

< 修 士 論 文 >

テレマティクスデータを用いた
ドライバー分類/認識モデルの構築
(要 旨)

滋 賀 大 学 大 学 院
デ ー タ サ イ エ ン ス 研 究 科
デ ー タ サ イ エ ン ス 専 攻

修了年度：2021年度

学籍番号：6020107

氏 名：佐藤 晴紀

指導教員：田中 琢真

提出年月日：2022年1月11日

1. 研究の背景

近年、損害保険会社は、顧客ごとにパーソナライズしたサービスを提供することで、保険契約に付加価値を与えることが求められている。損害保険種目の中で特に重要度が高いのが自動車保険である。自動車保険の特徴は、一台の契約車両に複数名のドライバーがいる可能性があることである。運転特徴の異なる複数名のドライバーを区分しない場合には、適切にパーソナライズされないおそれがある。そのため、顧客ごとにパーソナライズしたサービスを提供するためには、ドライバーを同定する必要がある。

いくつかの保険会社では、契約者に対して自動車に搭載する通信機能付き車載器を提供している。これにより、契約者の走行データを取得し、当該データを活用した様々なサービスを提供している。具体的には、顧客ごとの運転挙動から安全運転アドバイスを行うことや安全運転スコアを算出して保険料に反映させる仕組みを有する商品を販売している。ドライバーを同定して契約者単位からドライバー単位に細分化することにより、これらのサービスをより適正かつ効果的に行うことができる。テレマティクスデータから、ドライバーを同定することができれば、ドライバーを同定するために追加の機器やアプリの機能を必要としない。そのため、本研究においては、テレマティクスデータを活用して、ドライバーを同定することを考える。

2. データの概要

本研究では、ある日本の民間企業の社有車に取り付けられた車載器から取得したテレマティクスデータを用いて、ドライバー分類モデルの構築・評価を行った。走行開始から運転停止までを1単位としてデータが分かれている。取得される走行情報は、観測時点ごとのGPS速度、3軸方向の加速度、GPS進行方向、走行時間、走行距離、GPS位置情報などが測定される。これらのデータは最大15 Hzで観測される。

データの前処理として、入力データの次元を統一するために、トリップごとの観測時間および観測間隔を揃える。道路種別を揃える目的などから、速度は[2, 50] km/h、加速度は[-3, 3] m/s²の範囲に制限した。モデルへの入力は、速度、加速度、進行方向の変化を変換した変数、加加速度を用いる。

3. 分析

本研究ではResNet、ArcFace、AdaCosの3種類の深層学習モデルを実装し、以下の4つの分析を行った。

1つ目の分析は、以降の分析で使用するモデルへの入力時系列データの観測期間と観測間隔の決定方針を策定するための分析である。テレマティクスデータはビッグデータであるため、データ容量を制限するのがこの分析の目的である。具体的には、ランダムに選択され

た 5 人のドライバーの教師あり分類を通じて、観測時間および観測間隔と正答率との関係を調べた。結果としては、データ容量が同じ場合には、観測時間を長くするほど正答率が高くなる傾向が確認された。

2 つ目の分析では、3 つの深層学習モデルの性能の限界確認および性能比較をするために、1,000 人規模のドライバーの分類精度を検証した。結果として、ResNet の正答率は 56.9%、ArcFace の正答率は 51.9%、AdaCos の正答率は 54.8%であった。

3 つ目の分析では、2 つ目の分析で学習した AdaCos を用いて、2 人もしくは 3 人のドライバーの認識が可能であるかを確認した。結果としては、適切な閾値を定めることで、学習データに存在していないドライバーのデータを含めても分類可能であることが分かった。また、最適な閾値はドライバーの組み合わせに大きくは依存していないことが判明した。

4 つ目の分析では、教師なし分類モデルの構築を行った。具体的には、2 つ目の分析で学習したモデルを特徴量ベクトル生成器とし、テストデータから生成された特徴量ベクトルに対する球面 k-means++法により 2 人のドライバーの分類を行った。結果としては、AdaCos を用いることで、教師あり学習モデルと同等の精度となった。

4. 結論と今後の展望

1,000 人規模の分類を通じて、走行データには個人差に関する情報が含まれており、数分間の走行データからドライバーの運転特徴を抽出可能であることが確認された。また、深層距離学習モデルを用いることで、ドライバー認識モデルを構築した。多くのドライバーの組み合わせに対して、ドライバーの認識が可能であることを確認した。さらに、これらの深層距離学習モデルを特徴量ベクトル生成器として用いることで、高精度の教師なし分類モデルを構築できることが示された。

今後の展望としては、教師なし分類モデルにおいても閾値を定めることで想定外のドライバーも含めて判別するモデルを構築することが考えられる。また、運転挙動に影響を与えると考えられる要因を変化させた状態の走行データに対しても、適切に判別が可能であるのかに興味がある。