# 交通運輸技術開発推進制度 令和 4 年度業務実績説明書

先進安全技術による被害低減効果予測のための 車両の衝突直前挙動に基づく傷害予測モデルの構築

> 令和5年3月 一般財団法人 日本自動車研究所 東京工業大学 国立研究開発法人 産業技術総合研究所

# 研究成果要約

作成年月	令和5年3月
研究課題名	先進安全技術による被害低減効果予測のための車両の衝突直前挙動に基づく
	傷害予測モデルの構築
研究代表者名	一般財団法人 日本自動車研究所
	佐藤 房子
研究期間	令和2年9月4日~令和5年3月31日
研究の目的	自動車の先進運転支援システムや自動運転システムの高度化や普及には、
	これらシステムに搭載されている各先進安全技術の安全性や事故低減効果を
	適切に評価することが必要である. 本研究では、衝突直前の車の挙動から衝
	突後に発生する傷害までの関係を繋いだ傷害予測モデルを構築し、各先進安
	全技術が対象とする衝突直前の状況に基づき、その被害低減効果を、交通参
	加者の属性に応じて定量的に評価するための手法を提案する.
研究成果の要旨	令和2年度は、米国の大規模ミクロ事故データベースを用い、機械学習のた
	めの事故データベースを構築した.この事故データベースを用い,傷害予測モ
	デルのプロトタイプを、確率モデルの一つであるベイジアンネットワークによ
	り構築した.
	令和3年度は,前年度に構築した傷害予測モデルのプロトタイプをベースに,
	ベイジアンネットワークおよび LightGBM を実装した機械学習手法により傷害
	予測モデルを完成させた. 構築した傷害予測モデルは、米国の事故データに基
	づいていることから,国内ミクロ事故データを構築した傷害予測モデルに適用
	し、国内の事故に対する予測性能を確認した.
	令和4年度は、傷害予測モデルの構築に用いたデータセットを見直すことで、
	軽自動車関連の事故における予測性能を向上させ、国内の事故に対しても予測
	可能な傷害予測モデルを構築した. 続いて、構築した傷害予測モデルを構成し
	ている変数(事故データ項目)間の因果関係から、傷害予測に大きく寄与する
	事故データ項目を明らかにした. 最後に、本研究の成果をさらなる発展研究や
	社会実装へとつなぐため、本研究で構築した傷害予測モデルや傷害予測モデル
	構築までに得られた知見の活用先について検討した. その結果, 現在, 保険会
	社,自動車会社,ソフトウェア会社の3社と議論が進行中である.
知的財産権 取得状況	特許出願 1件
研究成果発表実績	論文発表: 国内 0件,海外 2件
	口頭発表:国内 2件,海外 2件
	その他:特になし

# 研究開発の目的および実施体制

# 研究開発の目的

自動車の先進運転支援システム (ADAS) や自動運転システム (ADS) の高度化や普及には、これらシステムに搭載されている各先進安全技術の安全性や事故低減効果を適切に評価することが必要である. 本研究では、衝突直前の車の挙動から衝突後に発生する傷害までの関係を繋いだ傷害予測モデルを構築し、各先進安全技術が対象とする衝突直前の状況に基づき、その被害低減効果を、乗員の属性に応じて定量的に評価するための手法を提案する.

# 研究実施体制

本研究は、日本自動車研究所を総括研究機関とし、東京工業大学、産業技術総合研究所と共同して実施 した. 担当機関の研究実施の流れを示すチャートは以下の通りである.

#### 全体課題名

「先進安全技術による被害低減効果予測のための 車両の衝突直前挙動に基づく傷害予測モデルの構築」

機関名及び研究代表者氏名

一般財団法人 日本自動車研究所 佐藤 房子

## 個別課題名

「米国の大規模ミクロ事故 データを学習させた傷害予 測モデルの構築」

機関研究代表者東京工業大学

宮崎 祐介

# 研究内容

#### 2020年度

- ・NASS-CDS, CISS を用いた機械 学習用データベースの構築
- NASS-CDS, CISS を用いた米国 の事故分布の把握
- ・ベイジアンネットワークによる傷害予測モデルのプロトタイプ構築

# 2021 年度

- ・ベイジアンネットワークによる傷害予測モデルの構築
- ・事故形態のパタン分析と国内 事故の特徴抽出

#### 2022 年度

- ・交通事故シミュレーションを 活用した傷害予測モデルの補正
- 研究成果の活用検討

## 個別課題名

「国内の事故実態に則した 傷害予測モデルの構築」と全 体総括

機関研究代表者日本自動車研究所

佐藤 房子

#### 研究内容

#### 2020年度

・マクロデータを用いた国内の事 故分布の把握

### 2021 年度

- 検証に用いるミクロデータの選 定と把握
- ・NASS-CDS, CISS を用いて構築 した傷害予測モデルの国内事故 データに対する適用性の検討

#### 2022 年度

- ・交通事故シミュレーションを活 用した傷害予測モデルの補正
- ・国内の事故実態に則した傷害予 測モデルの構築
- ・研究成果の活用検討

#### 個別課題名

「機械学習アルゴリズムの 実装と事故データ項目間の 因果関係の分析」

機関研究代表者 産業技術総合研究所

北村 光司

#### 研究内容

#### 2020年度

- ・機械学習のためのデータベース 構築手法の検討
- ・機械学習の実装

#### 2021 年度

- ・事故形態のパタン分析手法の検討
- ・傷害予測モデルの国内事故データに対する適用性検証手法の検討

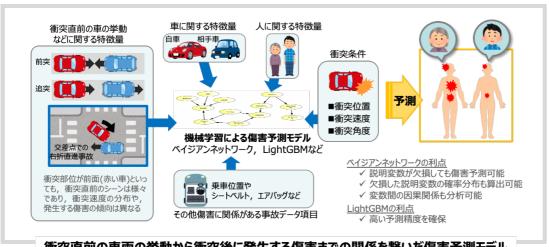
#### 2022 年度

- ・傷害予測モデルに寄与する事故 データ項目間の因果関係分析
- ・研究成果の活用検討

## 研究開発成果

#### 1. 序論

自動車の先進運転支援システム(ADAS)や自動運転システム(ADS)の高度化と普及に向けた様々 な場面において、シスムテによる被害低減効果の予測や、安全性の定量的な評価が必要となる[1].評価 が必要となる具体的な場面としては、システムを搭載した車両の安全性能の評価や、衝突が避けられな い場合の回避挙動を決定するアルゴリズムの開発等が挙げられる。そこで本研究では、衝突の一つ前の 段階である衝突直前の状況から衝突後に発生する傷害までの関係を繋いだ傷害予測モデルを構築し、各 先進安全技術が対象とする衝突直前の状況に基づき、その被害低減効果を、乗員の属性に応じて定量的 に評価するための手法を提案することを目的とした.



既存の傷害予測モデル (ロジスティック回帰モデル) ✓衝突形態が限定的

- ✓説明変数が少なく、多 様な衝突状況を考慮で きない
- ✓説明変数との非線形な 関係は表現できない
- ✓不明な説明変数がある と、傷害の予測はでき なくなる

衝突直前の車両の挙動から衝突後に発生する傷害までの関係を繋いだ傷害予測モデル

- 傷害予測モデルにベイジアンネットワークを用いることにより、取得した事故データ項目群から未取得事故データ項目の確率分布を算出する確率推論が可能となるため、衝突直前の車の挙動から、その他の事故データ項目との関係性に基づき、衝突速度等の確率分布が算出できるようになる。また変数間の因果関係も分析可能。
- ・傷害予測モデルにLightGBMを用いることにより, 高い予測精度を確保. 但し, 変数間の因果関係はわかりにくく, また予測には全変数の値が必要

図1 本研究の概要

本研究は3年計画のプロジェクトとし、下記①~④に沿って傷害予測モデルを構築した.

- ① 大規模ミクロ事故データを用いた機械学習のためのデータベース構築(令和2年度)
- ② 米国の交通事故データに基づく傷害予測モデルの構築と国内事故データに対する適用性の 検討(令和3年度)
- ③ 国内の事故実態に則した傷害予測モデルの構築(令和4年度)
- ④ 研究成果の活用検討(令和4年度)

具体的には、事故の状況を詳細に調査・分析した大規模ミクロ事故データである米国の National Automotive Sampling System- Crashworthiness Data System (NASS-CDS) [2][3][4]および Crash Investigation Sampling System (CISS) [5][6][7]を統合し、機械学習のためのデータベースを構築した、そして、確率モ デルの一つであるベイジアンネットワーク[8]や、決定木ベースの機械学習モデルの一つである LightGBM [9]等を実装した機械学習手法を適用することで、衝突直前の車の挙動に関する特徴量のほか、 年齢、性別などの人の特徴量、衝突条件や車の特徴量などの事故データ項目から、乗員の受傷リスクを 予測するための傷害予測モデルを構築した. さらに、傷害予測モデルの構築に用いているデータセット に国内のミクロ事故データを混合することにより、国内の事故実態に則した傷害予測モデルを構築した. 加えて、本予測モデルを構成する変数(事故データ項目)間の因果関係から、傷害予測に大きく寄与する事故データ項目を明らかにした。本書では、3年間で得られた研究成果について、最終年度である今年度に得られた成果を中心にその概要を述べる。なお、①で構築したデータベースは、②③の中で改良・変更等を加えており、ここではその最終版の概要を述べる。

# 2. 大規模ミクロ事故データを用いた機械学習のためのデータベース構築

傷害予測モデルを構築するにあたり、傷害の情報も含む詳細な交通事故データ(ミクロ事故データ)が必要となる。本研究では、事故の状況を様々な角度から調査・分析し、事故調査項目が豊富に揃っている米国の大規模ミクロ事故データベース NASS-CDS および CISS を用いた。NASS-CDS は 1979 年~2015 年に発生した事故のデータ(年間平均約 4700 件)で構成され、その後 NASS-CDS のデータシステム自体が見直され CISS へと移行し、CISS には 2016 年以降に発生した事故のデータが格納されている。本研究では新しい事故データも使用し、できるだけ現状の事故実態に近づけるべく、NASS-CDS と CISSを統合した。NASS-CDS および CISS の事故データを統合して 1 つの事故データベースを構築するにあたり、事故データ項目の対応付けを行った。また、傷害データである Abbreviated Injury Scale (AIS) [10]は、NASS-CDSでは AIS1990 – Update 98 (AIS1998) および AIS2005 – Update 2008 (AIS2008) のバージョンが用いられており、CISSでは AIS2015 バージョンのみが用いられている。そこで本研究では、NASS-CDS に記録されている AIS データを AIS2015 に変換することで、AIS のバージョンを統一した。

本研究では傷害予測モデルを構築するための米国データとして、NASS-CDS の事故発生年 2010 年~2015 年、および CISS の事故発生年 2017 年~2019 年のデータを対象とした(CISS の 2016 年データはパイロットスタディのため除外). この中から、「一般的な乗用車対乗用車の衝突事故における成人乗員の傷害」に着目し、また傷害予測が難しい特異な事故を除外するなどして予測性能の確保を図りつつ、該当する乗員の絞り込みを行った. さらに、本研究で用いた目的変数(AIS データ)および説明変数(表1)に1つでも欠損がある乗員データを除外し、最終的に残った乗員データを傷害予測モデル構築に用いた. また、NASS-CDS および CISS の各事故データには、全米の交通事故総数に換算した重み付け係数 Ratio Inflation Factor(NASS-CDS)、Case Weight(CISS)が付与されている. 本研究ではこれらを考慮することで全米規模のデータセットとした.

種類	特徴量	
	ŧ	デルイヤー
車両	車種	自車, 相手車
半凹	重量	自車, 相手車
	車	両重量比

種類	特徴量		
衝突	衝突シナリオ		
	衝突時の速度差 (ΔV)		
	衝突方向(PDOF)	自車, 相手車	
	衝突位置(GAD)	自車, 相手車	
	車両変形量		
	車室内侵入量		

	種類	特徴量
		性別
		年齢
	乗員	身長
	1	体重
		BMI
		傷害程度

種類	特徴量	
乗員 2	乗車位置(左右) 乗車位置(前後) 正面エアバッグ その他のエアバッグ シートベルトの装着	

表 1 説明変数の概要

# 3. 米国の交通事故データに基づく傷害予測モデルの構築と国内事故データに対する適用性

本研究では機械学習モデルとして、確率モデルの一つであるベイジアンネットワーク、および決定木アルゴリズムに基づく勾配ブースティング手法の一つである LightGBM を用いた. これまでの傷害予測モデル[11]は、線形ロジスティック回帰[12]にて構築されているものが大多数である. 従来の線形ロジスティック回帰による傷害予測モデルと比較し、ベイジアンネットワークを用いた傷害予測モデルには以下の利点がある.

- 1. ベイジアンネットワークでは、予測したい結果の変数を後から選択して予測することができるが、 既存手法(ロジスティック回帰モデル)では原因→結果の一方向だけの予測しかできない. また、 モデル構築前に結果の変数(目的変数)を指定してモデル化を行う必要がある.
- 2. ベイジアンネットワークでは、説明変数が欠損していても目的変数の予測が可能であるが、既存手法は予測モデルの説明変数が欠損すると目的変数の予測ができない.
- 3. ベイジアンネットワークでは、変数間の関係を条件付き確率としてグラフ構造で表現するため、変数間の因果関係を分析することができる.

一方で、ベイジアンネットワークを用いた傷害予測モデルは、予測精度の観点から、その他の機械学習 手法と比較して性能が劣る可能性もある。そこで、より分類性能が高い機械学習手法の一つである LightGBM を用いたモデルの構築も行った。ただし LightGBM は、ロジスティック回帰と同様に、モデル構築時に目的変数を指定する必要があること、説明変数が1つでも欠損すると目的変数が予測できないこと、また変数間の因果関係がわかりにくいといった特徴を有する。

これらベイジアンネットワークおよび LightGBM を実装した機械学習手法を, 前章で構築した米国の事故データベースに適用し, 乗員の傷害予測モデルを構築した. ここでは, 各車両において重傷度が最も高かった乗員(最大傷害乗員)を対象とし, モデル構築には最大傷害乗員のデータを用いた. 目的変数は, 国内の交通事故の定義でいう重傷以上の傷害に相当する Maximum Abbreviated Injury Scale 3+ (MAIS3+) を対象とし, 最大傷害乗員が MAIS3+であるか否かとした.

その結果、構築した傷害予測モデルは、従来の線形ロジスティック回帰を用いた傷害予測モデルURGENCY[11]と比較して、ベイジアンネットワークによるモデルでは同等、LightGBMによるモデルではより高い予測性能を確保できることを確認した。また、図2にベイジアンネットワークにより構築した最大傷害乗員 MAIS3+傷害予測モデルのグラフ構造を示す。本研究では従来の傷害予測モデルと比べ多様な特徴量を変数として考慮しており、ベイジアンネットワークによる傷害予測モデルではこれら変数間の因果関係がモデル化され、傷害の程度に直接関係する特徴量のほか、より上位の特徴量との関係もモデル化できていることがわかる。さらに図3に示すLightGBMの変数の重要度も併せて検討することで、衝突シナリオから傷害発生に至る過程において重要な特徴量を明らかにすることができた(詳細は次章)。

構築したベイジアンネットワークによる傷害予測モデルに、日本のミクロ事故データ(交通事故例調査データ、公益財団法人交通事故総合分析センター(ITARDA)所有)を適用し、国内の事故に対する予測精度を確認した。その結果、軽自動車同士の事例に絞り込んだデータセットにおいて予測精度が高くなることがわかった。これは、従来の傷害予測モデルに対し、本研究で新たに変数として追加した特徴量の一つである自車と相手車の車両重量比の効果だと考えられる。一方、軽乗用車・普通乗用車の組み合わせで抽出したデータセットにおいては予測精度が低かった。本章で構築した傷害予測モデルは米国データを用いて学習したモデルであり、日本と米国で事故データ分布が異なることから、米国データのみで学習したモデルでは、日本のデータに対して高い性能を発揮できないためと考えられる。次章「国内の事故実態に則した傷害予測モデルの構築」では、これらの事例に対し予測精度を確保することが課題となった。

また,①で抽出した乗員のうちドライバーを対象とし,全身の最大 AIS スコアに対する予測モデルとして目的変数を MAIS3+, MAIS2+としたモデルと,傷害部位ごとの予測モデルとして頭部 AIS1+および2+,頚部 AIS1+,胸部 AIS1+,2+,3+を目的変数とした予測モデルをそれぞれ構築した.その結果,傷害部位ごとの傷害予測モデルにおいて,ベイジアンネットワークを用いたモデルの予測性能が低下した.

一方, LightGBM を用いた傷害予測モデルは傷害部位別のモデルであっても、全身を対象とした傷害予測モデルである URGENCY よりも高性能であった. 従来のモデルでは全身の傷害予測のみしか行えなかったが、本研究により詳細な部位ごとの傷害予測が可能となった.

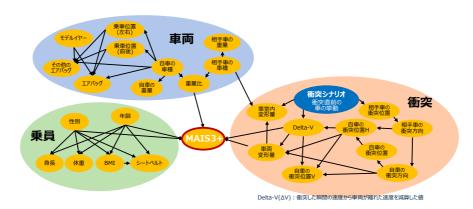


図 2 ベイジアンネットワークにより構築した最大傷害乗員 MAIS3+傷害予測モデルのグラフ構造 学習データ:米国データ

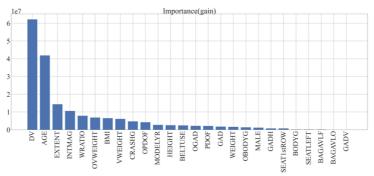


図3 LightGBM により構築した最大傷害乗員 MAIS3+傷害予測モデルの変数の重要度

# 4. 国内の事故実態に則した傷害予測モデルの構築

前章で構築した傷害予測モデルは、米国データを用いて構築しているため、国内の交通事故データを 適用すると、軽自動車絡みの事故において予測精度が低下した。日本におけるミクロ事故データは収集 地域や収集条件に偏りがあり、また米国のミクロ事故データと比較すると調査件数が少なく、傷害予測 モデル構築の利用には適していない(図 4)。そこで本研究では、学習に用いたデータセットの見直しと して、米国データに日本データを混合した。この日米混合データを用い、ベイジアンネットワークおよ び LightGBM を実装した機械学習手法により傷害予測モデルを構築することで、国内の事故実態に即し た傷害予測モデルとして、軽自動車関連の予測における性能向上を試みた。



図4 米国データおよび日本データを用いた場合の問題点

日本データには、日本のミクロ事故データである交通事故例調査データ(公益財団法人 交通事故総合分析センター(ITARDA)所有)を用いた. 1章の米国データの場合と同様に、事故発生年 2010 年~2019 年における「一般的な乗用車対乗用車の衝突事故における成人乗員の傷害」に着目し、また傷害予測が難しい特異な事故を除外するなどして予測性能の確保を図りつつ、該当する乗員の絞り込みを行った. さらに目的変数・説明変数に1つでも欠損がある乗員データを除外し、最終的に残った乗員データを傷害予測モデル構築に用いた. また、日米の事故データを混合するにあたり、説明変数の一つである衝突シナリオに該当する衝突形態の分類が異なるため、日米共通の分類を定義した. この定義に基づき、日本データおよび米国データの衝突シナリオの値を振り直した.

日米混合データは、混合比率を「米国データの重み合計値」に対する「日本データの重み合計値」と 定義し、混合比率を調整することにより作成した。この日米混合データに、ベイジアンネットワークおよび LightGBM を実装した機械学習手法を適用し、乗員の傷害予測モデルを構築した。そして、日本データの重みを増やし、日米混合データの混合比率の変化による予測性能の変化を観察した。また、比較のために米国データのみで学習した傷害予測モデルも構築した。なお、目的変数は2章と同様に最大傷害乗員が MAIS3+であるか否かとした。

米国データのみで学習したモデルでは、日本データで検証した結果は、米国データで検証した結果よりも予測性能が低下していた。日米混合データを用いて学習したモデルでは、混合比率 0.01 付近で予測性能が高くなり、米国データのみで学習し日本データで検証した結果と比べ、予測性能が改善した.

また、図5に検証に用いた日本データを自車・相手車の車種で分けた場合のROC-AUCの結果の一例を示す.軽乗用車と普通乗用車の組み合わせの事故において、米国データのみで学習したモデルでは予測性能が低下したが、日米混合データを用いて学習したモデルでは改善することを確認した。米国データのみの学習ではうまく予測することができなかった軽乗用車・普通乗用車の事故に対する予測性能が、日米混合データを用いることで改善することから、日米混合データを用いることで、データ数の多い米国データを活用しつつ、日本の事故実態に即した傷害予測モデルを構築することができたと言える。

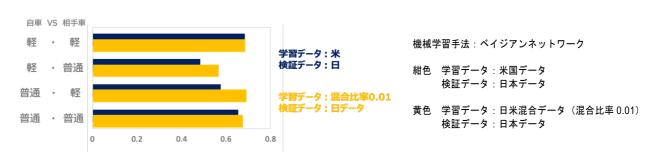


図 5 検証に用いた日本データを自車・相手車の車種で分けた場合の ROC-AUC

図 6, 図 7 にベイジアンネットワークにより構築した最大傷害乗員 MAIS3+傷害予測モデルのグラフ構造を示す. 日米混合データを学習したモデルと米国データのみで学習したモデルのグラフ構造は、車両変形量・相手車の車種間のリンクの有無に違いがあったが、それ以外は全く同じ結果を示した. 双方のモデルにおいて、MAIS3+と直接リンクを持った変数は、自車に対する相手車の車両重量比、性別、年齢、ΔV、車両変形量であった. これら 5 つの変数が、双方のモデルで MAIS3+に繋がったということは、乗員の傷害が重傷以上であるかを予測する上で、非常に重要な特徴量であると言える. 特に車両重量比は、米国データのみで学習したモデルでも軽乗用車同士の予測性能を確保できたことから(図 5)、米国データを活用しながら国内の実態に沿った傷害予測に大きく寄与する特徴量であると考えられる.

図8にLightGBMにより構築した最大傷害乗員MAIS3+傷害予測モデルの変数の重要度を示す.図8(a) と(b)に示す変数の重要度は同じグラフであり、図8(b)においては既存の傷害予測モデルURGENCYに使用されていた変数との比較を示す。図8(a)に示す通り、日米混合データを学習したモデルと米国データのみで学習したモデルの双方で上位5つの変数は同じであり、また次に続く5つの変数では4つの変数が共通でランクインしていた。ここから、 $\Delta V$ 、年齢、車室内侵入量、車両変形量、BMI、自車および相手車の車両重量、車両重量比、乗員の体重が重要であることがわかる。日米混合データで構築したモデルでは、車両重量比の重要度の順位が上がっていたことから、車両重量比がより傷害予測に寄与していると言え、これにより軽自動車と普通乗用車の事故、つまり車両重量差大きい事故における予測性能が向上したと考えられる。図8(b)では、上位2変数は、URGENCYでも使用している変数  $\Delta V$  と年齢であった。その後、本研究で新たに導入した変数が続いている。この結果から、本研究においてこれまで用いられてこなかった重要な特徴量を取り入れたことにより、既存の傷害予測モデル URGENCY よりも高性能なモデルを構築することができたと考えられる。

また、本研究で用いたベイジアンネットワークは予測値がそのまま確率値を表すモデルであるが、 LightGBM を含む2値分類機械学習モデルでは、予測値と実際の確率値に乖離が生じる可能性がある。 そこで2値分類機械学習モデルを用いた傷害予測モデルにおいて、予測値の確率値としての妥当性を確認した。ここでは、機械学習モデルとしてロジスティック回帰、ランダムフォレスト[13]、LightGBMの

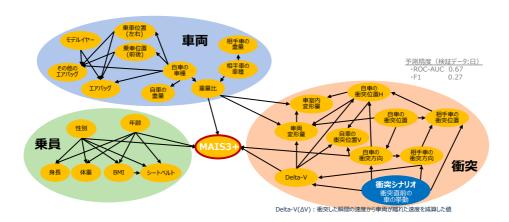


図 6 ベイジアンネットワークにより構築した最大傷害乗員 MAIS3+傷害予測モデルのグラフ構造 学習データ:日米混合データ(混合比率:0.01)

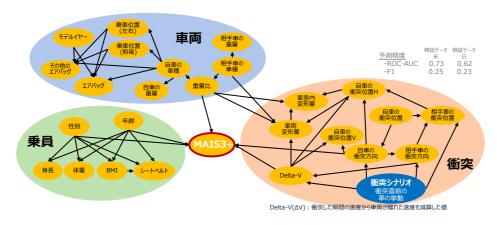
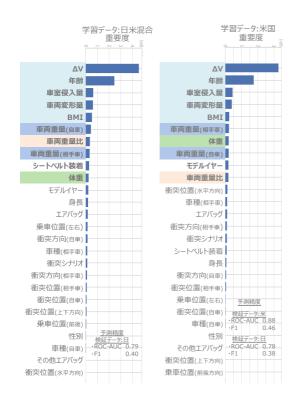
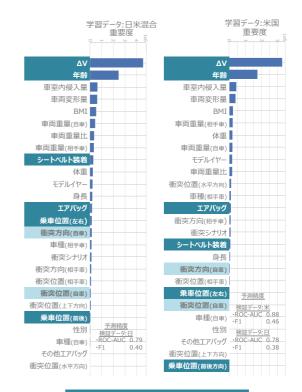


図 7 ベイジアンネットワークにより構築した最大傷害乗員 MAIS3+傷害予測モデルのグラフ構造 学習データ:米国データのみ (図 2 とは、衝突シナリオの分類定義が異なる)





URGENCYで使用している変数
URGENCYで限定的に使用している変数

#### (a) 学習データの違いによる重要度の比較

(b) URGENCY の変数との比較

図8 LightGBM により構築した最大傷害乗員 MAIS3+傷害予測モデルの変数の重要度 (日米混合データの混合比率: 0.01. 米国データは図3とは衝突シナリオの分類定義が異なる.)

3 種類のモデルを使用し、乗員の傷害予測モデルを構築した。データ数確保の点から抽出した全乗員の データ対象とし、目的変数はその乗員が MAIS3+であるか否かとした。その結果、ランダムフォレスト や LightGBM において、適切なデータの重み設定とハイパーパラメータ探索指標の組み合わせにて、高 い分類性能をもち、かつモデルの予測値と実際の確率値がほぼ等しくなるようなモデルの構築に成功し た。

#### 5. 研究成果の活用検討

本研究の成果をさらなる発展研究や社会実装へとつなぐため、本研究で構築した傷害予測モデルや 傷害予測モデル構築までに得られた知見の活用先について検討した。その結果、研究成果の活用先と して7つのアイテムを候補として挙げ、現在、保険会社、自動車会社、ソフトウェア会社等とその詳 細について議論を進めているところである。

#### 6. 結論

米国の大規模ミクロ事故データである NASS-CDS および CISS を統合し、機械学習のためのデータベースを構築した。この事故データベースを用い、ベイジアンネットワークおよび LightGBM 等を実装した機械学習手法により、自動車乗員を対象とした傷害予測モデルを構築した。構築した傷害予測モデルは、従来の線形ロジスティック回帰を用いた傷害予測モデルに対し、ベイジアンネットワークよるモデルでは同等の予測性能を確保しつつ、説明変数が欠損しても傷害予測を可能とし、LightGBM によるモ

デルではより高い予測性能を確保できることを確認した。本研究では従来の傷害予測モデルと比べ多様な特徴量を変数として考慮しており、ベイジアンネットワークによる傷害予測モデルではこれら変数間の因果関係がモデル化され、傷害の程度に直接関係する特徴量のほか、より上位の特徴量との関係もモデル化することができた。また、構築した傷害予測モデルに日本のミクロ事故データを適用し、国内の事故に対する予測性能を確認したところ、軽乗用車・普通乗用車の組み合わせにおいて精度が低下することがわかった。

そこで、国内の事故実態に即した傷害予測モデルを構築するため、米国データに日本データを混合した日米混合データを用い、傷害予測モデルを構築した。これにより米国データのみの学習ではうまく予測することができなかった軽乗用車・普通乗用車の事故に対する予測性能を向上させることができた。データ数の多い米国データを活用しつつ、日本データの混合により米国データには存在しない軽自動車関連の事故データを導入することで、日本の事故実態に即した傷害予測モデルを構築することができたと言える。

また、ベイジアンネットワークのグラフ構造で示された変数(事故データ項目)間の因果関係および LightGBM の変数の重要度から、乗員の傷害予測に寄与する特徴量を明らかにした. 傷害予測に大きく寄与する特徴量として、双方のモデルに共通していたものは  $\Delta V$ 、年齢、車両重量比、車両変形量であった. このうち特に車両重量比は、軽乗用車・普通乗用車の組み合わせといった車両重量比の大きい事故における予測に貢献していると考えられ、国内の事故実態に即した傷害予測モデル構築においては重要な特徴量であることがわかった.

# 7. 知的財産権取得状況

特許出願 1件

発明者:佐藤房子, 宮崎祐介, 高橋啓太, 北村光司

発明の名称: 衝突傷害予測モデル構築方法及び衝突傷害予測方法

出願日:2023年3月1日 出願番号:特願2023-30832

#### 8. 研究成果発表実績

1) 論文発表

国内 0件,海外 2件

- i. Wei J., Miyazaki Y., Kitamura K., Sato F. (2022) Construction of collision-type prediction models based on pre-crash data for advanced driver assistance systems. International Journal of Automotive Engineering, Volume 13 Issue 4 Pages 163-168, DOI: 10.20485/jsaeijae.13.4 163.
- ii. Wei J., Miyazaki Y., Kitamura K., Sato F. (2022) Construction of collision-type prediction model based on light gradient boosting machine algorithm and elucidation of data causality and relevance by Bayesian network. Traffic Injury prevention, 23:sup1, S181-S182, DOI: 10.1080/15389588.2022.2124807.

# 2) 口頭発表

国内 2件,海外 2件

i. Otsuka K., Sakamoto K., Miyazaki Y., Sato F., Kitamura K. (2020) Construction of Injury Prediction Model Using Random Forest Algorithm with Resampling Methods for Imbalanced Accident Data.

- Student Research Symposium of Association for the Advancement of Automotive Medicine (AAAM) 2020.
- ii. 高橋, 宮崎, 北村, 佐藤 (2022) 勾配ブースティング決定木を用いた乗員傷害予測モデル の構築, 自動車技術会 2022 年春季大会学術講演会.
- iii. Wei J., Miyazaki Y., Kitamura K., Sato F. Construction of collision-type prediction models based on pre-crash data for advanced driver assistance systems,自動車技術会 2022 年春季大会学術講演会.
- iv. Wei J., Miyazaki Y., Kitamura K., Sato F. Construction of collision-type prediction model based on light gradient boosting machine algorithm and elucidation of data causality and relevance by Bayesian network, Annual Conference of Association for the Advancement of Automotive Medicine (AAAM) 2022.
- 3) その他(研究内容報告書、機関誌発表、プレス発表等) ・特になし

#### 9. 参考文献

- [1] National Highway Traffic Safety Administration (NHTSA), Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey, 2015.
- [2] National Highway Traffic Safety Administration (NHTSA), National Automotive Sampling System (NASS). <a href="https://www.nhtsa.gov/research-data/national-automotive-sampling-system-nass">https://www.nhtsa.gov/research-data/national-automotive-sampling-system-nass</a>.
- [3] National Highway Traffic Safety Administration (NHTSA), National Automotive Sampling System Crashworthiness Data System 2015 Analytical User's Manual, 2016, DOT HS 812 321.
- [4] National Highway Traffic Safety Administration (NHTSA), National Automotive Sampling System Crashworthiness Data System 2014 Coding and Editing Manual, 2015, DOT HS 812 195.
- [5] National Highway Traffic Safety Administration (NHTSA), Crash Investigation Sampling System (CISS). <a href="https://www.nhtsa.gov/crash-data-systems/crash-investigation-sampling-system">https://www.nhtsa.gov/crash-data-systems/crash-investigation-sampling-system</a>
- [6] National Highway Traffic Safety Administration (NHTSA), Crash Investigation Sampling System 2019 Data Manual, 2020, DOT HS 813 040.
- [7] National Highway Traffic Safety Administration (NHTSA), Crash Investigation Sampling System 2019 Analytical User's Manual, 2020, DOT HS 813 039.
- [8] 本村陽一, 佐藤泰介. ベイジアンネットワーク: 不確定性のモデリング技術. 人工知能学会誌, 15(4), 575-582, 2000.
- [9] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye and T.-Y. Liu, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," Neural Information Processing Systems Conference, 2016.
- [10] Association for the Advancement of Automotive Medicine. The Abbreviated Injury Scale 2015 Revision. Des Plaines, IL: Author; 2016.
- [11] Malliaris A.C., Digges K.H., DeBlois J.H., Relationships between crash casualties and crash attributes, SAE Technical Paper, No. 970393, 1997.
- [12] D.R.Cox, "The Regression Analysis of Binary Sequences," Journal of the Royal Statistical Society, 1958.
- [13] L. Breiman, "Random forests," Machine Learning, Vol. 45, No. 1, pp. 5-32, 2001.