるとしている。

このようなことから、神奈川県内には特異な地域性や季節性があると考え、季節による影響がある可能性を考慮して、１年間の 40 日を抽出した。

(ｷ) 予測対象メッシュ(一般道)

予測評価メッシュ数：32,648,000

内訳：対象メッシュ(居住区＋道路＋鉄道)数 102,025×予測対象日数 40 日

×１日当たりの時間分割数 8＝32,648,000

第２回実験については、当初の設定のとおり予測時間単位を３時間で行った。

また、予測の対象とするメッシュは、第１回実験の「神奈川県内全域」から変更し、「道路メッシュ」「居住区メッシュ」及び「鉄道メッシュ」に対象を限定することとする。

これらは、国土交通省の国土数値情報オープンデータで提供されているものを使用した。

道路メッシュは、「緊急交通路」と「土地利用区分情報のうち、道路などで、面的に捉えられているところ」と定義されている。

居住区メッシュは、土地利用区情報のうち、「住宅地・市街地等で建物が密集しているところ」と定義される。

鉄道メッシュは、「鉄道、操車場などで、面的に捉えられているところ」と定義される。

100 メートル分割標準メッシュの「道路メッシュ」は 24,229 メッシュ、居住区に区分されるメッシュは 76,246 メッシュ、そして、「鉄道メッシュ」に区分されるメッシュは 1,550 メッシュである。合計すると 102,025 メッシュとなり、神奈川県全域の約 40%を占める。

メッシュ数としては約60%の削減となるが、削減された利用区分は森林、河川、海浜、海水域といった本調査研究における予測対象事故類型が発生する可能性が低い場所である。

実際に交通事故のデータで確認すると、実際に発生したメッシュのカバー率は約87%となり、削減率に対して発生場所の未カバー率は低く抑えられた。

メッシュ数を絞り込むことで得られる効果としては

○　予測対象数が減少することによる処理速度の向上

○　発生メッシュ、未発生メッシュの不均衡（未発生メッシュが多数を占める）という状況の改善

○　発生可能性の少ないメッシュの除外による予測精度の向上

が考えられる。

(ｸ) 予測対象メッシュ(高速道)

予測評価メッシュ数：1,706,560

内訳：対象メッシュ：高速道路メッシュ数5,333×予測対象日数40日×1日当たりの時間分割数8＝1,706,560

高速道路メッシュについては高速道路を目視にてメッシュとして抽出した。

イ　実験結果

(ｱ) 予備実験結果

予測対象メッシュ（居住区メッシュ）全体で予測のための学習をさせたところ、サポートベクターマシン（以下「ＳＶM」という）の学習処理が現実的な時間で終了しないという問題が発生した。

ＳＶMは学習データ量の２乗に比例した学習時間が必要となることから、データ量が増えると急激に学習時間が伸びることが原因と考えられた。

そこで、予備実験として予測対象罪種は空き巣とし、ＳＶMが現実的な時間内で学習処理が可能なデータ量である 1,000 件程度に学習データを減らし、4 つの機械学習アルゴリズムで予測精度の比較を行った。

予備実験の結果は次のとおりである。

◆　図表２－４　第２回実験（予備実験）結果

アルゴリズム	TP	FP	FN	TN	正解率	適合率	再現率	リスク比	F 値
ロジスティック回帰	88	14,602,633	18	11,846,933	44.79%	0.00060%	83.02%	3.97	0.00121%
XGBoost	74	10,152,301	32	16,297,265	61.62%	0.00073%	69.81%	3.71	0.00146%
SVM	90	14,498,814	16	11,950,752	45.18%	0.00062%	84.91%	4.64	0.00124%
ディープラーニング	104	25,259,478	2	1,190,088	4.50%	0.00041%	98.11%	2.45	0.00082%

予測対象日については、絞り込んだ学習データで実際に使用している変数の関係から、予測ができなかった３日分（学習データの大幅な削減のため、予測結果と評価データの不整合が生じたため。）を除外した 37 日分で評価した。

この結果から、ＳＶMについては現実的な学習データ数で学習するとロジスティック回帰と同程度の予測性能であった。また、一般に大量の学習データ数

を必要とすると言われるディープラーニングは、学習データ量が絶対的に不足しているのか、予測が発生方向に大きく振れていた。

XGBoost については、再現率はやや低いものの適合率は高く、少ない学習データ量でも発生の特徴をつかんだ学習が出来ていると見ることができた。

以上の予備実験から、ＳＶＭは

○　実際に運用を考えると、現実的な時間での学習処理が難しい。

○　現実的な処理が可能な件数で比較しても精度は XGBoost に及ばない。

ということが分かった。

そのため、ＳＶＭに関してはこの予備実験の結果をもって比較実験の対象から除外することとした。

なお、ＳＶＭに関する交通事故の予測については、犯罪の予備試験と同様の結果が見込まれることから実施していない。そのため交通事故予測での予測精度については不明である。

(イ) 本実験結果

予備実験の結果を踏まえ、ＳＶＭ以外の３つの機械学習アルゴリズムを用いて予測実験を行った。これを本実験とする。

結果は、次の表のとおりである。

◆ 図表２－５　第２回実証実験結果　人対車両

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
ロジスティック回帰	391	16,310,009	37	16,337,563	0.0024%	91.36%	10.59	0.0048%
XGBoost	307	7,106,790	121	25,540,782	0.0043%	71.73%	9.12	0.0086%
ディープラーニング	363	14,661,859	65	17,985,713	0.0025%	84.81%	6.85	0.0050%

◆ 図表２－６　第２回実証実験結果　車両相互

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
ロジスティック回帰	1720	15,782,347	185	16,863,748	0.0109%	90.29%	9.93	0.0218%
XGBoost	1344	5,872,075	561	26,774,020	0.0229%	70.55%	10.92	0.0458%
ディープラーニング	1764	16,928,525	141	15,717,570	0.0104%	92.60%	11.61	0.0208%

◆ 図表２－７　第２回実証実験結果　車両単独

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
ロジスティック回帰	52	18,255,283	15	14,392,650	0.0003%	77.61%	2.73	0.0006%
XGBoost	34	7,056,512	33	25,591,421	0.0005%	50.75%	3.74	0.0010%
ディープラーニング	47	16,940,954	20	15,706,979	0.0003%	70.15%	2.18	0.0006%

◆ 図表２－８　第２回実証実験結果　高速道路等

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
ロジスティック回帰	10	1,027,384	0	679,166	0.0010%	100.00%	―	0.0019%
XGBoost	8	233,147	2	1,473,403	0.0034%	80.00%	25.28	0.0069%
ディープラーニング	9	842,003	1	864,547	0.0011%	90.00%	9.24	0.0021%

◆ 図表２－９　第２回実証実験結果　人対車両（取締り）

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
ロジスティック回帰	259	11,395,506	21	13,090,214	0.0023%	92.50%	14.17	0.0045%
XGBoost	211	6,311,910	69	18,173,810	0.0033%	75.36%	8.80	0.0067%
ディープラーニング	181	7,894,680	99	16,591,040	0.0023%	64.64%	3.84	0.0046%

◆ 図表２－10　第２回実証実験結果　車両相互（取締り）

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
ロジスティック回帰	1103	11,360,571	90	13,124,236	0.0097%	92.46%	14.16	0.0194%
XGBoost	859	5,031,265	334	19,453,542	0.0171%	72.00%	9.94	0.0341%
ディープラーニング	797	5,019,734	396	19,465,073	0.0159%	66.81%	7.80	0.0317%

◆ 図表２－11　第２回実証実験結果　車両単独（取締り）

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
ロジスティック回帰	30	12,224,035	5	12,261,930	0.0002%	85.71%	6.02	0.0005%
XGBoost	29	9,137,900	6	15,348,065	0.0003%	82.86%	8.12	0.0006%
ディープラーニング	0	1,442,447	35	23,043,518	0.0000%	0.00%	0	0.0000%

第２回の方が第１回よりも問題設定としては難しくなっている（時間の分割数が多い）が、それでも大幅に精度が低下しなかったのは、対象メッシュ数の削減による予測対象数の減少や新たに加えた特徴量が予測結果に貢献しているためと考えられる。

機械学習アルゴリズムの予測性能について比較すると、全ての対象事故類型でXGBoostが最も高いＦ値を示した。

再現率に関しては他のアルゴリズムよりもやや劣る値であったが、リスク比は他のアルゴリズムよりも高い数値を出しており、総合的に考えれば、「発生する」という予測結果が当たりやすいモデルが構築できているという結果となった。

ディープラーニングに関しては、適合率・再現率ともに概ねロジスティック回帰と同程度の数値に止まる結果となった。

(3) 第3回実験

第2回実験の結果を踏まえ、機械学習アルゴリズム及び特徴量の絞り込み、さらに松浦の著書[2)]において、高齢者の交通事故は特異な傾向があると指摘されていることから「人対車両」の中から年齢65歳以上の第一当事者に該当する交通事故を抽出し「人対車両（高齢者）」を事故類型に追加した。

実験設定は、次のとおりである。

ア　実験設定

(ア) 機械学習アルゴリズム

第2回実験の結果から、機械学習アルゴリズムをXGBoost及びディープラーニングに絞り込み、予測精度を比較した。

第2回実験では、ディープラーニングはXGBoostには及ばなかったが、パラメーターを調整することで更なる精度向上の可能性があるディープラーニングと最も精度の良かったXGBoostを用いて実験を行うこととした。

この結果から交通事故予測に適当と考えられる機械学習アルゴリズムを検討する。

(イ) 特徴量

第2回実験に特徴量として含めていなかった、交通事故と関係があると考えられる、バス路線、バス停、バスの運行本数などを追加した。

また、交通事故と関係性が高いと思われる物件交通事故の発生件数について、第２回実験で追加した特徴量に加え、予測事故類型の「人対車両(取締り)」「車両相互(取締り)」に、データを加工して作成することとし、

・直近１年の物損事故の発生件数

・直近６ヶ月の物損事故の発生件数

・直近１ヶ月の物損事故の発生件数

を新たに特徴量として追加した。

なお、交通事故抑止に効果があるとされる交通取締りについても同様に、

・直近１ヶ月以内の最寄交通取締り場所までの距離

・直近１ヶ月以内の最寄交通取締りの違反時間帯からの経過時間

・直近１ヶ月以内の１ｋｍ四方の交通取締り件数

を新たに特徴量として追加した。

これは、第２回実験で取締りの件数に関係する特徴量重要度が低かったため、交通事故と取締りでは時空間的な要因があると考えられることから、新しい特徴量として「直近１ヶ月以内の最寄交通取締り場所までの距離」「直近１ヶ月以内の最寄交通取締りの違反時間」「直近１ヶ月以内の１ｋｍ四方の交通取締り件数」を追加したものである。

さらに、第２回実験の結果から予測に際して重要度が低かった特徴量を除外し、相関性が高い特徴量は統合及び除外した。また、予測に重視しているがその理由付けが交通事故に関係していない特徴量についても除外した。

使用した特徴量は、別添２「特徴量一覧」の第３回実証実験を参照。

(ｳ) 取締りまでの距離・時間について

「直近１ヶ月以内の最寄り交通取締り場所までの距離」は、それぞれのメッシュの中央から直近１ヶ月以内の最寄り交通取締り場所までの距離を測定し特徴量とした。

「直近1ヶ月以内の最寄り交通取締り場所の違反時間帯からの経過時間」は、それぞれのメッシュの中央から直近１ヶ月以内の最寄り交通取締り場所の違反時間帯からの経過時間を測定し特徴量とした。

(ｴ) 高速道路の交通量について

高速道路等の交通量については、道路管理者（中日本高速道路株式会社、東日本高速道路株式会社、首都高速道路株式会社）の協力により使用することが可能となった。

高速道路等の交通量は、道路管理者がジャンクション、インターチェンジ、サービスエリア、パーキングエリア等の間に設置されている 315 か所の車両感知器（トラフィックカウンター）から車両通過台数を測定した１時間単位のデータを使用した。

車両感知器の緯度経度については既に特定されていたことから、測定した交通量をデータ化して対象メッシュ内に測定地点（高速）として設定した。

実証実験を行うにあたり、高速道路の測定地点（高速）で測定した車両台数を有効に活用して測定地点以外に交通量の特徴量とすることとした。

(ｵ) 予測対象事故類型

予測対象事故類型については、「人対車両」「人対車両(高齢者)」「人対車両(取締り無し)」「人対車両(取締り)」「車両相互」「車両相互(取締り無し)」「車両相互(取締り)」「高速道路等」を対象とした。

「人対車両(高齢者)」の実験設定は以下のとおりである。

・事故類型　人対車両

・年齢　65 歳以上

・当事者種別　指定なし

・当事者分類　第一当事者

(ｶ) 対象期間、予測対象範囲

対象とした期間、予測対象範囲は第２回実験(本実験)と同一設定とする。

「人対車両(高齢者)」の学習期間は以下のとおりである。

◆　図表２－12　第３回実験学習追加設定

事故類型	学習期間	正例学習件数	負例学習件数	予測評価期間
人対車両(高齢者)	2008 年 1 月～2017 年 6 月	8,879	10,606	2017 年 7 月～2018 年 6 月

イ　実験結果

第３回実験の結果は、次のとおりである。

◆　図表２－13　第３回実証実験結果　人対車両

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
XGBoost	321	7,557,611	107	25,089,961	0.0042%	75.00%	9.96	0.0085%
ディープラーニング	350	14,104,308	78	18,543,264	0.0025%	81.78%	5.90	0.0050%

◆　図表２－14　第３回実証実験結果　人対車両(高齢者)

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
XGBoost	83	4,063,387	29	28,584,501	0.0020%	74.11%	20.13	0.0041%
ディープラーニング	94	13,388,791	18	19,259,097	0.0007%	83.93%	7.51	0.0014%

◆　図表２－15　第３回実証実験結果　人対車両(取締り無し)

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
XGBoost	212	6,255,080	68	18,230,640	0.0020%	74.11%	20.13	0.0041%
ディープラーニング	228	8,526,044	52	15,959,676	0.0007%	83.93%	7.51	0.0014%

◆　図表２－16　第３回実証実験結果　人対車両(取締り)

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
XGBoost	213	6,520,256	67	17,965,464	0.0034%	75.71%	9.09	0.0068%
ディープラーニング	214	8,022,579	66	16,463,141	0.0027%	81.43%	8.21	0.0054%

◆　図表２－17　第３回実証実験結果　車両相互

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
XGBoost	1425	7,159,996	480	25,486,099	0.0199%	74.80%	10.57	0.0398%

◆　図表２－18　第３回実証実験結果　車両相互(取締り無し)

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
XGBoost	879	5,171,224	314	19,313,583	0.0170%	73.68%	10.45	0.0340%

◆　図表２－19　第３回実証実験結果　車両相互(取締り)

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
XGBoost	856	4,966,433	337	19,518,374	0.0172%	71.75%	9.98	0.0345%

◆　図表２－20　第３回実証実験結果　高速道路等

アルゴリズム	TP	FP	FN	TN	適合率	再現率	リスク比	F 値
XGBoost	9	317,797	1	1,388,753	0.0028%	90.00%	39.33	0.0057%
ディープラーニング	8	682,561	2	1,023,989	0.0012%	80.00%	6.00	0.0023%

全体的に見ると、第２回実験と比較して適合率、再現率ともに若干ではあるが低下傾向となってしまった。

また、一般道の各モデルと高速道路等のモデルは特徴量の重要度の順位に大きな違いがあるが、適合率、再現率について実験結果を比較すると、大きな違いはなく、同様の結果となっている。

「人対車両(取締り無し)」「人対車両(取締り)」「車両相互(取締り無し)」「車両相互(取締り)」については、交通取締関係の特徴量を追加したモデルと追加していないモデルでの特異な数値の変化は見られなかった。

機械学習アルゴリズム間の性能比較結果としては、第２回実験と同じく、どの対象事故類型に関してもXGBoostが高い予測精度を示した。

ディープラーニングはパラメーターの調整を行ったが、XGBoostの精度には及ばなかった。

3　特徴量について

特徴量については、第２回実験で用いた特徴量が全体的に見ると高い精度の予測結果を出した。

図表３－２～図表３－10 に示す第２回実験の特徴量重要度を確認すると、全体的な傾向として、時間、天候などの変動性がある特徴量が上位にあり、メッシュ内の施設数のような特徴量はほとんど効果を及ぼしていないことが分かった。

また、メッシュ間の差があまりない特徴量は、メッシュごとの特徴（差異）を説明することができないことが、予測に対する重要度が低くなる理由と考えられる。

(1) 特徴量の重要度の変動について

本調査研究の交通事故発生予測では、特徴量を第１回から第３回の実証実験で図表３－１のように異なる数の特徴量を使用した。そのため、第２回実証実験と第３回実証実験時については、特徴量を追加、統合又は除外した場合に重要度がどのように変化したか、また、予測結果がどのように影響したかを確認した。

◆　図表３－１　特徴量数一覧

実証実験	特徴量の数	増加数	削減数	統合数
第1回実証実験	26	0	0	0
第２回（人対・車両）	191	167	0	0
第３回（人対・車両）	106	23	104	4
第２回（高速道路等）	188	162	0	0
第３回（高速道路等）	80	8	112	4

ア　特徴量の変更等について

(ｱ) 第２回実証実験

業務データやオープンデータにあるものを相関関係に関係なく特徴量として使用した。そのため、ガソリン価格関係（ハイオク、レギュラー、軽油、灯油）や年齢関係（人口、昼間人口）などの相関係数が 0.9 以上であっても、特徴量として使用した。

(ｲ) 第３回実証実験

交通事故と関係のあると考えられる「最寄バス停までの距離」「バス路線　平日運行本数」「バス路線　土曜日運行本数」「バス路線　日・祝日運行本数」「交通量（高速道）」を追加した。

また、年齢関係を「男 20～24 歳人口」のような５歳間隔を「男 20～29 歳人口」の 10 歳間隔に変更し追加した。

イ　第２回と第３回の実証実験の特徴量重要度比較

第２回と第３回の実証実験の特徴量重要度を比較した。

比較方法については、第３回実証実験の特徴量重要度の上位 30 位を基に、第２回実証実験時の特徴量重要度の順位と重要度を比較した。

第３回実験の特徴量の重要度を見ると、一般道のモデルは時間帯、交通量（一般道）や交通事故と各施設等の距離を考慮した特徴量が予測に高く貢献しており、高速道路等のモデルでは交通量（高速道）、サービスエリアまでの距離等、高速道路等に関係の強いと思われるものが高く貢献しているという結果となった。

特徴量の重要度の第２回実験と第３回実験の上位 30 位の順位では、一般道の各モデルで時間帯は上位のままであったが、高速道路等のモデルでは第２回実験

の6位から第3回実験の15位と順位を下げ、変わりに高速道路の交通量等が時間帯より大きく寄与する結果となった。

これは、神奈川県内全体を予測している一般道の各モデルでは、時間帯が地域性や季節性に関係して寄与度が高いと考えられるが、高速道路等のモデルでは路線的な視点で予測している点が強いことから交通量（高速道）が時間帯よりも寄与度が高かったと考えられる。

比較した結果については、次のとおりである。

◆　図表３－２　人対車両　特徴量重要度ランキング

No	項目名	第３回 ランキング	第３回 特徴量重要度	第２回 ランキング	第２回 特徴量重要度
1	時間帯	1	0.0382475	1	0.0351468
2	気温	2	0.0226980	2	0.0157224
3	交通量(一般道)	3	0.0183885	3	0.0140586
4	ガソリン価格_レギュラー	4	0.0182842	142	0.0023190
5	最寄り駅までの距離	5	0.0177470	12	0.0111375
6	最寄交差点までの距離	6	0.0174858	8	0.0120860
7	最寄バス停までの距離	7	0.0174296	第2回未入力	第2回未入力
8	昼間人口	8	0.0162338	48	0.0076267
9	最寄深夜飲食店までの距離	9	0.0161091	6	0.0124114
10	道路面積比率	10	0.0156500	4	0.0136366
11	湿度	11	0.0148902	10	0.0113294
12	地区容積率	12	0.0147969	18	0.0101119
13	最寄金融機関までの距離	13	0.0145739	9	0.0114419
14	最寄交番・派出所・ 駐在所までの距離	14	0.0145739	27	0.0089692
15	最寄小学校までの距離	15	0.0143145	17	0.0101838
16	最寄駐車場までの距離	16	0.0141211	13	0.0109932
17	最寄コンビニまでの距離	17	0.0139990	11	0.0112998
18	地価	18	0.0138358	22	0.0092276
19	最寄病院までの距離	19	0.0137334	25	0.0090609
20	最寄公園までの距離	20	0.0137209	16	0.0102525
21	最寄ゴルフ場までの距離	21	0.0133357	32	0.0087896
22	商業施設面積比率	22	0.0130906	5	0.0125270
23	最寄貴金属店までの距離	23	0.0129076	26	0.0090114
24	最寄特別支援学校までの距離	24	0.0128847	49	0.0076034
25	昼夜間人口比率	25	0.0128065	76	0.0061493
26	全産業従業員数	26	0.0127550	35	0.0085160
27	最寄スーパーまでの距離	27	0.0126900	24	0.0091468
28	最寄公衆浴場までの距離	28	0.0125936	54	0.0073787
29	最寄幼稚園までの距離	29	0.0125761	43	0.0079066
30	最寄中学・高校までの距離	30	0.0124982	30	0.0088538

◆　図表３－３　人対車両(第３回でランク外となったもの)

No	項目名	第３回 ランキング	第３回 特徴量重要度	第２回 ランキング	第２回 特徴量重要度
1	月	43	0.011247061	20	0.00976187
2	最寄神社までの距離	47	0.0110029	29	0.0088656
3	直近１年での全事象発生件数	54	0.0106367	7	0.0121043
4	最寄ドラッグストアまでの距離	57	0.0103947	19	0.0097920
5	住宅地面積比率	103	0.0009348	15	0.0102739

◆　図表３－４　人対車両(第３回で除外したもの)

No	項目名	第２回ランキング	第２回特徴量重要度	理由
1	ガソリン価格_灯油店頭	14	0.0105558	ガソリン価格は レギュラーのみ とした
2	ガソリン価格_灯油配達	23	0.0091844	
3	ガソリン価格_軽油	28	0.0089334	

◆ 図表３－５　車両相互　特徴量重要度ランキング

No	項目名	第３回 ランキング	第３回 特徴量重要度	第２回 ランキング	第２回 特徴量重要度
1	時間帯	1	0.0434605	3	0.0281017
2	ガソリン価格_レギュラー	2	0.0373671	134	0.0029560
3	道路面積比率	3	0.0320531	2	0.0346409
4	交通量(一般道)	4	0.0299801	1	0.0349389
5	最寄交差点までの距離	5	0.0290565	4	0.0217088
6	気温	6	0.0187506	7	0.0175524
7	最寄バス停までの距離	7	0.0185284	第2回未入力	第2回未入力
8	湿度	8	0.0178470	8	0.0164023
9	最寄サービスエリアまでの距離	9	0.0163591	14	0.0108633
10	最寄ジャンクションまでの距離	10	0.0153752	17	0.0104149
11	昼間人口	11	0.0151880	91	0.0046400
12	直近１年での全事象発生件数	12	0.0150903	9	0.0158641
13	最寄ラブホテルまでの距離	13	0.0145638	22	0.0092596
14	商業施設面積比率	14	0.0143969	11	0.0123624
15	最寄コンビニまでの距離	15	0.0130763	13	0.0112170
16	最寄大学までの距離	16	0.0129994	28	0.0082941
17	最寄ガソリンスタンドまでの距離	17	0.0129007	15	0.0107967
18	最寄特別支援学校までの距離	18	0.0128803	30	0.0082139
19	最寄公衆浴場までの距離	19	0.0126314	36	0.0078166
20	最寄深夜飲食店までの距離	20	0.0124817	27	0.0084667
21	最寄ゲームセンターまでの距離	21	0.0124722	43	0.0071758
22	最寄地下通路までの距離	22	0.0124089	34	0.0079428
23	最寄ゴルフ場までの距離	23	0.0122133	24	0.0088722
24	最寄マンガ喫茶までの距離	24	0.0121781	26	0.0086861
25	最寄中学・高校までの距離	25	0.0121633	20	0.0093337
26	最寄公園までの距離	26	0.0121607	25	0.0087347
27	最寄駐車場までの距離	27	0.0119341	41	0.0074276
28	地価	28	0.0118685	37	0.0078022
29	最寄コインランドリーまでの距離	29	0.0116759	32	0.0080920
30	最寄パチンコまでの距離	30	0.0115901	42	0.0073633

◆ 図表３－６　車両相互（第３回でランク外になったもの）

No	項目名	第３回 ランキング	第３回 特徴量重要度	第２回 ランキング	第２回 特徴量重要度
1	月	32	0.0113590	21	0.0093180
2	直近１年での 予測対象事象発生件数	38	0.0105613	18	0.0102867
3	風速	41	0.0103886	16	0.0106068
4	全産業従業員数	43	0.0103483	29	0.0082258
5	曜日	62	0.0086548	23	0.0091978
6	住宅地面積比率	104	0.0007018	12	0.0113736

◆ 図表３－７　車両相互（第３回で除外したもの）

No	項目名	第２回ランキング	第２回特徴量重要度	理由
1	ガソリン価格_灯油店頭	5	0.0217059	ガソリン価格は レギュラーのみ とした
2	ガソリン価格_軽油	6	0.0200409	
3	ガソリン価格_灯油配達	10	0.0155683	
4	ガソリン価格_ハイオク	19	0.0102495	

◆ 図表３－８　高速道路等　特徴量重要度ランキング

No	項目名	第３回 ランキング	第３回 特徴量重要度	第２回 ランキング	第２回 特徴量重要度
1	最寄ジャンクションまでの距離	1	0.0361579	11	0.0153493
2	交通量(高速道)	2	0.0359531	第2回未入力	第2回未入力
3	道路面積比率	3	0.0325679	5	0.0260086
4	ガソリン価格_レギュラー	4	0.0304672	70	0.0052727
5	湿度	5	0.0286289	2	0.0348920
6	最寄サービスエリアまでの距離	6	0.0254074	13	0.0148956
7	気温	7	0.0242215	1	0.0349267
8	直近１年での予測対象事象発生件数	8	0.0206066	22	0.0095000
9	最寄り駅の乗降客数	9	0.0193998	54	0.0066683
10	最寄警察署までの距離	10	0.0189896	18	0.0104971
11	地価	11	0.0189791	15	0.0129974
12	風速	12	0.0188414	3	0.0318782
13	交通量(一般道)	13	0.0188286	4	0.0285982
14	バス路線_平日運行本数	14	0.0186445	第2回未入力	第2回未入力
15	時間帯	15	0.0186240	6	0.0187674
16	最寄り駅までの距離	16	0.0185468	17	0.0110150
17	最寄バス停までの距離	17	0.0179787	第2回未入力	第2回未入力
18	最寄公共施設までの距離	18	0.0176038	26	0.0087974
19	最寄ゴルフ場までの距離	19	0.0174519	34	0.0076567
20	最寄イベント施設までの距離	20	0.0170076	53	0.0067563
21	バス路線_日・祝日運行本数	21	0.0167680	第2回未入力	第2回未入力
22	最寄協同組合までの距離	22	0.0166815	31	0.0078479
23	最寄特別支援学校までの距離	23	0.0148982	47	0.0071590
24	最寄宿泊施設までの距離	24	0.0147955	23	0.0093896
25	最寄公園までの距離	25	0.0145651	19	0.0104382
26	人口増加率	26	0.0143669	29	0.0079974
27	昼夜間人口比率	27	0.0139228	67	0.0055508
28	最寄コインランドリーまでの距離	28	0.0139125	58	0.0064486
29	最寄交差点までの距離	29	0.0133647	16	0.0116315
30	最寄中学・高校までの距離	30	0.0133486	30	0.0078974

◆　図表３－９　高速道路等（第３回でランク外になったもの）

No	項目名	第３回 ランキング	第３回 特徴量重要度	第２回 ランキング	第２回 特徴量重要度
1	最寄駐輪場までの距離	31	0.0132639	25	0.0093299
2	最寄ガソリンスタンドまでの距離	37	0.0120778	21	0.0095980
3	最寄空き地までの距離	42	0.0110510	27	0.0086741
4	曜日	46	0.0106320	12	0.0148963
5	最寄駐車場までの距離	59	0.0086576	24	0.0093625
6	月	64	0.0073504	14	0.0142349

◆　図表３－10　高速道路等（第３回で除外したもの）

No	項目名	第２回ランキング	第２回特徴量重要度	理由
1	ガソリン価格_灯油店頭	7	0.0187567	ガソリン価格はレギュラーのみとした
2	ガソリン価格_ハイオク	8	0.0184671	
3	ガソリン価格_軽油	9	0.0180138	
4	ガソリン価格_灯油配達	10	0.0171569	
5	最寄自治会までの距離	28	0.0084316	廃止

特徴量については、第２回実験で用いた特徴量を学習データとしたモデルが全体的に高精度の予測結果を出した。

これは、第２回実験の特徴量には、既存研究による知見が得られておらず、現時点では人間が理由を説明することが難しい特徴量であっても、機械学習による学習処理において何らかの相関関係を見いだし、予測の根拠としていたことを示す結果となった。

第３回実験では、一部特徴量を追加したが、人工知能が予測に際して重要度が低かった特徴量を除外し、相関性が高い特徴量は統合及び除外したため、予測結果が第２回実証実験より若干低下する結果となった。

また、特徴量の重要度の順位や数値に変化があったが、「時間帯」「気温」「交通量」「道路面積比率」の特徴量については、第２回実証実験と同様に予測に対して高い貢献度を示す結果となった。

また、ガソリン価格関係については、ガソリン価格関係（ハイオク、軽油、灯油）を「ガソリン価格　レギュラー」を選択した結果、人対車両では第２回 142 位から第３回４位に、車両相互では第２回 134 位から第３回２位に、高速道路等では第２回 70 位から第３回４位に、いずれのモデルも大幅に順位を上げる結果となった。

第２回実証実験と第３回実証実験の結果、例えば「最寄自治会までの距離」等のような交通事故の発生との関係性や特徴量とする理由が見られない特徴量であっても、予測精度が向上する可能性は否定できない。

また、それらの特徴量の重要度が高い場合にそのまま交通事故の発生要因であると誤った判断を生む危険性があり、これは人工知能が因果関係を見いだすことはできないという弱点があるためである。

説明性や分かりやすさを最重視し、予測精度をある程度犠牲にしてでも因果関係が確認できる特徴量のみを使用するか、「疑似相関」や「逆の因果関係」の可能性があることを前提とした上で予測精度を重視してより多くの特徴量を使用するかは、特徴量の内容を吟味して最終的には人間が判断する必要がある。

いずれにしても、大量のデータの中から、今まで人間が考えもしなかった要因を示唆してくれるということは、予測に基づいて交通指導取締りなどの街頭活動を行うというような警察活動による対策だけでなく、地域と連携し、効果的な交通安全教育や根本的な交通規制の見直し等の交通事故防止３Ｅのプロセスにおいても、適切かつ有効に活用することで大きく役立つものとなるであろう。

ウ　学習データの特徴量分布について

本調査研究の特徴量には、「曜日」「時間帯」等のカテゴリ値と「温度」「最寄○○までの距離」等の連続値がある。交通事故分析のためそれぞれの特徴量ごとの発生・未発生数の分布を算出した。

学習データの特徴量分布については、第３回の実証実験の学習データ（アンダーサンプリング済み）を使用し、事故類型については、人対車両、人対車両(高齢者)、車両相互、高速道路等とした。

例えば発生が未発生を上回る、事故類型「人対車両」、特徴量が「最寄コンビニまでの距離」、数値が「０～100m」、事故発生数が31%、未発生数９%であった場合、交通事故の発生は未発生の約3.4倍であることから、コンビニ付近（０～100m）での人対車両事故の発生確率は約3.4倍高いと考えられる。

逆に、発生より未発生が上回る数値、例えば事故類型「人対車両」、特徴量が「最寄交差点までの距離」、数値が「400～500m」、事故発生数 2.3%、未発生数 6.2%であった場合、事故の発生は未発生の約 1/3 であることを考えれば、発生確率が低いと考えられる。

特徴量の分布については、次のとおりである。

◆ 図表３－11　特徴量分布　人対車両

時間帯(時)

	0～3	3～6	6～9	9～12	12～15	15～18	18～21	21～24
発生	3.12%	2.19%	13.66%	15.43%	14.98%	23.05%	19.61%	7.96%
未発生	11.69%	11.05%	12.74%	12.83%	12.92%	13.13%	12.79%	12.85%

曜日(曜日)

	月	火	水	木	金	土	日
発生	14.62%	15.29%	15.28%	15.35%	16.48%	13.10%	9.88%
未発生	14.25%	14.25%	14.19%	14.40%	14.48%	14.43%	14.00%

月(月)

	1	2	3	4	5	6
発生	8.87%	8.31%	9.44%	8.64%	8.25%	7.77%
未発生	8.87%	8.33%	9.02%	8.67%	8.72%	8.75%
	7	8	9	10	11	12
発生	6.96%	6.08%	7.16%	8.82%	9.24%	10.46%
未発生	8.10%	7.95%	7.83%	8.00%	7.66%	8.09%

気温(度)

	-5 未満	-5～0	0～5	5～10	10～15	15～20	20～25	25～30	30～35	35～40	40 以上
発生	0.0%	0.4%	8.1%	20.0%	17.7%	19.1%	19.0%	12.7%	3.0%	0.0%	0.0%
未発生	0.0%	1.3%	9.9%	17.8%	16.1%	19.1%	19.8%	13.1%	2.8%	0.0%	0.0%

ガソリン価格_レギュラー(円)

	100 未満	100～105	105～110	110～115	115～120	120～125	125～130	130～135	135～140
発生	0.00%	0.00%	3.36%	3.62%	4.18%	7.07%	12.01%	12.47%	8.47%
未発生	0.00%	0.00%	2.96%	3.45%	4.95%	7.14%	11.31%	12.49%	8.63%
	140～145	145～150	150～155	155～160	160～165	165～170	170～175	175～180	180 以上
発生	9.20%	10.97%	10.36%	8.25%	3.30%	3.66%	1.11%	0.82%	1.14%
未発生	8.72%	11.08%	9.85%	8.27%	3.74%	4.35%	1.05%	0.83%	1.19%

道路面積比率(%)

	0～10	10～20	20～30	30～40	40～50	50～60	60～70	70～80	80～90	90～100
発生	4.60%	34.95%	36.69%	16.38%	5.06%	1.76%	0.39%	0.13%	0.04%	0.00%
未発生	28.95%	44.78%	18.86%	5.04%	1.51%	0.57%	0.16%	0.07%	0.04%	0.02%

商業施設面積比率(%)

	0～10	10～20	20～30	30～40	40～50	50～60	60～70	70～80	80～90	90～100
発生	72.33%	13.84%	6.88%	3.31%	1.66%	0.84%	0.60%	0.22%	0.17%	0.14%
未発生	93.10%	3.69%	1.54%	0.74%	0.38%	0.26%	0.13%	0.07%	0.05%	0.05%

交通事故 人対車両 10年 :未発生 :発生

単位:m	0-100	100-200	200-300	300-400	400-500	500-1000	1000-2000	2000以上
最寄コンビニまでの距離_差分(%)	31.27	27.96	17.75	10.43	6.05	6.02	0.45	0.07
最寄コンビニまでの距離_差分(%)	9.05	18.67	19.43	15.75	11.14	18.12	5.75	2.08
最寄駐輪場までの距離_差分(%)	47.88	29.53	11.41	5.12	2.43	2.84	0.57	0.22
最寄駐輪場までの距離_差分(%)	22.35	26.02	16.54	9.76	5.87	11.93	4.72	2.84
最寄風俗営業所までの距離_差分(%)	6.40	6.52	6.23	6.06	5.63	24.16	32.05	12.94
最寄風俗営業所までの距離_差分(%)	0.94	1.97	2.84	3.47	3.83	22.24	37.45	27.26
最寄イベント施設までの距離_差分(%)	7.31	14.15	13.63	13.08	11.80	32.52	7.20	0.31
最寄イベント施設までの距離_差分(%)	2.71	6.60	9.05	10.38	10.82	40.54	16.45	3.45
最寄カラオケボックスまでの距離_差分(%)	3.93	5.55	6.45	7.07	6.94	27.87	30.36	11.82
最寄カラオケボックスまでの距離_差分(%)	0.85	1.85	2.47	3.21	3.77	23.11	36.18	28.56
最寄ゲームセンターまでの距離_差分(%)	5.53	7.01	6.69	6.48	5.44	24.71	31.22	12.92
最寄ゲームセンターまでの距離_差分(%)	1.05	2.21	2.84	3.44	4.12	22.42	35.78	28.15
最寄ゴルフ場までの距離_差分(%)	0.16	0.50	0.97	1.39	1.53	10.68	32.40	52.37
最寄ゴルフ場までの距離_差分(%)	0.14	0.52	0.84	1.23	1.52	11.55	31.53	52.68
最寄スーパーまでの距離_差分(%)	21.18	23.94	16.46	12.18	8.69	15.00	2.06	0.47
最寄スーパーまでの距離_差分(%)	5.91	12.03	13.62	13.26	11.30	27.81	11.23	4.84
最寄ドラッグストアまでの距離_差分(%)	10.30	15.27	13.77	12.59	10.36	28.00	8.64	1.06
最寄ドラッグストアまでの距離_差分(%)	2.60	6.05	8.34	9.52	9.55	35.34	19.87	8.73
最寄パチンコまでの距離_差分(%)	14.65	13.16	10.39	8.87	7.47	25.77	15.87	3.83
最寄パチンコまでの距離_差分(%)	2.85	5.23	6.43	7.17	7.29	31.15	26.75	13.13
最寄ホームセンターまでの距離_差分(%)	2.75	5.58	6.15	6.32	5.93	31.73	31.84	9.71
最寄ホームセンターまでの距離_差分(%)	0.90	2.43	3.55	4.25	4.65	26.65	36.27	21.30
最寄マンガ喫茶までの距離_差分(%)	3.64	4.80	4.55	3.97	3.86	16.48	29.91	32.79
最寄マンガ喫茶までの距離_差分(%)	0.46	0.99	1.51	1.81	2.20	13.40	29.32	50.30
最寄レンタルビデオ店までの距離_差分(%)	3.13	5.53	6.14	6.31	5.99	27.68	32.70	12.51
最寄レンタルビデオ店までの距離_差分(%)	0.63	1.60	2.60	3.25	3.69	23.91	37.67	26.65
最寄宿泊施設までの距離_差分(%)	3.95	5.66	4.66	4.05	4.12	18.80	33.57	25.19
最寄宿泊施設までの距離_差分(%)	0.95	1.94	2.41	2.67	3.25	18.31	36.25	34.22
最寄飲食店までの距離_差分(%)	35.18	24.53	14.65	9.41	6.23	8.74	1.14	0.11
最寄飲食店までの距離_差分(%)	10.84	17.14	16.39	13.58	10.24	21.30	8.03	2.48
最寄家電量販店までの距離_差分(%)	2.66	4.40	5.02	5.18	5.36	26.68	33.90	16.79
最寄家電量販店までの距離_差分(%)	0.62	1.46	2.33	2.96	3.40	21.28	36.53	31.43
最寄貴金属店までの距離_差分(%)	4.01	5.88	5.56	5.86	5.18	21.57	31.31	20.62
最寄貴金属店までの距離_差分(%)	0.60	1.35	2.10	2.54	2.86	18.13	35.56	36.86
最寄ガソリンスタンドまでの距離_差分(%)	4.61	8.10	10.07	11.26	11.32	41.12	12.53	0.99
最寄ガソリンスタンドまでの距離_差分(%)	1.95	5.16	7.33	9.10	9.40	40.19	21.69	5.17
最寄金融機関までの距離_差分(%)	22.62	19.93	15.57	12.60	9.43	17.38	2.32	0.15
最寄金融機関までの距離_差分(%)	5.50	10.67	12.66	13.20	11.89	31.38	11.60	3.10
最寄空き地までの距離_差分(%)	3.04	6.85	9.72	10.76	10.76	39.93	18.40	0.54
最寄空き地までの距離_差分(%)	1.74	4.94	7.60	9.18	9.85	40.46	22.23	4.00
最寄古物店までの距離_差分(%)	9.07	12.43	10.51	9.36	8.36	29.41	16.79	4.08
最寄古物店までの距離_差分(%)	2.02	4.53	5.81	7.07	7.32	32.21	26.40	14.64
最寄深夜飲食店までの距離_差分(%)	35.77	19.29	12.05	8.86	6.51	13.90	3.24	0.37
最寄深夜飲食店までの距離_差分(%)	10.61	13.69	12.97	11.53	9.36	25.81	12.23	3.80
最寄神社までの距離_差分(%)	2.85	8.52	11.90	12.59	12.07	39.17	12.58	0.32
最寄神社までの距離_差分(%)	2.16	5.86	8.30	9.72	10.36	39.17	20.61	3.81
最寄地下通路までの距離_差分(%)	1.27	2.01	2.29	2.18	1.99	9.89	23.51	56.87
最寄地下通路までの距離_差分(%)	0.19	0.41	0.66	0.80	0.93	6.84	20.10	70.07

単位m	0-100	100-200	200-300	300-400	400-500	500-1000	1000-2000	2000以上
最寄駐車場までの距離_差分(%)	66.35	25.32	5.08	1.68	0.73	0.68	0.14	0.02
最寄駐車場までの距離_差分(%)	40.17	30.71	11.69	5.27	2.91	6.46	2.31	0.50
最寄協同組合までの距離_差分(%)	0.90	1.94	2.49	3.02	3.08	17.09	36.41	35.07
最寄協同組合までの距離_差分(%)	0.33	0.98	1.54	2.00	2.50	17.25	37.87	37.55
最寄病院までの距離_差分(%)	15.10	21.47	16.21	12.25	8.68	20.61	4.79	0.88
最寄病院までの距離_差分(%)	4.40	9.93	11.44	11.49	10.48	30.72	14.27	7.28
最寄ラブホテルまでの距離_差分(%)	1.78	2.79	3.27	3.26	3.15	15.31	29.76	40.68
最寄ラブホテルまでの距離_差分(%)	0.50	1.15	1.57	1.75	2.12	12.77	28.63	51.51
最寄郵便局までの距離_差分(%)	3.61	8.37	9.56	10.00	9.06	33.02	21.77	4.60
最寄郵便局までの距離_差分(%)	0.99	2.91	4.63	5.68	6.36	31.57	33.43	14.44
最寄コインランドリーまでの距離_差分(%)	2.58	4.65	5.88	7.06	7.48	31.40	31.25	9.70
最寄コインランドリーまでの距離_差分(%)	0.80	2.21	3.37	4.51	5.18	27.35	36.44	20.13
最寄公衆浴場までの距離_差分(%)	1.30	3.16	4.05	4.67	4.71	22.16	35.66	24.29
最寄公衆浴場までの距離_差分(%)	0.49	1.28	1.98	2.51	3.09	17.71	35.18	37.78
最寄商店までの距離_差分(%)	41.01	26.93	14.32	7.60	4.27	5.19	0.60	0.06
最寄商店までの距離_差分(%)	14.79	21.69	18.27	13.18	8.47	15.43	6.19	1.99
最寄幼稚園までの距離_差分(%)	8.14	20.64	22.29	17.56	11.76	17.61	1.62	0.38
最寄幼稚園までの距離_差分(%)	4.61	13.05	16.01	15.06	12.26	26.62	8.28	4.10
最寄公共施設までの距離_差分(%)	1.37	3.65	4.53	5.30	5.50	23.48	32.53	23.65
最寄公共施設までの距離_差分(%)	0.59	1.51	2.21	2.93	3.50	20.09	36.34	32.83
最寄社交飲食店までの距離_差分(%)	10.45	8.69	6.84	5.76	5.18	22.82	26.01	14.24
最寄社交飲食店までの距離_差分(%)	1.64	2.76	3.25	3.64	4.12	22.42	34.19	27.98
最寄ジャンクションまでの距離_差分(%)	0.13	0.34	0.50	0.50	0.51	3.03	7.67	87.32
最寄ジャンクションまでの距離_差分(%)	0.09	0.25	0.30	0.30	0.38	1.96	5.66	91.06
最寄サービスエリアまでの距離_差分(%)	0.01	0.04	0.07	0.12	0.13	1.21	4.57	93.85
最寄サービスエリアまでの距離_差分(%)	0.12	0.19	0.23	0.24	0.23	1.54	5.08	92.36
最寄交差点までの距離_差分(%)	54.25	25.12	11.05	5.31	2.28	1.80	0.15	0.03
最寄交差点までの距離_差分(%)	19.57	30.37	20.47	11.79	6.21	8.18	2.45	0.96
最寄り駅までの距離_差分(%)	4.76	9.77	9.49	8.87	7.77	27.00	22.83	9.51
最寄り駅までの距離_差分(%)	0.97	2.81	4.10	5.33	5.92	28.61	31.60	20.66
最寄警察署までの距離_差分(%)	0.58	1.24	1.61	2.12	2.50	15.33	32.17	44.46
最寄警察署までの距離_差分(%)	0.14	0.40	0.69	0.92	1.14	9.13	27.65	59.94
最寄交番・派出所・駐在所までの距離_差分(%)	6.96	11.27	12.47	12.23	11.40	35.29	10.08	0.30
最寄交番・派出所・駐在所までの距離_差分(%)	1.74	4.58	6.93	8.80	9.88	42.17	22.77	3.13
最寄小学校までの距離_差分(%)	1.62	8.46	13.01	16.80	16.63	39.94	3.36	0.19
最寄小学校までの距離_差分(%)	1.49	6.29	10.08	12.61	13.12	40.22	13.20	2.99
最寄中学・高校までの距離_差分(%)	0.82	4.70	8.80	11.07	12.83	49.57	11.71	0.49
最寄中学・高校までの距離_差分(%)	1.09	4.11	6.79	8.70	9.83	43.29	21.11	5.08
最寄大学までの距離_差分(%)	0.19	0.53	0.99	1.50	1.74	11.82	26.10	57.14
最寄大学までの距離_差分(%)	0.16	0.38	0.67	0.88	1.20	8.67	22.85	65.18
最寄特別支援学校までの距離_差分(%)	0.08	0.23	0.53	1.10	1.35	10.50	27.93	58.28
最寄特別支援学校までの距離_差分(%)	0.11	0.28	0.54	0.80	0.97	7.87	22.55	66.89
最寄公園までの距離_差分(%)	14.25	33.70	26.08	13.57	5.57	5.92	0.68	0.22
最寄公園までの距離_差分(%)	12.95	25.40	19.52	11.76	7.39	14.48	5.27	3.23
最寄バス停までの距離_差分(%)	44.16	35.10	12.74	4.47	1.75	1.66	0.11	0.00
最寄バス停までの距離_差分(%)	19.84	35.09	20.91	10.19	4.67	6.78	2.21	0.30

◆ 図表３－12　特徴量分布　車両相互

時間帯(時)

	0～3	3～6	6～9	9～12	12～15	15～18	18～21	21～24
発生	2.39%	2.20%	18.12%	17.57%	17.63%	21.38%	14.78%	5.94%
未発生	12.54%	12.43%	12.49%	12.51%	12.53%	12.42%	12.56%	12.51%

曜日(曜日)

	月	火	水	木	金	土	日
発生	14.48%	14.94%	14.51%	14.59%	15.95%	14.30%	11.23%
未発生	14.22%	14.34%	14.27%	14.37%	14.28%	14.32%	14.21%

月(月)

	1	2	3	4	5	6
発生	8.02%	7.72%	8.92%	8.31%	8.47%	8.46%
未発生	8.89%	8.17%	8.96%	8.63%	9.00%	8.66%
	7	8	9	10	11	12
発生	8.73%	8.27%	8.00%	8.19%	8.26%	8.66%
未発生	7.97%	8.08%	7.77%	8.02%	7.86%	7.99%

気温(度)

	-5 未満	-5～0	0～5	5～10	10～15	15～20	20～25	25～30	30～35	35～40	40 以上
発生	0.0%	0.5%	6.9%	17.6%	16.5%	18.5%	20.1%	15.5%	4.4%	0.1%	0.0%
未発生	0.0%	1.3%	10.1%	17.7%	16.2%	19.1%	19.7%	13.0%	2.6%	0.0%	0.0%

ガソリン価格_レギュラー(円)

	100 未満	100～105	105～110	110～115	115～120	120～125	125～130	130～135	135～140
発生	0.00%	0.00%	2.96%	3.32%	4.30%	6.78%	11.69%	12.95%	8.59%
未発生	0.00%	0.00%	3.05%	3.39%	4.97%	7.25%	11.42%	12.34%	8.78%
	140～145	145～150	150～155	155～160	160～165	165～170	170～175	175～180	180 以上
発生	8.78%	11.38%	10.21%	7.84%	3.35%	3.97%	1.33%	1.01%	1.53%
未発生	8.72%	10.92%	9.68%	8.34%	3.71%	4.40%	1.02%	0.81%	1.20%

道路面積比率(%)

	0～10	10～20	20～30	30～40	40～50	50～60	60～70	70～80	80～90	90～100
発生	3.09%	30.06%	36.44%	19.18%	7.18%	2.95%	0.88%	0.18%	0.03%	0.01%
未発生	28.89%	44.80%	19.05%	4.84%	1.57%	0.57%	0.17%	0.06%	0.03%	0.02%

商業施設面積比率(%)

	0～10	10～20	20～30	30～40	40～50	50～60	60～70	70～80	80～90	90～100
発生	74.31%	13.54%	6.10%	3.11%	1.52%	0.71%	0.38%	0.19%	0.10%	0.05%
未発生	93.23%	3.65%	1.48%	0.74%	0.37%	0.20%	0.13%	0.07%	0.07%	0.05%

交通事故 車両相互 10年 :未発生 :発生

単位:m	0-100	100-200	200-300	300-400	400-500	500-1000	1000-2000	2000以上
最寄コンビニまでの距離_差分(%)	25.32	27.54	20.01	11.79	6.97	7.47	0.70	0.20
最寄コンビニまでの距離_差分(%)	8.93	18.68	19.43	15.80	11.00	18.28	5.75	2.13
最寄駐輪場までの距離_差分(%)	40.86	31.00	13.18	6.02	3.28	4.35	0.87	0.44
最寄駐輪場までの距離_差分(%)	22.29	25.98	16.48	9.76	5.89	12.09	4.63	2.88
最寄風俗営業所までの距離_差分(%)	3.52	4.94	5.30	5.72	5.72	25.81	34.00	14.99
最寄風俗営業所までの距離_差分(%)	0.98	1.97	2.73	3.31	3.84	22.51	37.16	27.51
最寄イベント施設までの距離_差分(%)	5.14	11.51	12.82	13.20	12.10	36.08	8.59	0.56
最寄イベント施設までの距離_差分(%)	2.62	6.56	9.03	10.39	10.84	40.67	16.51	3.38
最寄カラオケボックスまでの距離_差分(%)	2.76	4.86	5.12	5.67	5.81	28.29	33.43	14.07
最寄カラオケボックスまでの距離_差分(%)	0.83	1.84	2.60	3.25	3.82	22.75	36.00	28.91
最寄ゲームセンターまでの距離_差分(%)	3.19	5.61	6.07	5.99	5.91	27.19	31.99	14.05
最寄ゲームセンターまでの距離_差分(%)	1.00	2.17	2.93	3.49	3.92	22.40	35.87	28.20
最寄ゴルフ場までの距離_差分(%)	0.15	0.49	0.92	1.39	1.51	11.40	32.62	51.52
最寄ゴルフ場までの距離_差分(%)	0.15	0.50	0.85	1.21	1.45	11.46	31.73	52.66
最寄スーパーまでの距離_差分(%)	14.43	21.23	16.80	13.38	9.95	19.82	3.57	0.81
最寄スーパーまでの距離_差分(%)	5.92	12.20	13.62	13.11	11.22	28.04	11.09	4.81
最寄ドラッグストアまでの距離_差分(%)	6.05	11.43	12.49	12.49	10.76	33.33	11.68	1.77
最寄ドラッグストアまでの距離_差分(%)	2.51	6.09	8.49	9.54	9.67	35.09	19.91	8.71
最寄パチンコまでの距離_差分(%)	8.73	11.44	10.00	9.27	8.37	29.60	17.90	4.69
最寄パチンコまでの距離_差分(%)	2.84	5.31	6.48	7.04	7.22	31.03	26.85	13.24
最寄ホームセンターまでの距離_差分(%)	2.27	5.25	6.06	6.00	6.24	31.76	32.15	10.26
最寄ホームセンターまでの距離_差分(%)	0.92	2.44	3.46	4.19	4.70	26.60	36.39	21.30
最寄マンガ喫茶までの距離_差分(%)	1.83	3.08	3.52	3.62	3.75	17.70	31.49	35.01
最寄マンガ喫茶までの距離_差分(%)	0.50	1.03	1.46	1.79	2.17	13.27	29.37	50.40
最寄レンタルビデオ店までの距離_差分(%)	2.15	4.03	4.75	5.60	5.58	28.83	34.76	14.30
最寄レンタルビデオ店までの距離_差分(%)	0.59	1.69	2.55	3.26	3.72	23.53	38.03	26.62
最寄宿泊施設までの距離_差分(%)	2.78	4.79	3.97	3.88	4.00	20.62	34.32	25.64
最寄宿泊施設までの距離_差分(%)	0.94	1.93	2.32	2.74	3.09	18.64	36.15	34.18
最寄飲食店までの距離_差分(%)	29.25	25.85	15.72	10.05	6.79	10.37	1.67	0.29
最寄飲食店までの距離_差分(%)	10.80	17.07	16.40	13.45	10.33	21.49	8.01	2.45
最寄家電量販店までの距離_差分(%)	1.97	3.86	4.71	5.03	5.32	26.65	34.54	17.93
最寄家電量販店までの距離_差分(%)	0.61	1.54	2.32	2.90	3.42	21.28	36.65	31.29
最寄貴金属店までの距離_差分(%)	1.97	3.48	4.26	4.92	4.78	23.09	34.09	23.40
最寄貴金属店までの距離_差分(%)	0.57	1.41	1.97	2.50	2.90	18.02	35.67	36.95
最寄ガソリンスタンドまでの距離_差分(%)	6.53	9.64	10.48	10.93	10.87	37.57	12.82	1.16
最寄ガソリンスタンドまでの距離_差分(%)	1.96	5.16	7.38	8.82	9.70	40.08	21.66	5.23
最寄金融機関までの距離_差分(%)	13.95	17.43	15.73	13.50	10.83	23.80	4.42	0.34
最寄金融機関までの距離_差分(%)	5.46	10.68	12.89	13.11	11.81	31.33	11.67	3.06
最寄空き地までの距離_差分(%)	2.87	6.77	9.33	10.42	10.67	40.86	18.32	0.76
最寄空き地までの距離_差分(%)	1.84	5.08	7.54	9.09	9.81	40.65	22.09	3.91
最寄古物店までの距離_差分(%)	6.45	10.28	9.74	9.94	9.06	31.23	18.68	4.62
最寄古物店までの距離_差分(%)	1.95	4.60	6.01	6.84	7.34	32.33	26.30	14.63
最寄深夜飲食店までの距離_差分(%)	26.18	19.62	13.35	10.63	7.68	17.20	4.66	0.67
最寄深夜飲食店までの距離_差分(%)	10.59	13.84	12.93	11.43	9.61	25.85	12.03	3.73
最寄神社までの距離_差分(%)	2.50	7.74	10.80	12.11	11.67	40.53	14.09	0.55
最寄神社までの距離_差分(%)	2.10	5.78	8.20	9.66	10.11	39.86	20.50	3.78
最寄地下通路までの距離_差分(%)	0.66	1.07	1.49	1.56	1.84	10.32	23.74	59.31
最寄地下通路までの距離_差分(%)	0.19	0.41	0.60	0.80	0.95	6.70	20.13	70.22

単位m	0-100	100-200	200-300	300-400	400-500	500-1000	1000-2000	2000以上
最寄駐車場までの距離_差分(%)	61.26	27.61	6.25	2.28	1.03	1.28	0.23	0.06
最寄駐車場までの距離_差分(%)	40.22	30.49	11.53	5.29	3.11	6.53	2.31	0.53
最寄協同組合までの距離_差分(%)	0.70	1.54	2.34	2.85	3.22	18.20	38.56	32.58
最寄協同組合までの距離_差分(%)	0.36	0.99	1.59	2.07	2.45	17.22	38.06	37.28
最寄病院までの距離_差分(%)	10.64	17.40	15.42	12.65	10.00	25.48	6.99	1.43
最寄病院までの距離_差分(%)	4.44	9.83	11.55	11.31	10.54	30.78	14.45	7.11
最寄ラブホテルまでの距離_差分(%)	1.55	2.77	3.01	3.30	3.04	16.40	30.03	39.89
最寄ラブホテルまでの距離_差分(%)	0.48	1.09	1.53	1.85	2.08	12.82	28.49	51.66
最寄郵便局までの距離_差分(%)	2.33	5.86	7.54	8.08	8.29	35.02	26.84	6.03
最寄郵便局までの距離_差分(%)	1.04	2.92	4.49	5.73	6.39	31.54	33.26	14.63
最寄コインランドリーまでの距離_差分(%)	2.13	4.50	5.54	6.89	6.98	31.31	33.21	9.45
最寄コインランドリーまでの距離_差分(%)	0.84	2.31	3.49	4.39	5.03	27.24	36.60	20.09
最寄公衆浴場までの距離_差分(%)	1.05	2.53	3.32	4.16	4.25	22.84	35.92	25.93
最寄公衆浴場までの距離_差分(%)	0.49	1.25	1.95	2.54	2.92	17.67	35.16	38.03
最寄商店までの距離_差分(%)	34.69	28.37	15.81	8.66	4.93	6.56	0.82	0.16
最寄商店までの距離_差分(%)	14.67	21.95	18.35	12.65	8.31	15.84	6.23	2.00
最寄幼稚園までの距離_差分(%)	6.72	17.77	21.51	17.41	12.71	20.83	2.42	0.62
最寄幼稚園までの距離_差分(%)	4.55	12.84	16.26	15.01	12.21	26.92	8.23	3.99
最寄公共施設までの距離_差分(%)	1.05	2.83	3.76	4.61	5.08	23.83	34.01	24.84
最寄公共施設までの距離_差分(%)	0.54	1.43	2.24	2.89	3.33	20.17	36.39	33.00
最寄社交飲食店までの距離_差分(%)	5.36	6.18	5.50	5.57	5.42	25.14	29.60	17.23
最寄社交飲食店までの距離_差分(%)	1.67	2.69	3.26	3.65	4.09	22.31	34.41	27.92
最寄ジャンクションまでの距離_差分(%)	0.18	0.31	0.46	0.38	0.47	2.88	6.91	88.40
最寄ジャンクションまでの距離_差分(%)	0.09	0.23	0.28	0.31	0.36	2.04	5.56	91.12
最寄サービスエリアまでの距離_差分(%)	0.03	0.09	0.09	0.12	0.11	1.30	4.41	93.85
最寄サービスエリアまでの距離_差分(%)	0.10	0.17	0.18	0.23	0.25	1.56	5.00	92.52
最寄交差点までの距離_差分(%)	57.27	23.98	10.12	4.53	1.99	1.74	0.29	0.08
最寄交差点までの距離_差分(%)	19.66	30.09	20.69	11.59	6.22	8.27	2.57	0.92
最寄り駅までの距離_差分(%)	2.16	5.40	6.40	7.45	7.58	30.04	28.39	12.58
最寄り駅までの距離_差分(%)	0.98	2.78	4.19	5.32	5.84	28.38	31.86	20.65
最寄警察署までの距離_差分(%)	0.66	1.15	1.42	1.67	2.06	13.13	30.91	48.99
最寄警察署までの距離_差分(%)	0.14	0.43	0.68	0.90	1.17	9.02	27.45	60.21
最寄交番・派出所・駐在所までの距離_差分(%)	4.92	7.98	9.62	10.79	11.14	40.80	14.27	0.48
最寄交番・派出所・駐在所までの距離_差分(%)	1.65	4.64	6.96	8.67	9.67	42.26	23.17	2.99
最寄小学校までの距離_差分(%)	1.27	7.24	12.02	15.29	15.92	42.29	5.62	0.35
最寄小学校までの距離_差分(%)	1.43	6.13	10.00	12.53	13.25	40.39	13.33	2.93
最寄中学・高校までの距離_差分(%)	0.72	4.63	8.54	10.76	12.33	48.47	13.80	0.74
最寄中学・高校までの距離_差分(%)	1.10	4.01	6.71	8.80	9.97	43.00	21.37	5.05
最寄大学までの距離_差分(%)	0.12	0.36	0.82	1.22	1.38	10.63	25.89	59.57
最寄大学までの距離_差分(%)	0.15	0.38	0.67	0.90	1.18	8.65	22.88	65.18
最寄特別支援学校までの距離_差分(%)	0.06	0.30	0.64	1.09	1.43	10.16	27.78	58.53
最寄特別支援学校までの距離_差分(%)	0.11	0.32	0.54	0.80	0.97	7.86	22.44	66.97
最寄公園までの距離_差分(%)	12.38	31.07	25.18	14.33	7.00	8.47	1.10	0.47
最寄公園までの距離_差分(%)	13.04	25.53	19.30	11.76	7.23	14.53	5.41	3.21
最寄バス停までの距離_差分(%)	41.93	34.97	13.50	4.99	2.26	2.20	0.15	0.00
最寄バス停までの距離_差分(%)	19.95	34.81	21.23	9.95	4.71	6.90	2.14	0.29

◆ 図表３－13　特徴量分布　人対車両(高齢者)

時間帯(時)

	0～3	3～6	6～9	9～12	12～15	15～18	18～21	21～24
発生	2.18%	1.85%	10.87%	22.39%	18.46%	24.26%	15.01%	4.98%
未発生	7.30%	7.26%	13.11%	15.72%	14.81%	16.19%	14.84%	10.77%

曜日(曜日)

	月	火	水	木	金	土	日
発生	14.83%	15.52%	15.70%	15.55%	17.34%	12.44%	8.61%
未発生	13.87%	14.50%	14.66%	14.80%	15.36%	14.63%	12.18%

月(月)

	1	2	3	4	5	6
発生	9.00%	8.62%	9.65%	8.53%	8.20%	7.35%
未発生	8.66%	8.10%	8.65%	8.52%	9.57%	8.76%
	7	8	9	10	11	12
発生	6.38%	5.72%	6.63%	8.54%	9.99%	11.38%
未発生	7.43%	7.81%	7.44%	8.55%	8.08%	8.43%

気温(度)

	-5 未満	-5～0	0～5	5～10	10～15	15～20	20～25	25～30	30～35	35～40	40 以上
発生	0.0%	0.3%	6.7%	21.0%	18.9%	18.9%	18.6%	12.2%	3.3%	0.0%	0.0%
未発生	0.0%	1.2%	8.4%	18.1%	16.0%	19.4%	20.3%	13.5%	3.1%	0.0%	0.0%

ガソリン価格_レギュラー(円)

	100 未満	100～105	105～110	110～115	115～120	120～125	125～130	130～135	135～140
発生	0.00%	0.00%	3.07%	3.73%	4.28%	7.71%	12.22%	12.23%	7.96%
未発生	0.00%	0.00%	2.94%	3.44%	5.22%	7.77%	11.63%	12.63%	8.79%
	140～145	145～150	150～155	155～160	160～165	165～170	170～175	175～180	180 以上
発生	9.27%	10.56%	10.21%	9.05%	3.47%	4.07%	0.74%	0.57%	0.88%
未発生	8.34%	10.90%	9.48%	8.24%	3.35%	4.41%	0.97%	0.71%	1.17%

道路面積比率(%)

	0～10	10～20	20～30	30～40	40～50	50～60	60～70	70～80	80～90	90～100
発生	4.40%	34.30%	37.82%	16.56%	4.86%	1.52%	0.37%	0.13%	0.04%	0.00%
未発生	28.33%	45.22%	19.29%	4.75%	1.53%	0.66%	0.16%	0.03%	0.03%	0.01%

商業施設面積比率(%)

	0～10	10～20	20～30	30～40	40～50	50～60	60～70	70～80	80～90	90～100
発生	71.87%	14.13%	6.86%	3.20%	1.76%	0.93%	0.74%	0.20%	0.17%	0.13%
未発生	93.13%	3.89%	1.44%	0.76%	0.33%	0.15%	0.13%	0.04%	0.08%	0.05%

交通事故 人対車両 10年 高齢者 :未発生 :発生

単位:m	0-100	100-200	200-300	300-400	400-500	500-1000	1000-2000	2000以上
最寄コンビニまでの距離_差分(%)	30.55	28.17	17.58	10.56	6.25	6.31	0.48	0.11
最寄コンビニまでの距離_差分(%)	9.00	18.53	19.76	15.92	10.83	18.19	5.71	2.07
最寄駐輪場までの距離_差分(%)	47.76	29.76	10.67	5.34	2.37	3.28	0.52	0.30
最寄駐輪場までの距離_差分(%)	21.92	25.90	16.73	10.14	5.90	12.21	4.44	2.76
最寄風俗営業所までの距離_差分(%)	6.62	6.58	6.19	6.09	5.48	23.84	32.40	12.80
最寄風俗営業所までの距離_差分(%)	1.00	1.72	2.63	3.37	3.91	22.51	37.20	27.66
最寄イベント施設までの距離_差分(%)	7.42	13.95	13.55	13.16	12.10	32.53	6.97	0.32
最寄イベント施設までの距離_差分(%)	2.44	7.22	8.81	10.22	10.71	40.66	16.58	3.34
最寄カラオケボックスまでの距離_差分(%)	5.59	6.68	6.97	6.55	5.87	24.83	29.77	13.74
最寄カラオケボックスまでの距離_差分(%)	0.81	1.79	2.64	3.18	4.02	23.39	35.31	28.85
最寄ゲームセンターまでの距離_差分(%)	5.56	6.59	6.52	6.19	5.24	23.95	31.29	14.67
最寄ゲームセンターまでの距離_差分(%)	0.93	2.53	2.96	3.64	3.98	22.44	35.35	28.17
最寄ゴルフ場までの距離_差分(%)	0.10	0.49	0.77	1.36	1.54	10.16	32.75	52.83
最寄ゴルフ場までの距離_差分(%)	0.14	0.55	0.85	1.13	1.23	11.47	32.07	52.55
最寄スーパーまでの距離_差分(%)	21.98	24.77	15.88	12.00	8.31	14.40	2.06	0.61
最寄スーパーまでの距離_差分(%)	5.81	12.17	13.42	13.35	10.84	28.61	10.82	4.97
最寄ドラッグストアまでの距離_差分(%)	10.49	15.54	13.97	12.14	10.34	27.48	8.85	1.19
最寄ドラッグストアまでの距離_差分(%)	2.68	6.23	8.43	9.34	9.68	34.88	20.09	8.67
最寄パチンコまでの距離_差分(%)	14.78	13.82	9.99	9.04	7.06	24.69	16.47	4.15
最寄パチンコまでの距離_差分(%)	2.78	5.35	6.46	7.43	7.19	31.10	26.47	13.23
最寄ホームセンターまでの距離_差分(%)	2.76	5.08	5.88	6.42	5.50	31.52	32.43	10.42
最寄ホームセンターまでの距離_差分(%)	1.03	2.49	3.36	4.33	4.67	26.51	35.80	21.81
最寄マンガ喫茶までの距離_差分(%)	3.26	4.73	4.27	4.09	3.84	16.06	29.46	34.29
最寄マンガ喫茶までの距離_差分(%)	0.47	1.23	1.56	1.73	2.24	13.88	28.52	50.37
最寄レンタルビデオ店までの距離_差分(%)	2.89	5.52	6.14	6.25	6.18	27.22	32.23	13.58
最寄レンタルビデオ店までの距離_差分(%)	0.71	1.87	2.55	3.44	3.79	24.13	36.96	26.55
最寄宿泊施設までの距離_差分(%)	3.84	5.79	4.78	4.19	3.99	19.31	33.31	24.78
最寄宿泊施設までの距離_差分(%)	1.02	1.93	2.03	2.62	3.16	18.81	36.10	34.32
最寄飲食店までの距離_差分(%)	34.96	25.11	14.80	9.03	5.80	9.00	1.17	0.14
最寄飲食店までの距離_差分(%)	10.94	16.99	15.91	13.24	10.73	22.00	7.78	2.42
最寄家電量販店までの距離_差分(%)	2.35	4.15	5.00	5.13	4.78	26.00	34.33	18.26
最寄家電量販店までの距離_差分(%)	0.47	1.74	2.22	2.76	3.60	21.23	36.29	31.69
最寄貴金属店までの距離_差分(%)	4.42	5.89	5.31	5.72	5.40	21.27	31.02	20.97
最寄貴金属店までの距離_差分(%)	0.53	1.43	1.88	2.44	3.07	18.08	35.61	36.96
最寄ガソリンスタンドまでの距離_差分(%)	4.45	8.24	10.54	10.29	10.88	41.49	12.90	1.21
最寄ガソリンスタンドまでの距離_差分(%)	1.92	4.91	7.45	9.12	10.22	39.22	21.82	5.33
最寄金融機関までの距離_差分(%)	23.87	20.12	15.43	12.24	9.22	16.56	2.43	0.13
最寄金融機関までの距離_差分(%)	5.17	10.63	12.97	12.64	12.03	31.88	11.52	3.16
最寄空き地までの距離_差分(%)	2.87	6.75	9.65	10.21	11.03	39.69	19.17	0.62
最寄空き地までの距離_差分(%)	1.81	5.04	7.49	9.46	9.74	40.67	21.69	4.09
最寄古物店までの距離_差分(%)	9.05	12.29	9.91	9.57	8.22	29.05	17.46	4.45
最寄古物店までの距離_差分(%)	2.04	4.91	5.81	6.54	7.05	32.77	26.45	14.42
最寄深夜飲食店までの距離_差分(%)	37.36	18.48	12.17	8.40	6.17	13.66	3.35	0.40
最寄深夜飲食店までの距離_差分(%)	10.59	14.44	12.83	11.52	8.95	26.21	11.76	3.69
最寄神社までの距離_差分(%)	3.04	8.92	12.14	12.50	11.91	38.38	12.81	0.30
最寄神社までの距離_差分(%)	2.13	5.89	8.45	9.84	10.08	39.47	20.36	3.79
最寄地下通路までの距離_差分(%)	1.12	1.90	2.41	2.17	2.02	9.64	23.39	57.35
最寄地下通路までの距離_差分(%)	0.24	0.45	0.57	0.69	0.91	6.75	20.39	70.01

単位m	0-100	100-200	200-300	300-400	400-500	500-1000	1000-2000	2000以上
最寄駐車場までの距離_差分(%)	66.09	25.76	4.88	1.65	0.72	0.74	0.11	0.06
最寄駐車場までの距離_差分(%)	40.00	30.31	12.31	5.32	3.09	6.42	2.03	0.51
最寄協同組合までの距離_差分(%)	0.95	1.69	2.36	2.75	3.21	16.62	35.33	37.09
最寄協同組合までの距離_差分(%)	0.35	0.82	1.51	1.97	2.60	17.56	38.49	36.69
最寄病院までの距離_差分(%)	16.07	22.05	16.05	12.09	8.34	19.62	4.76	1.02
最寄病院までの距離_差分(%)	4.51	9.79	11.47	11.00	10.61	31.21	14.12	7.29
最寄ラブホテルまでの距離_差分(%)	1.56	2.79	2.94	3.15	3.21	15.77	29.86	40.72
最寄ラブホテルまでの距離_差分(%)	0.50	1.11	1.50	1.71	2.08	13.06	28.23	51.80
最寄郵便局までの距離_差分(%)	3.76	8.78	9.88	10.36	8.98	32.55	20.83	4.87
最寄郵便局までの距離_差分(%)	0.97	2.98	4.41	5.80	6.27	31.63	33.29	14.65
最寄コインランドリーまでの距離_差分(%)	2.62	4.85	5.87	6.78	7.50	31.57	30.87	9.94
最寄コインランドリーまでの距離_差分(%)	0.96	2.36	3.35	4.33	4.79	27.53	37.43	19.25
最寄公衆浴場までの距離_差分(%)	1.27	3.56	3.95	4.55	4.60	21.70	34.81	25.56
最寄公衆浴場までの距離_差分(%)	0.45	1.16	1.72	2.52	3.23	18.41	34.64	37.87
最寄商店までの距離_差分(%)	41.55	27.06	13.69	7.63	4.06	5.31	0.60	0.10
最寄商店までの距離_差分(%)	14.99	22.25	17.77	12.52	8.19	16.19	6.14	1.96
最寄幼稚園までの距離_差分(%)	7.67	20.31	23.11	18.09	11.96	16.76	1.65	0.45
最寄幼稚園までの距離_差分(%)	4.54	13.20	16.04	15.30	12.15	26.69	7.96	4.11
最寄公共施設までの距離_差分(%)	1.41	3.70	4.20	5.59	5.47	23.77	32.76	23.10
最寄公共施設までの距離_差分(%)	0.56	1.41	2.23	2.91	3.33	19.59	36.86	33.11
最寄社交飲食店までの距離_差分(%)	10.74	8.79	6.89	5.98	5.00	21.26	26.48	14.86
最寄社交飲食店までの距離_差分(%)	1.70	2.59	3.42	3.76	4.02	22.20	34.37	27.94
最寄ジャンクションまでの距離_差分(%)	0.10	0.32	0.47	0.46	0.59	2.90	7.70	87.45
最寄ジャンクションまでの距離_差分(%)	0.10	0.21	0.30	0.32	0.39	1.93	5.70	91.04
最寄サービスエリアまでの距離_差分(%)	0.01	0.02	0.08	0.06	0.13	1.20	4.45	94.05
最寄サービスエリアまでの距離_差分(%)	0.11	0.19	0.17	0.18	0.29	1.57	5.04	92.45
最寄交差点までの距離_差分(%)	55.65	24.48	10.23	5.49	2.24	1.69	0.19	0.03
最寄交差点までの距離_差分(%)	19.20	30.36	20.78	11.55	6.19	8.34	2.83	0.75
最寄り駅までの距離_差分(%)	4.75	10.05	9.37	9.50	7.91	26.36	22.92	9.14
最寄り駅までの距離_差分(%)	0.85	2.80	4.07	4.81	6.00	29.02	31.15	21.30
最寄警察署までの距離_差分(%)	0.54	1.31	1.80	2.00	2.36	15.37	32.35	44.26
最寄警察署までの距離_差分(%)	0.17	0.39	0.71	0.77	1.12	9.07	27.44	60.34
最寄交番・派出所・駐在所までの距離_差分(%)	7.39	11.42	12.85	12.76	11.43	34.26	9.59	0.28
最寄交番・派出所・駐在所までの距離_差分(%)	1.62	4.49	7.32	8.42	10.17	42.25	22.77	2.97
最寄小学校までの距離_差分(%)	1.67	8.97	13.33	16.94	16.63	38.86	3.40	0.19
最寄小学校までの距離_差分(%)	1.27	6.16	10.21	12.76	13.25	39.69	13.67	2.98
最寄中学・高校までの距離_差分(%)	0.64	4.84	9.05	10.98	13.16	49.40	11.40	0.53
最寄中学・高校までの距離_差分(%)	0.98	3.95	6.85	9.23	9.71	42.77	21.41	5.10
最寄大学までの距離_差分(%)	0.21	0.67	1.06	1.54	1.79	11.70	25.72	57.31
最寄大学までの距離_差分(%)	0.13	0.39	0.76	0.88	1.17	8.95	22.34	65.37
最寄特別支援学校までの距離_差分(%)	0.12	0.25	0.41	1.13	1.38	10.34	27.85	58.52
最寄特別支援学校までの距離_差分(%)	0.13	0.29	0.62	0.70	1.02	8.01	22.37	66.85
最寄公園までの距離_差分(%)	15.02	33.98	26.57	12.75	5.05	5.63	0.79	0.21
最寄公園までの距離_差分(%)	12.97	25.12	19.70	12.02	6.96	14.51	5.57	3.14
最寄バス停までの距離_差分(%)	45.43	34.48	12.26	4.44	1.64	1.61	0.13	0.00
最寄バス停までの距離_差分(%)	20.37	34.66	21.44	9.79	4.64	6.94	1.85	0.33

◆ 図表３－14　特徴量分布　高速道路等

時間帯(時)

	0～3	3～6	6～9	9～12	12～15	15～18	18～21	21～24
発生	3.68%	3.56%	19.97%	16.47%	15.56%	21.45%	13.67%	5.63%
未発生	9.30%	7.92%	14.87%	13.77%	15.01%	16.12%	12.60%	10.42%

曜日(曜日)

	月	火	水	木	金	土	日
発生	13.27%	13.00%	12.36%	12.39%	14.82%	17.47%	16.68%
未発生	12.70%	13.57%	13.77%	13.74%	14.26%	16.41%	15.54%

月(月)

	1	2	3	4	5	6
発生	7.25%	6.78%	10.17%	7.21%	7.69%	6.63%
未発生	7.96%	7.65%	9.63%	7.92%	8.40%	7.12%
	7	8	9	10	11	12
発生	8.45%	10.50%	8.85%	8.54%	9.05%	8.87%
未発生	8.06%	9.17%	8.24%	8.83%	8.42%	8.59%

気温(度)

	-5 未満	-5～0	0～5	5～10	10～15	15～20	20～25	25～30	30～35	35～40	40 以上
発生	0.0%	0.3%	6.0%	17.8%	17.2%	19.3%	19.0%	15.2%	5.1%	0.2%	0.0%
未発生	0.0%	1.2%	8.0%	17.6%	17.9%	18.2%	19.3%	14.6%	3.3%	0.0%	0.0%

ガソリン価格_レギュラー(円)

	100 未満	100～105	105～110	110～115	115～120	120～125	125～130	130～135	135～140
発生	0.00%	0.00%	2.38%	3.12%	4.18%	7.11%	11.62%	13.12%	8.47%
未発生	0.00%	0.00%	2.67%	3.08%	4.73%	6.92%	11.62%	12.49%	8.41%
	140～145	145～150	150～155	155～160	160～165	165～170	170～175	175～180	180 以上
発生	8.28%	12.24%	10.57%	8.49%	3.74%	3.87%	0.76%	0.88%	1.19%
未発生	8.10%	11.77%	10.91%	8.14%	3.95%	4.23%	0.93%	1.01%	1.03%

道路面積比率(%)

	0～10	10～20	20～30	30～40	40～50	50～60	60～70	70～80	80～90	90～100
発生	4.03%	6.92%	12.98%	21.33%	23.88%	17.13%	8.11%	3.05%	2.12%	0.46%
未発生	21.31%	17.13%	20.04%	17.16%	12.69%	6.96%	2.80%	1.14%	0.60%	0.17%

交通事故 高速道路等 :未発生 :発生

単位:m	0-100	100-200	200-300	300-400	400-500	500-1000	1000-2000	2000以上
最寄コンビニまでの距離_差分(%)	6.58	15.17	21.97	14.87	10.40	22.41	5.13	3.48
最寄コンビニまでの距離_差分(%)	5.41	12.48	16.13	14.78	10.60	24.40	10.73	5.46
最寄駐輪場までの距離_差分(%)	16.22	26.13	18.70	11.71	6.61	11.72	6.47	2.44
最寄駐輪場までの距離_差分(%)	12.29	21.20	16.10	10.11	7.61	18.97	10.52	3.19
最寄風俗営業所までの距離_差分(%)	1.60	2.50	2.82	3.06	3.34	20.47	36.50	29.71
最寄風俗営業所までの距離_差分(%)	0.56	1.93	1.87	2.31	2.29	14.89	37.05	39.10
最寄イベント施設までの距離_差分(%)	3.13	8.38	7.21	7.95	8.42	42.37	15.30	7.23
最寄イベント施設までの距離_差分(%)	2.12	5.26	6.69	7.50	8.03	38.45	21.13	10.82
最寄カラオケボックスまでの距離_差分(%)	0.83	1.03	1.33	2.07	3.43	19.38	46.51	25.43
最寄カラオケボックスまでの距離_差分(%)	0.44	0.75	1.15	1.56	2.38	16.36	37.76	39.61
最寄ゲームセンターまでの距離_差分(%)	0.77	1.33	1.39	1.79	4.10	18.57	39.01	33.03
最寄ゲームセンターまでの距離_差分(%)	0.42	0.94	1.36	1.42	2.46	16.33	37.64	39.42
最寄ゴルフ場までの距離_差分(%)	0.00	0.10	0.43	0.31	1.98	14.63	29.16	53.38
最寄ゴルフ場までの距離_差分(%)	0.07	0.39	0.62	0.77	1.21	11.18	24.44	61.31
最寄スーパーまでの距離_差分(%)	4.01	10.83	13.24	14.87	10.64	30.16	10.33	5.92
最寄スーパーまでの距離_差分(%)	2.46	7.44	9.66	10.00	9.69	32.10	20.23	8.42
最寄ドラッグストアまでの距離_差分(%)	0.91	3.32	5.97	7.80	8.61	42.54	20.47	10.38
最寄ドラッグストアまでの距離_差分(%)	0.98	2.71	5.12	5.98	6.68	32.59	28.24	17.69
最寄パチンコまでの距離_差分(%)	1.82	4.89	6.82	9.26	7.44	32.38	23.74	13.65
最寄パチンコまでの距離_差分(%)	1.76	4.39	5.01	6.39	5.98	28.62	24.35	23.52
最寄ホームセンターまでの距離_差分(%)	1.05	2.17	3.65	5.66	5.99	26.72	34.93	19.83
最寄ホームセンターまでの距離_差分(%)	0.41	1.65	2.41	2.73	4.26	23.29	35.11	30.15
最寄マンガ喫茶までの距離_差分(%)	0.46	0.67	0.77	0.57	0.69	9.67	30.75	56.41
最寄マンガ喫茶までの距離_差分(%)	0.06	0.53	0.42	0.58	0.86	8.06	24.87	64.63
最寄レンタルビデオ店までの距離_差分(%)	0.14	0.29	0.79	1.72	2.24	15.99	47.74	31.09
最寄レンタルビデオ店までの距離_差分(%)	0.14	0.52	1.41	1.58	2.69	15.13	36.72	41.81
最寄宿泊施設までの距離_差分(%)	1.29	4.92	3.20	3.60	4.46	19.45	35.65	27.42
最寄宿泊施設までの距離_差分(%)	1.18	2.81	3.07	1.88	2.91	16.54	34.82	36.78
最寄飲食店までの距離_差分(%)	7.40	15.27	18.18	13.07	11.36	25.36	4.80	4.56
最寄飲食店までの距離_差分(%)	5.33	12.52	14.23	11.66	11.70	27.33	10.01	7.22
最寄家電量販店までの距離_差分(%)	0.60	1.22	2.24	2.43	4.32	22.93	35.10	31.16
最寄家電量販店までの距離_差分(%)	0.32	1.31	1.69	1.90	2.70	17.23	34.33	40.52
最寄貴金属店までの距離_差分(%)	0.48	1.27	1.64	1.36	2.91	13.86	36.00	42.49
最寄貴金属店までの距離_差分(%)	0.44	0.89	1.15	1.74	1.63	10.58	29.56	54.01
最寄ガソリンスタンドまでの距離_差分(%)	1.84	6.23	7.88	8.35	9.80	46.57	16.63	2.70
最寄ガソリンスタンドまでの距離_差分(%)	1.43	4.28	6.60	7.81	9.40	38.21	26.78	5.50
最寄金融機関までの距離_差分(%)	2.22	6.73	9.80	11.88	8.80	35.43	17.56	7.59
最寄金融機関までの距離_差分(%)	1.88	5.84	8.31	9.48	9.11	34.71	19.72	10.94
最寄空き地までの距離_差分(%)	1.29	4.72	6.94	8.11	9.69	40.88	22.86	5.51
最寄空き地までの距離_差分(%)	1.77	4.96	6.22	7.67	8.89	36.74	25.23	8.52
最寄古物店までの距離_差分(%)	0.62	2.31	3.49	5.65	7.82	43.36	25.05	11.71
最寄古物店までの距離_差分(%)	0.56	2.24	3.47	4.40	5.64	31.84	31.32	20.52
最寄深夜飲食店までの距離_差分(%)	5.87	11.93	11.43	10.78	8.78	29.68	13.36	8.18
最寄深夜飲食店までの距離_差分(%)	4.26	9.66	10.48	10.01	8.24	29.56	18.92	8.86
最寄神社までの距離_差分(%)	0.38	3.20	6.90	8.18	7.78	43.50	22.79	7.26
最寄神社までの距離_差分(%)	0.76	3.81	7.02	9.09	9.35	38.99	21.50	9.48
最寄地下通路までの距離_差分(%)	0.67	0.72	1.43	1.26	1.51	4.94	15.60	73.87
最寄地下通路までの距離_差分(%)	0.31	0.62	0.49	0.86	0.72	4.22	12.28	80.51

単位m	0-100	100-200	200-300	300-400	400-500	500-1000	1000-2000	2000以上
最寄駐車場までの距離_差分(%)	33.57	33.52	14.44	6.47	2.94	5.89	2.70	0.46
最寄駐車場までの距離_差分(%)	26.08	29.21	15.25	7.88	5.68	10.56	4.15	1.20
最寄協同組合までの距離_差分(%)	0.31	0.09	0.79	1.50	0.71	16.56	32.83	47.22
最寄協同組合までの距離_差分(%)	0.13	0.58	1.11	1.18	1.63	13.11	33.22	49.04
最寄病院までの距離_差分(%)	1.55	6.47	7.88	11.65	13.81	37.75	9.95	10.93
最寄病院までの距離_差分(%)	1.21	4.88	7.90	7.44	9.82	32.50	21.49	14.75
最寄ラブホテルまでの距離_差分(%)	1.95	5.46	5.06	4.96	3.24	13.27	27.08	38.99
最寄ラブホテルまでの距離_差分(%)	1.07	2.33	2.76	2.22	2.18	11.90	26.99	50.55
最寄郵便局までの距離_差分(%)	0.53	1.76	4.11	4.73	4.89	32.36	38.01	13.60
最寄郵便局までの距離_差分(%)	0.31	1.79	3.38	3.73	4.16	27.29	40.06	19.30
最寄コインランドリーまでの距離_差分(%)	0.10	0.50	1.12	2.91	3.01	23.12	46.34	22.90
最寄コインランドリーまでの距離_差分(%)	0.13	0.68	1.73	2.18	3.45	20.87	44.05	26.92
最寄公衆浴場までの距離_差分(%)	1.17	1.48	1.45	2.94	4.77	18.71	29.88	39.59
最寄公衆浴場までの距離_差分(%)	0.35	0.98	1.87	1.97	2.97	13.90	30.72	47.24
最寄商店までの距離_差分(%)	8.35	16.13	21.55	16.54	8.33	19.04	6.52	3.53
最寄商店までの距離_差分(%)	6.79	14.53	16.69	13.29	10.03	23.39	10.45	4.82
最寄幼稚園までの距離_差分(%)	2.15	9.86	14.20	11.57	13.27	36.24	5.75	6.95
最寄幼稚園までの距離_差分(%)	2.17	8.33	11.74	11.22	10.83	35.06	11.76	8.89
最寄公共施設までの距離_差分(%)	0.36	2.00	3.53	3.84	3.03	11.59	42.86	32.79
最寄公共施設までの距離_差分(%)	0.62	1.58	2.29	2.70	2.98	13.40	36.30	40.13
最寄社交飲食店までの距離_差分(%)	0.81	1.82	1.65	2.00	4.72	20.09	35.31	33.60
最寄社交飲食店までの距離_差分(%)	0.42	1.36	1.48	1.87	3.19	16.32	34.92	40.44
最寄ジャンクションまでの距離_差分(%)	4.42	3.05	2.60	3.79	3.49	9.57	10.62	62.45
最寄ジャンクションまでの距離_差分(%)	1.14	2.04	1.67	1.94	1.88	6.15	8.44	76.74
最寄サービスエリアまでの距離_差分(%)	3.34	1.70	1.26	1.15	0.40	3.27	7.92	80.96
最寄サービスエリアまでの距離_差分(%)	1.13	1.53	0.90	0.82	1.01	4.35	7.90	82.36
最寄交差点までの距離_差分(%)	19.80	26.10	21.12	15.06	5.28	8.71	2.91	1.02
最寄交差点までの距離_差分(%)	17.33	23.45	17.33	12.05	8.23	15.09	5.49	1.04
最寄り駅までの距離_差分(%)	0.40	1.91	4.20	3.41	3.62	21.81	38.35	26.30
最寄り駅までの距離_差分(%)	0.82	2.39	3.11	3.80	4.89	21.90	38.37	24.73
最寄警察署までの距離_差分(%)	0.09	0.28	1.14	1.60	0.83	9.57	27.39	59.12
最寄警察署までの距離_差分(%)	0.08	0.30	0.37	0.73	0.86	6.34	21.53	69.79
最寄交番・派出所・駐在所までの距離_差分(%)	0.74	3.15	6.90	8.62	8.93	43.33	21.83	6.49
最寄交番・派出所・駐在所までの距離_差分(%)	0.77	3.16	5.32	6.78	7.95	40.87	26.98	8.17
最寄小学校までの距離_差分(%)	0.15	1.67	6.06	10.17	13.34	45.98	15.42	7.20
最寄小学校までの距離_差分(%)	0.73	2.76	5.95	8.55	10.52	40.75	21.79	8.96
最寄中学・高校までの距離_差分(%)	0.12	2.50	4.72	6.54	8.61	48.63	21.09	7.80
最寄中学・高校までの距離_差分(%)	0.41	2.22	4.49	5.18	6.46	44.74	26.05	10.46
最寄大学までの距離_差分(%)	0.22	0.26	0.46	0.69	0.69	7.23	24.07	66.38
最寄大学までの距離_差分(%)	0.10	0.13	0.35	0.34	0.44	6.33	17.38	74.94
最寄特別支援学校までの距離_差分(%)	0.03	0.28	0.64	0.52	1.81	11.17	29.42	56.14
最寄特別支援学校までの距離_差分(%)	0.15	0.18	0.60	0.82	1.18	9.04	22.11	65.91
最寄公園までの距離_差分(%)	7.09	21.85	21.52	12.12	7.63	19.81	3.89	6.09
最寄公園までの距離_差分(%)	9.02	17.85	16.41	12.07	8.19	22.10	6.84	7.54
最寄バス停までの距離_差分(%)	16.18	28.94	24.03	12.53	6.09	9.36	2.43	0.43
最寄バス停までの距離_差分(%)	13.35	27.05	22.17	12.22	6.86	13.36	4.05	0.94

4　ＸＡＩについて

人工知能の目覚ましい進歩を受け、ディープラーニングやアンサンブル学習などの高度な機械学習アルゴリズムが高い予測性能を示す一方で、人工知能を活用する際の問題点の一つとして

・人工知能の判断過程の「ブラックボックス」化

が挙げられる。

ディープラーニングをはじめとする昨今の機械学習アルゴリズムは、その学習ロジックが複雑で、予測モデルの解釈も困難であるため

○　答えを導き出した根拠となる特徴量が何だったのか

○　それがどのように影響したのか

が人間には理解することができず、判断に至った理由の説明ができないことが大きな問題となってきた。

人工知能を用いて交通事故発生予測を行う場合を考えると、人工知能の予測結果を警察業務に用いるためには、予測の根拠を理解することは不可欠な問題である。

「人工知能がそう予測しているが、その理由はわからない。」という理屈は、地域住民に対しても警察内部に対しても説明責任を果たさなければならない、という点では到底通用しない話だということである。

このような人工知能のブラックボックス化の問題に対して、一定の規制をかける動きも出てきている。

EU では人工知能が個人に対して下した判断の根拠を求める権利を認める規制(欧州一般データ保護規制（GDPR））が施行された。

日本においても総務省が「AI 開発ガイドライン案」を策定し、その第２条の「透明性の原則」において「開発者は AI システムの入出力の検証可能性及び判断結果の

説明可能性に留意すること」と示されるなど、人工知能の説明可能性に関する規制やガイドラインの明文化が進みつつある。

このような人工知能の判断過程のブラックボックス化という課題への対応策として研究が進んでいるのが人工知能の意思決定の根拠を分析するＸＡＩ（eXplainable AI：説明可能な人工知能）と呼ばれる技術である。

ＸＡＩ技術の分類としては、

○　学習済の任意のブラックボックス型の人工知能に対してこのＸＡＩを適用し、この人工知能が重視した特徴量や学習データ等を根拠として提示する方法（Model Induction）

○　人間が理解可能な特徴量や構造を学習させる新しい高性能な機械学習アルゴリズムの開発（Interpretable Model）

などの方法が提案されており、研究開発が行われている。

本調査研究においては、「SHAP (SHapley Additive exPlanations)」と呼ばれる手法を用いて、どのような特徴量が予測結果にどのように影響を与えたのかを分析することにする。

SHAP とは、Model Induction 型のＸＡＩ技術であり、ゲーム理論（複数主体の意思決定や行動に関する利得を数学的モデルで表し、分析する理論）に基づき、人工知能の予測値を各特徴量の貢献度の和に分解する手法であり、予測値に対する各特徴量による感度を分析するような方法である。

簡単にいえば、特徴量が少しずつ違うデータを生成して、それらを入力して人工知能に予測をさせた場合に、予測結果がどのように変化するかを観測し、各特徴量の貢献度を算出していくというものである。

現在、SHAP はオープンソースソフトウエアとしても公開されており、誰でも使用することが可能であるが、本調査研究では日立製作所によって一部の処理が高速化されるなど改良版の SHAP を使用することとした。

SHAP を第 3 回実証実験の予測モデルに適用し、予測根拠を分析した結果のうち、予測に際して重視したものや特徴のあるものをいくつか抽出して添付する。詳細は別添 1 「SHAP による人工知能の学習傾向分析結果」を参照。

SHAP を適用し、人工知能が予測に際して重視した特徴量を確認した結果、これまでに交通事故との関連性に関する知見がなく、なぜ予測に影響するのか説明が困難な特徴量が影響していることや影響の及び方が予想と逆である特徴量があることが示された。

このような特徴量は、新たな交通事故発生の要因である可能性を示唆するものとも考えられる。しかし、ここで注意しなくてはならないのが人工知能は大量のデータから特徴量間の相関関係を探し出すことは得意とする反面、その因果関係を見いだすことはできないという弱点があるということである。

このため、交通事故予測に関して重要度が高いと人工知能が判断した特徴量が、そのまま交通事故の発生要因であると考えてしまうと誤った判断を生む危険性があることを十分に理解しておかなくてはならない。

人工知能が発生予測に際して重視した特徴量は、「疑似相関」や「逆の因果関係」である可能性を十分に考え、交通事故の発生に因果関係のある要因であるのか、単に疑似相関等の関係に過ぎないのかを最終的には人間が判断する必要がある。

その上で、説明性や分かりやすさを重視し、予測精度をある程度犠牲にしてでも因果関係が確認できる特徴量のみを使用するか、疑似相関の可能性があることを前提としたうえで、予測精度を重視してより多くの特徴量を使用するかは、特徴量の内容を吟味して個別に判断する必要があると考えられる。

いずれにしても、人工知能の予測根拠を示すことが可能な技術は、非常に有益なものであり、今後ますます重要なものとなるであろう。

【 用語集 】

○ 特徴量（Feature）

予測対象にどのような特徴があるかを、人工知能が扱うことが可能なように数値化したものをいう。

面積、気温などの数値そのものに意味のある「数値型」の特徴量と、色や形などのデータに一定の決まりを設けてラベル付けすることで数値化した「カテゴリ型」の特徴量に分けられる。

カテゴリ型は、更にカテゴリ内で大小などの優劣や順序付けが可能な「順序特徴量」と、優劣の関係を持たずに識別子として用いる「名義特徴量」に分けることができる。

○ 回帰（Regression）と分類（Classification）

過去のデータを用いて将来の値を予測するものを回帰、グループ分けするものを分類という。

犯罪が「発生する」「発生しない」を予測する問題は分類、「何件の犯罪が発生するか」を予測する問題は回帰問題として考えることになる。

分類の中でも「発生する」「発生しない」のような、2つのグループに分けるものを「二値分類」、3つ以上のグループに分けるものを「多値分類」と呼ぶ。

○ ハイパーパラメータ

機械学習を行う際、予め設定しておかなければならない変数（パラメータ）のことをいう。

ハイパーパラメータの設定次第で機械学習の予測精度が変化することがあるので、この調整は重要な要素である。

しかし、機械学習により適切な数値を得られない変数であるため、人間の経験則や試行錯誤で決定していく必要があり、その調整には技術や相当な労力を要する場合がある。

○　過学習（オーバーフィッティング）

学習データに対する分類能力は高いものの、学習していない未知のデータに対しては正しく分類できない（汎化できていない）状態になることをいう。

○　交差検証（クロスバリデーション）

学習データを分割し、それぞれを順に学習・評価することで過学習していないかを評価する手法のことをいう。

○　L1 正則化

正則化とは、モデルが複雑化することに対してペナルティを課すなどして、過学習を防ぐ手法である。

L1 正則化の場合は、相関の高い説明変数が複数含まれている場合、１つを残して他の変数をゼロとすることで、結果として特徴量選択を行うことになり、汎化能力が向上する。

○　汎化能力

様々な異なる対象に共通する性質、共通して適用できる法則などを見いだす能力のことをいう。機械学習に関して言えば、学習データに対する分類性能だけではなく、未知の新たなデータに対しても正しく答えを返すことができる能力のことである。

○　アンダーサンプリング

正例と負例のデータ数が不均衡な場合、数の少ないデータに合わせて多数派のデータからランダムにデータを抽出して同数程度に合わせることをいう。

不均衡なデータをそのまま学習させてモデルを作ると、少数派データの分類能力が低くなることが知られているため、アンダーサンプリングを行うことで正負双方の分類性能が高くなる。

○　標準メッシュ（標準地域メッシュ）

標準地域メッシュとは、統計利用のために、緯度・経度に基づいて地域をほぼ同じ大きさの網の目（メッシュ）に分割したもので、その分割方法については「統計に用いる標準地域メッシュおよび標準地域メッシュ・コード（昭和 48 年行政管理庁告示第 143 号）」により定められている。

総務省統計局などの行政機関が作成しているメッシュ統計に用いる標準地域メッシュは「基準地域メッシュ」「分割地域メッシュ」「統合地域メッシュ」の３種類と定め、メッシュの区分とコードの表示方法を定めている。

基準地域メッシュ（３次メッシュ区画）は、一辺約１キロメートル（緯度間隔 30 秒、経度間隔 45 秒）の範囲であり、基準地域メッシュを緯度・経度方向に２分割したものを「２分の１地域メッシュ」（一辺約 500 メートル）、２分の１地域メッシュを更に緯度・経度方向に２分割したものを「４分の１地域メッシュ」（一辺約 250 メートル）として、統計情報が公開されている。

100 メートル分割標準メッシュとは、基準地域メッシュ（３次メッシュ区画）を緯線方向及び経線方向に 10 等分してできる区域として利用している。

○　オープンソースソフトウエア（Open Source Software）

ソースコード（プログラム文）が作者により無償で一般公開されており、一部の例外を除き、商用・非商用の目的を限らずソースコードの利用、修正、拡張、再頒布等が可能となっているソフトウエアをいう。

フリーソフトウエアとは、ソースコードの公開・非公開の違いやプログラムの改変が可能か否かという点で違いがある。

有名なものでは UNIX 系ＯＳである「Linux」、スマートフォン用ＯＳである「Android」などがオープンソースソフトウエアとして知られており、多くの個人や企業が利用し、プログラムの改良などを行っている。

数多くの機械学習アルゴリズムや関係するツールなどもオープンソースソフトウエアとして公開されており、本研究においてもこれらのソフトウエアを利用している。

○　疑似相関

実際には２つの事象の間には何の因果関係もないが、双方の間に存在する見えない要因（「交絡因子」「潜在変数」などという。）の介在により、統計的な分析を行うと相関関係が見いだされ、あたかも因果関係があるかのよう推測されてしまうというものをいう。

例として「アイスクリームが売れると、水難事故が増える」というものがある。

この２つの事象の間には相関関係があるが、実際にはこの間に「気温」という交絡因子が存在し、これが「アイスクリームの販売量」と「水難事故の件数」のそれぞれに因果関係があるに過ぎない。

「気温」という因子を介して「アイスクリームの販売量」と「水難事故の件数」が疑似的に相関関係を示してしまっているだけであり、実際には何の因果関係もないことから、アイスクリームの販売量を減らしても水難事故が減ることはないので、問題解決にはつながらない。

○　逆の因果関係

相関関係が存在し、なんらかの因果関係もあるが、原因であると考えられていた事象と結果であると考えられていた事象が実際には逆であるというものをいう。

例として「警察官が増えると、犯罪が増える」というものがある。

警察官が増えたことが原因で結果として犯罪が増えるのではなく、犯罪が増えたことからその対策として警察官の配置を増やしたという原因と結果を取り違えていることから、間違った仮説を導き出してしまった例である。

この仮説に基づいて警察官を減らしてしまうと、普通に考えれば犯罪は増加してしまうと考えられる。

【注釈】

(1) Python のオープンソース機械学習ライブラリ。様々な機械学習アルゴリズムをサポートしている。https://scikit-learn.org/

(2) Python による XGBoost をサポートするオープンソースライブラリ。https://anaconda.org/anaconda/py-xgboost/

(3) Google によるオープンソースソフトウェアライブラリ。ディープラーニングをはじめ、様々な機械学習アルゴリズムをサポートする。https://www.tensorflow.org/

(4) Tensor-Flow の上部で動作するオープンソースのラッパーライブラリ。https://keras.io/ja/

(5) 国土交通省「国土数値情報 土地利用細分メッシュ」で定義されるもの。http://nlftp.milt.go.jp/ksj/jpgis/datalist/KsjTmplt-L03-b.html

【参考文献】

1)　西颯人,樋野公宏『交通事故発生件数のトレンドと季節性の類型化－神奈川県の車両相互事故を対象として－』、公益社団法人日本都市計画学会都市計画報告集 No.17 pp237-242 2018.8

2)　松浦常夫『統計データが語る交通事故防止のヒント』東京法令出版株式会社　2014.6.20