

# 機械学習を用いた物損事故発生地点における人身事故発生リスクの要因分析

知能情報システム工学講座

学籍番号 1814019 近 祐大

指導教員 本吉 達郎 准教授

## 1. はじめに

### 1.1 人身事故予防に対する新たな取り組みの必要性

各国における交通事故による状態別の死者数の構成比率 (Fig.1) によると、日本では交通事故による死者のうち 52%が歩行中もしくは自転車乗用中である。この割合は、Fig.1 に示した他の国と比較して高いと言える。

また近年では、交通事故死者、重傷者における高齢者の割合は年々増加している<sup>1)</sup>。

日本では近年交通事故の発生は減少傾向にあるが、今後さらに高齢化が進行すると考えられていることや、コロナ禍による徒歩、自転車の需要の増加<sup>2)</sup>を考慮すると、人身事故予測・予防に対するさらなる取り組みが必要であると考えられる。

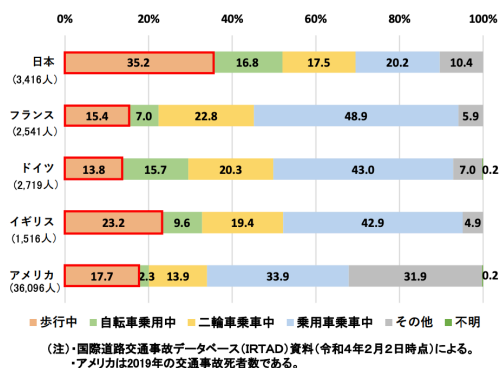


Fig. 1 国別状態別 30 日以内死者数の構成率比較<sup>1)</sup>

### 1.2 人身事故予測・予防の困難性

人身事故が発生しうる地点は無数に存在しており、全ての地点での人身事故発生予測・予防は困難である。そのため、交通事故が発生する可能性が高いと思われる地点を絞った上で人身事故が発生する要因を分析する手法が必要であると考えられる。

また、人身事故の発生に影響を与えると思われる要素は無数にあり、全ての要因を排除することは困難である。そのため、要因排除における優先度を設定する必要があると考えられる。

### 1.3 本研究のアプローチ

本研究では、過去に発生した交通事故データに対してクラス分類を用いることで人身事故か物損事故かを決定づける特徴を分析することで、人身事故の発生に大きな影響を与える要因を発見する手法を提案する。

## 2. 本研究で用いるデータ

本研究では、Fig.2 に示す富山県内で発生した人身事故に関するデータ (以下人身事故データ) と Fig.3 に示す富山県内で発生した物損事故に関するデータ (以下物損事故データ) を用いる。以下、人身事故データと物損事故データを合わせて交通事故データとする。

人身事故データは平成 29 年から令和 3 年の間に富山県内で発生した人身事故に関して、道路や環境の特徴、当事者の特徴など 98 個の属性が記録されたものである。

物損事故データは令和 3 年に富山県内で発生した物損事故に関して、道路の特徴など 25 個の属性が記録されたものである。

なおこれらのデータはいずれも富山県警から入手したものである。

事故概況	発生日	発生時刻	昼夜	曜日	天候	市町村CD
第 1 当事	2017/12/31	11:45	昼	日	曇	小矢部市
第 1 当	2017/12/31	8:10	昼	日	曇	射水市
第 1 当事	2017/12/30	15:05	昼	土	曇	砺波市
第一当事	2017/12/31	3:30	夜	日	晴	小矢部市
第一当事	2017/12/30	14:00	昼	土	晴	滑川市
第 1 当	2017/12/29	12:20	昼	金	晴	砺波市
第 1 当	2017/12/28	6:30	明	木	雪	高岡市
第 1 当事	2017/12/28	3:15	夜	木	雪	射水市
第一当	2017/12/28	17:50	夜	木	曇	入善町

Fig. 2 人身事故データ (一部抜粋)

昼夜	曜日	天候	市町村	地形	路面状態	道路形状	信号機
昼	日	曇	入善町	非市街地	乾燥	その他	施設なし
昼	日	曇	入善町	非市街地	乾燥	カーブ・屈折	施設なし
昼	金	曇	入善町	非市街地	湿潤	その他	施設なし
暮	月	曇	入善町	その他	積雪	その他	施設なし
明	火	曇	入善町	非市街地	乾燥	交差点付近	点灯 (3 灯式)
昼	月	雪	朝日町	非市街地	積雪	交差点	施設なし
明	木	雨	朝日町	その他	湿潤	交差点	施設なし
暮	金	曇	入善町	非市街地	乾燥	その他	施設なし
昼	土	晴	入善町	非市街地	乾燥	その他	施設なし

Fig. 3 物損事故データ (一部抜粋)

## 3. 本研究で用いる分析手法

### 3.1 クラス分類

本研究では、クラス分類を用いて交通事故データに対し人身事故 (第一当事者が怪我)、人身事故 (第二当事者が怪我)、物損事故の 3 種類に分類する。

クラス分類の手法として、線形サポートベクターマシン、非線形サポートベクターマシン、勾配ブースティング決定木などが挙げられる。

本研究では、複数のクラス分類手法を用いて学習を行い、分析に用いるクラス分類手法を選定する。

### 3.2 形式概念分析

形式概念分析 ( Formal Concept Analysis ) は 1981 年の Banff 会議 "Ordered Sets" において Rudolf Wille 氏によって提案されたデータ分析手法である<sup>3)</sup>。Fig.4 に示すようなデータ集合に対して形式概念分析を用いることで、複雑なデータの構造の分析や、属性間に成り立つ含意論理の抽出によるデータ全体が持つ傾向分析を行うことが可能である。

### 3.3 Implication と Association Rule

形式概念分析では、「A という属性を持つものは必ず B である」というルールを抽出することで、データの集合における属性間の関係を可視化する。このルールのことを Implication (含意論理) という。動物の例における Implication を Fig.5 の番号 1 及び 2 に示す。

また形式概念分析の補助として、「A という属性を持つオブジェクトのうち何%は B という属性を持つ」というルールを抽出することで、外れ値によって Implication として抽出されない属性感の関係を抽出することが可能である。このルールのことを Association Rule (関連) という。動物の例における Association Rule を Fig.5 の番号 3 及び 4 に示す。

本研究では交通事故データにおける Association Rule を抽出することで、各種類の交通事故が持つ傾向を抽出する。

## 4. 分析用データ収集手法

### 4.1 物損事故, 人身事故間のデータの差異

本研究で用いる人身事故データと物損事故データの間には記録されている属性に違いがあり、分析にあたって保有する属性を一致させる必要がある。本研究では各事故データに対して属性を追加するためのツールとして、座標周辺探索プログラム及び地図画像分類システムを作成した。

	二足	四足	ヒレ	陸上	水中	空中	胎生	卵生	哺乳類
イヌ		x		x			x		x
ネコ		x		x			x		x
カラス	x			x		x			
ニワトリ	x			x				x	
カエル		x		x	x			x	
メダカ			x		x			x	
カモノハシ		x	x	x	x			x	x
オオサンショウウオ		x	x	x	x			x	
クジラ			x		x		x		x

Fig. 4 動物の例におけるコンテキスト表

番号	前提部	確信度	結論部
1	胎生	100%(3/3)	哺乳類
2	陸上・水中	100%(3/3)	四足・卵生
3	水中	80%(4/5)	卵生
4	哺乳類	75%(3/4)	胎生
⋮	⋮	⋮	⋮

Fig. 5 動物の例における Implication と Association Rule の一部

## 4.2 座標周辺探索プログラム

本プログラムでは、国土交通省の国土数値情報ダウンロードサービス<sup>4)</sup>より入手した各施設に関する外部データを用いて、交通事故発生地点周辺における特定の施設の有無を判定する。

本プログラムを用いて交通事故発生地点が小学校や都市公園などの施設の周辺であるかどうかを判別し、属性として保存する。

## 4.3 地図画像分類システム

本システムでは、道路線形や道路形状、信号の有無などといった人身事故データのものに記録されている属性を機械学習を用いて物損事故データに追加する。Fig.6に簡易的なサーバ構成を示す。

本システムは、「地図画像の取得及び整形用スクリプト配布サーバ」(以下 WebAP サーバ)と、「タイル画像配布サーバ」(以下地図配布サーバ)によって構成される。

WebAP サーバは Java プラットフォーム向けのフレームワークである SpringBoot を用いて作成した。また地図配布サーバにはリクエストに応じてタイル画像を送信する機能を持つオープンソースソフトウェアである mbtiles-server<sup>5)</sup>を用いた。今回は地図画像データベースとして、ダウンロード及び個人利用が許可されている OpenMapTiles<sup>6)</sup>と地理院タイル<sup>7)</sup>を用いた。本システムを用いて取得した交通事故発生地点周辺地図の例を Fig.7 に示す。

本システムでは、Python で記述された画像分類アプリケーションがブラウザを介して物損事故発生地点周辺の地図画像を取得して、人身事故データを用いて学習した結果をもとに物損事故発生地点の情報を追加する。なお画像分類アプリケーションには、分類モデルとして学習済みネットワークである VGG-16 をもとに人身事故データを用いて学習したモデルを用いる。

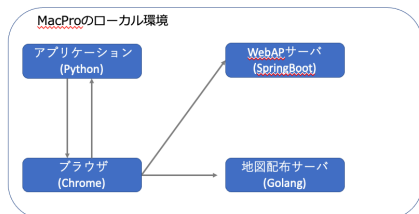


Fig. 6 地図画像分類システムの構成 (概要)

## 5. 分析用データ収集の結果

### 5.1 座標周辺探索プログラムによるデータ収集の結果

国土数値情報ダウンロードシステムの”小学校区 (ポリゴン) を用いて交通事故発生地点周辺の小学校の有無を判定し、交



Fig. 7 交通事故発生地点周辺地図の例

通事故データに追加した。人身事故発生地点のうち小学校周囲 100m 以内と判定された地点の地図 (Google Map) を Fig.8 に示す。



Fig. 8 小学校周辺の交通事故発生地点の例

## 5.2 地図画像分類システムによるデータ追加の結果

構築した地図画像分類システムを用いて交通事故データの座標に対して道路上か道路以外か分類した。人身事故データを用いた学習の結果を Table.1 に示す。学習結果をモデルとして保存し、物損事故データに対し推定を行った。

Table 1 交通事故発生位置に対する学習の結果

データの種類	正解率
道路上で発生した事故データ	0.8942
道路外で発生した事故データ	0.9016
全体の事故データ	0.8945

## 6. まとめ

本研究では、過去に物損事故が発生した地点において人身事故発生リスクの要因を分析する手法として、交通事故データに対してクラス分類を行う手法を提案する。

交通事故データに対してクラス分類を行う準備として、人身事故及び物損事故発生地点に関する特徴を追加するツールを作成した。

今後の作業として、作成したツールを用いたデータ収集を行い、収集したデータを用いて作成した交通事故データに対してクラス分類を行い、人身事故と判定されるデータに共通する特徴を抽出し、形式概念分析を用いて抽出された特徴が Association Rule として抽出されることを確認する。

## 文 献

- 1) 警察庁交通局, 令和 3 年における交通事故の発生状況等について, 2022
- 2) 国土交通省, 令和 2 年度第 2 回自転車活用の活用推進に向けた有識者会議 / 自転車を巡る現状等【参考資料】
- 3) 鈴木治, 室伏俊明, 形式概念分析-入門・支援ソフト・応用-, 日本知能情報ファジィ学会誌, Vol.19, No.2, pp.103-142, 2007
- 4) 国土交通省, 国土数値情報ダウンロードサービス, [https://nlftp.mlit.go.jp/ksj/], (2022/11/14)
- 5) consbio, mbtilesserver, [https://github.com/consbio/mbtilesserver], (2022/11/14)
- 6) MapTiler, OpenMapTiles, [https://openmaptiles.org/], (2022/11/14)
- 7) 国土地理院, 電子国土基本図