

Classification de séries temporelles basée sur des “shapelets” interprétables par réseaux de neurones antagonistes

Yichang Wang¹, Rémi Emonet², Elisa Fromont¹, Simon Malinowski¹, Etienne Menager¹,
Loïc Mosser¹, et Romain Tavenard³

¹Univ. Rennes, Inria, CNRS, IRISA, Rennes, France

²Univ. Lyon, UJM Saint-Etienne, CNRS, Institut d’Optique Graduate School, Laboratoire Hubert Curien UMR 5516, Saint-Etienne, France

³Univ. Rennes, LETG, IRISA, Rennes, France

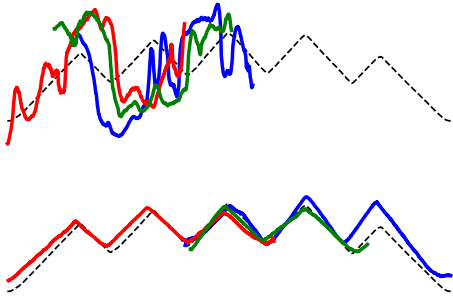


Figure 1: Exemple d’une série (en pointillés noirs) pour le jeux de données HandOutlines sur laquelle sont surimposées les trois shapelets les plus discriminantes pour cette série apprises avec la méthode de Grabocka et. al. [GSWST14] (en haut) et pour notre méthode AI \leftrightarrow PR-CNN (en bas).

Résumé étendu

Une série temporelle (TS) est une suite de données, chacune décrite par un vecteur d’attributs, ordonnées dans le temps. Dans cet article, nous nous intéressons à la classification de séries temporelles univariées (i.e. le vecteur d’attribut est de dimension 1). Etant donné un ensemble de séries temporelles, chacune associée à une classe, notre but est d’apprendre un modèle qui pourra prédire la classe d’une nouvelle série temporelle.

Ce problème a de nombreuses applications (voir par exemple [SS05], [HRH⁺] ou [HCL⁺18]) allant du domaine de la bourse à celui de l’optimisation de consommation électrique en passant par le domaine médical, le domaine du jeu vidéo, la télé surveillance, etc.

De nombreuses méthodes ont été proposées dans la littérature pour le résoudre (voir see [BLB⁺17] pour un état de l’art). Un groupe de méthodes, qui obtient de bonnes performances en classification, s’attache à représenter une série en se basant sur sa similarité à des sous-séries discriminantes appelées “shapelets” (appries ou non) puis à classifier automatiquement la série en fonction de cette nouvelle représentation [YK09, RK13, LDHB12, GSWST14]. Les premiers travaux à ce sujet [YK09, RK13] proposaient d’énumérer les “shapelets” possibles directement à partir des séries existantes. Même si ces méthodes sont trop complexes ou très peu précises, elles ont l’avantage de fournir un critère interpretable direct (la shapelet est une sous partie d’une série existante) permettant d’expliquer la décision d’un classifieur pour un exemple donné.

Les travaux qui ont suivi se sont attachés à apprendre conjointement les shapelets et le modèle qui permet d’obtenir, efficacement, de bonnes performances en classification [LDHB12, GSWST14]. Malheureusement, ces méthodes ne garantissent plus le caractère “interpretable” des shapelets apprises comme le faisaient les premières méthodes citées puisqu’il n’y a plus de lien direct entre les shapelets et les séries de l’ensemble d’entraînement. Un exemple de série et de shapelet apprise par la méthode proposée par Grabocka et. al. et par la méthode que nous proposons dans cet article est donné Figure 1 : il est visuellement beaucoup plus difficile, dans le cas de Grabocka et. al., d’expliquer pourquoi le classifieur, en se basant sur ces shapelets, a proposé une décision de classification particulière. On peut noter que le même type de problème d’interprétabilité intervient avec des ensembles de classifieurs [BLHB15] pour lesquels la décision finale, même

si elle est très précise, est expliquée par la présence ou l'absence de trop nombreuses shapelets. Nos considérations d'interprétabilité interviennent dans un contexte Européen (GDPR) et mondial (Darpa XAI) où il est de moins en moins accepté de laisser des classifieurs "boîtes noires" comme les réseaux de neurones", aussi performants soient-ils, prendre des décisions opaques sur des problématiques critiques [GMR⁺18]. Il nous paraît donc important de fournir des méthodes qui pourront apporter une explication à la décision prise.

Nous proposons une nouvelle architecture, nommée AI \leftrightarrow PR-CNN, composée de réseaux de neurones convolutifs (CNN) antagonistes qui adresse ce problème. Notre réseau possède deux composants: un classifieur et un discriminateur. Le classifieur est un CNN, il sert à classer des séries. Les convolutions sont des motifs discriminants appris à partir des données qui permettent une représentation des séries temporelles plus discriminante (de manière similaire à celle des shapelets présentée initialement). Pour pouvoir expliquer la décision du classifieur, nous voudrions imposer que les convolutions utilisées soient des vraies "shapelets" c'est à dire qu'elles soient proches de sous-série réelles présentes dans l'ensemble d'entraînement. Cette contrainte est mise en oeuvre par le réseau antagoniste dont le but sera de déterminer à quel point les matrices de poids des convolutions du classifieur sont proches de sous-séries de l'ensemble d'entraînement. La Figure 1 montre que notre méthode permet de satisfaire la contrainte d'interprétabilité que nous nous sommes fixés et notre architecture fournit des résultats de classification comparables aux résultats de l'état de l'art.

Mots-clef: Série temporelle, shapelets, modèle interprétable, réseaux de neurones antagonistes.

References

- [BLB⁺17] Anthony Bagnall, Jason Lines, Aaron Bostrom, James Large, and Eamonn Keogh. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data Mining and Knowledge Discovery*, 31(3):606–660, May 2017.
- [BLHB15] Anthony Bagnall, Jason Lines, Jon Hills, and Aaron Bostrom. Time-series classification with cote: the collective of transformation-based ensembles. *IEEE Transactions on Knowledge and Data Engineering*, 27(9):2522–2535, 2015.
- [GMR⁺18] Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. A survey of methods for explaining black box models. *ACM Comput. Survey*, 51(5), 2018.
- [GSWST14] Josif Grabocka, Nicolas Schilling, Martin Wistuba, and Lars Schmidt-Thieme. Learning time-series shapelets. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data mining*, pages 392–401, 2014.
- [HCL⁺18] Kyle Hundman, Valentino Constantinou, Christopher Laporte, Ian Colwell, and Tom Soderstrom. Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '18, pages 387–395, New York, NY, USA, 2018. ACM.
- [HRH⁺] Awni Y. Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H. Tison, Codie Bourn, Mintu P. Turakhia, and Andrew Y. Ng. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. 25(1):65–69.
- [LDHB12] Jason Lines, Luke M Davis, Jon Hills, and Anthony Bagnall. A shapelet transform for time series classification. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data mining*, pages 289–297, 2012.
- [RK13] Thanawin Rakthanmanon and Eamonn Keogh. Fast shapelets: A scalable algorithm for discovering time series shapelets. pages 668–676, 05 2013.
- [SS05] Robert H. Shumway and David S. Stoffer. *Time Series Analysis and Its Applications (Springer Texts in Statistics)*. Springer-Verlag, Berlin, Heidelberg, 2005.
- [YK09] Lexiang Ye and Eamonn Keogh. Time series shapelets: a new primitive for data mining. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data mining*, pages 947–956, 2009.