

# Deep Neural Networks for Classification of Riding Patterns: with a focus on explainability

Milad Leyli abadi<sup>1</sup> and Abderrahmane Boubezoul<sup>2</sup>

1- IRT SystemX, Saclay, France

2- Gustave Eiffel University - TS2, Simu&Moto, France

**Abstract.** The powered two-wheelers (PTW) are among the most vulnerable transport users. It is crucial to identify the appropriate action that should be undertaken during a specific situation to reduce the risk. In this article, the aim is to improve the current state of the art in identification of riding patterns through neural network architectures and to explain how a decision is made by a model which is considered as a black box. In this regard, a new visualization tool specific to time series is suggested to help identify the most influential factors and hopefully to develop appropriate risk mitigation strategies.

## 1 Introduction

In past years, road safety has become a priority for the governments of European countries and abroad. While road safety in the EU has improved considerably in recent decades, the number of deaths and injuries is still unacceptably high. While statistics show a substantial decline in the number of fatalities on the road for four-wheeled vehicles, in the category of Powered Two-Wheelers (PTWs), the statistics show only a minor reduction with a constant decrease (approximately 2%), in France, for instance<sup>1</sup>.

Road accidents are events that can be based on complex and hardly predictable mechanisms. However, it has been shown in the literature that driver behavior and driver errors are major causes of vehicular accidents [1]. Therefore, observing and understanding driver behavior has attracted much attention from researchers. Many studies have been made to determine what factors are associated with critical events that occur before crashes. In the context of naturalistic riding study (NRS), a huge amount of data is collected which provides useful information concerning the rider's behavior, i.e., how the rider interacts with the vehicle and other road users.

The machine learning techniques are used widely to determine driving styles characterization [2]. So far, applications of deep learning approaches to Rider behavior data analysis are limited compared to Driver data analysis. Recently, in [3], the authors give an overview of the use of the Deep Learning algorithms to learn driving style from trajectory data.

Inspired by our recent work [4], in this article we adopt a two-step methodology for identification of riding patterns and we experiment several recurrent neural network architectures and methods. In order to understand the decision made by the models, we propose to analyze the latent representation obtained

---

<sup>1</sup><https://www.onisr.securite-routiere.gouv.fr/>

by an autoencoder. Furthermore, the Local Interpretable Model-agnostic Explanations (LIME) [5] is used to provide a global understanding of the model through an individual prediction. Finally, a visualization tool is proposed to better understand the features impact when analyzing time series data.

## 2 Data

The data analyzed in this study is issued from a natural riding study comprising 11 riders. Each rider performs a different number of riding experience (3 in average). The motorbikes are equipped with two types of sensors (a gyroscope and an accelerometer) which capture the information concerning the riding dynamics (with 10 Hz frequency), e.g., accelerations alongside different axis and also lateral and longitudinal movements (see Figure 1). An extract of the features for one of the rider is also shown in this figure.

The annotations for each temporal point in this dataset are provided. It comprises an ensemble of 26,310 points with 65% of straight segments, 6% of left turns points, 16% of right turns points and 11% of roundabouts points. Each temporal point  $\mathbf{x}^{(t)} \in \mathbb{R}^6$  is described by  $\mathbf{x}^{(t)} = \{A_x^{(t)}, A_y^{(t)}, A_z^{(t)}, R_x^{(t)}, R_y^{(t)}, R_z^{(t)}\}$ . As it can be noticed, we encounter the imbalanced data challenge, with straight segment and left turn patterns as majority and minority classes respectively.

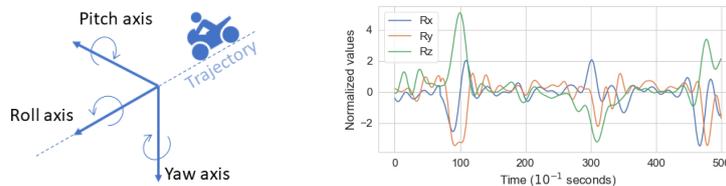


Fig. 1: 3D gyroscope features in left and their values extracted over 50 seconds

**Descriptive analysis.** In order to study the riding situations in greater detail, the joint distributions of two features, namely the lateral acceleration ( $A_y$ ) and yaw axis ( $R_z$ ), are traced for riding patterns (see Figure 2). These joint distribution highlight clearly the difference between the four riding situations. The left and right turns display opposite behaviors while the roundabout patterns is a fusion of both (more similar distribution to that of left turns).

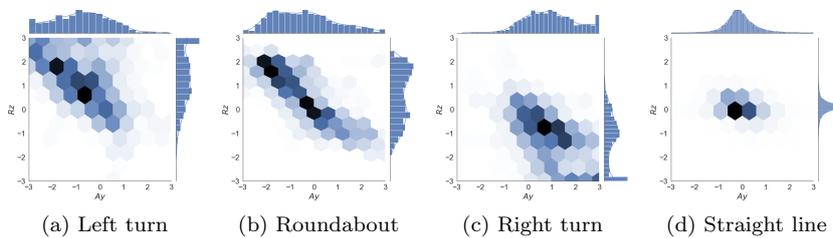


Fig. 2: Joint distribution of 4 riding situations with respect  $A_y$  and  $R_z$

### 3 Methodology

In order to classify the four previously mentioned riding patterns, we have adopted a two-step approach. First, the temporal points are segmented through an overlapped window. The reason behind this segmentation step is that the riding situations takes at least few seconds over time to be performed. As the result from this step, we obtain a set of  $T$  overlapped segments  $\mathbf{S} = \{S_1, S_2, \dots, S_T\}$  where  $S_t = \{\mathbf{x}^{(t)}, \dots, \mathbf{x}^{(t+w)}\}$  with the window length  $w = 50$  (5 seconds).

The resulted segments  $S$  are then used as the input of a classification module  $f$  as follows:

$$\hat{y} = f(S, \theta); \quad \text{with} \quad \hat{y} \in \begin{cases} SL & \text{Straight Line} \\ LT & \text{Left Turn} \\ RT & \text{Right Turn} \\ RA & \text{Roundabout} \end{cases}, \quad (1)$$

where  $\theta$  is the model parameters to be estimated. As the riding patterns are issued from the successive patterns observed over time, the temporal dependency between the points inside each segment should be taken into account for a better identification of such patterns. Two recurrent neural network architectures are considered for the function  $f$ , which are briefly explained in the following:

- Long shot-term memory network (LSTM) : A recurrent neural network architecture with three gates which allows the information persistence over time through a cell state;
- Gated Recurrent Unit (GRU) : A simpler recurrent neural network architecture than LSTM including two gates and without the memory cell. As result, it requires less parameters to be estimated and is more adapted for small datasets.

To explain the decision made by these models, a Deep Temporal Autoencoder (DTA) is used to visualize the learned latent representations. This is inspired by [6] which is used basically for unsupervised temporal data clustering. The DTA's scheme is shown in Figure 3. The latent temporal representation of segments is learned in unsupervised fashion using an autoencoder consisting of recurrent layers (LSTM). It should be noted that this approach may also be used directly for classification purpose. In this article, it is used to demonstrate the underlying representation.

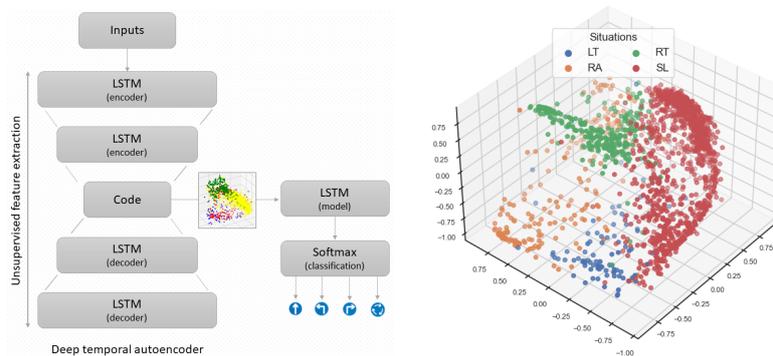


Fig. 3: Deep Temporal Autoencoder (DTA) and its embeddings

## 4 Results

In the following, we present the experimental setup based on which the results are obtained. Finally, we try to explain the models through an example and using different tools.

**Experimental setup.** The experimentation is done using leave-one-out sequence cross validation approach. Alongside the methods based on neural networks, we have also evaluated the Gradient Boosting (GB) approach. To fine-tune the hyper parameters of each of the previously mentioned methods, we have used a grid search over a set of parameters. For more information, the readers may refer to [4]. Various evaluation criteria are also used to compare between the approaches which are F-measure, Cohen’s kappa and accuracy expressed alongside its standard deviation over several runs of experiments.

**Riding patterns embeddings and classification.** In order to show the difficulty of the riding patterns classification, the latent representation obtained using DTA is shown in Figure 3. As it can be seen, the frontier between some patterns is ambiguous, e.g., the LT pattern’s embeddings are very close to those of the SL and RA patterns. The finding that could be also expected from the analysis in section 2. This phenomena could be explained by the fact that each riding pattern is constituted of multiple segments (for example a roundabout involves different manoeuvres such as right turn, straight segment and left turn) where each segment is represented by a set of points over time and could be similar between patterns. In the following section, one of these ambiguous patterns is used to explain the model decision.

The performances of three previously mentioned methods are also compared using various criteria in Table 1 with best performances highlighted in bold. It can be observed that recurrent neural networks work better than GB which does not consider the temporal nature of riding patterns. The performance of LSTM and GRU remains very close with a relatively smaller number of parameters for GRU (implying less computation time). It should also be noted that the performances are lower for LT patterns (minority class) which are very close to SL and RA patterns in the latent space (see Figure 3).

Table 1: Comparison of the performances for the riding patterns classification task

	F-measure per class				F-measure	$\kappa$	Accuracy ( $R$ ) $\pm$ ( $std$ )	#parameters
	SL	LT	RT	RA				
GB	92.2	56.3	80.7	73.8	86.2	0.74	88.6 $\pm$ 2.9	100 estimators
LSTM	<b>94.9</b>	<b>67.8</b>	87.5	<b>86.3</b>	89.4	<b>0.83</b>	<b>93.1</b> $\pm$ 3.1	123,604
GRU	94.8	65.9	<b>88.6</b>	85.6	<b>91.5</b>	0.83	91.5 $\pm$ .03	93,404

The predictions of the LSTM model are also compared visually with the ground truth labels for one of the sequences (see Figure 4). The third graphic in this figure shows the Softmax layer probabilities used to highlight the ambiguous cases. Most of the situations are correctly identified. A period of ambiguous decisions are highlighted by a dashed rectangle, which concerns a roundabout situation. The RA situations are particularly challenging as they are constituted of multiple manoeuvres similar to other patterns.

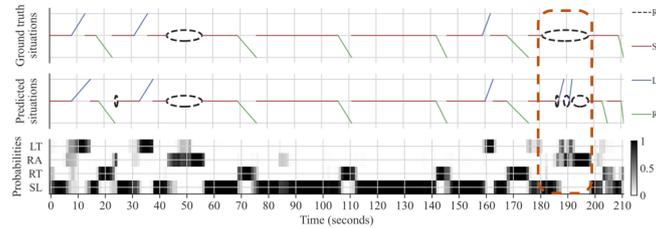


Fig. 4: A riding sequence alongside the predictions and probabilities obtained using LSTM **Model and predictions explanation.** This section provides explanations concerning the RA situation highlighted in Figure 4 using the LIME method [5]. More precisely, the first wrongly identified left pattern is considered for the sake of this analysis, and we call it *test segment* in the following. The corresponding features for this test segment and prediction probabilities are shown in the bottom part of the Figure 5. In order to better understand the decision made by the LSTM model, the mean behaviors of different features for each riding pattern are also shown in the top part of this figure. The confusion concerning this segment can be observed through the assigned probabilities.

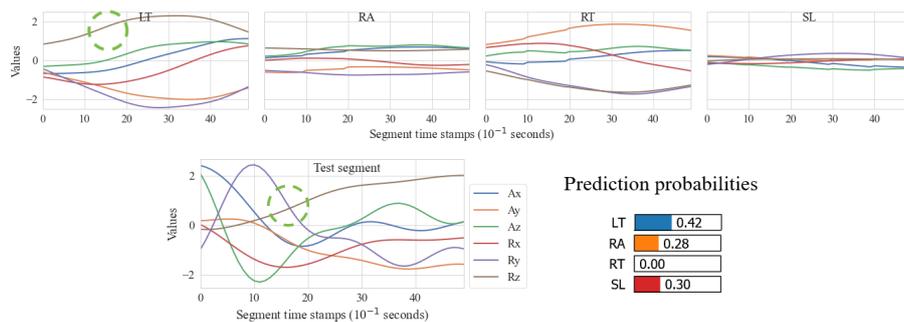


Fig. 5: Averaged features for each riding pattern in the top part. A test segment (187th segment inside the dashed rectangle in Figure 4) which is a RA segment and is misidentified as LT by the model and for which we search for an explanation

The local explanations obtained using LIME for the test segment are shown in Figure 6. The categorized features are shown on y-axis of these figures, indicating a relation at a fixed point in time. As an example, ‘ $R_z_{t-16} > 0.27$ ’ indicates that the feature  $R_z$  at time stamp 16 ( $10^{-1}$  second) with a value greater than 0.27 contributed as the length of the corresponding red (negative contribution) or green bar (positive contribution).

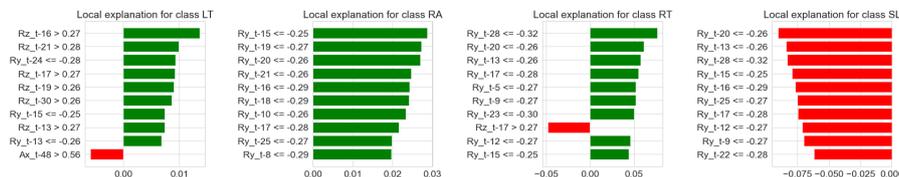


Fig. 6: Local explanation using Lime for the test segment depicted in Figure 5

It can be observed that the  $R_z$  feature which is highly related to handlebar position, shows a high contribution at time point 16 for the LT pattern. This

can be confirmed by the increasing behavior of this feature at time stamp 16 for the test segment and its similarity with the averaged  $R_z$  for LT pattern (see green encircled areas in Figure 5). On the other hand, the various instants of  $R_y$  feature are contributed to identifying the test segment as RA and SL.

Although the LIME method gives some local insights about the decision of the model, it becomes cumbersome to trace the impact of features for long time series. For this reason, we propose a new visualization tool based on the values of hidden units and model weights for each feature over time which are shown in second and third graphics of Figure 7. The global relation between features and riding patterns could be observed (e.g.,  $R_z$  feature with RT and SL).

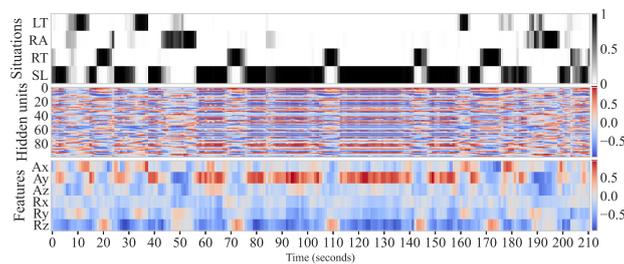


Fig. 7: Proposed visualization tool for explanation of time series data

## 5 Conclusion and perspectives

In this article, a two-step methodology for the classification of riding patterns from time series data is proposed. The recurrent neural network based methods have shown good performances for this task. Furthermore, the model's decision is explained locally and at a global level. In order to overcome the imbalanced class challenge encountered in this work, as an insight for future works, we investigate the use of generative methods as GAN and VAE.

## References

- [1] Luca Mallia and al. Crash risk and aberrant driving behaviors among bus drivers: the role of personality and attitudes towards traffic safety. *Accident Analysis & Prevention*, 79:145–151, 2015.
- [2] C. Marina Martinez and al. Driving style recognition for intelligent vehicle control and advanced driver assistance: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 19(3):666–676, 2018.
- [3] Sobhan Moosavi and al. Driving style representation in convolutional recurrent neural network model of driver identification. *CoRR*, abs/2102.05843, 2021.
- [4] Milad Leyli-abadi and al. Riding pattern recognition for powered two-wheelers using a long short-term memory network. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–7, 2020.
- [5] Marco Tulio Ribeiro and al. "why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144, 2016.
- [6] Naveen Sai Madiraju, Seid M Sadat, Dmitry Fisher, and Homa Karimabadi. Deep temporal clustering: Fully unsupervised learning of time-domain features. *arXiv preprint arXiv:1802.01059*, 2018.