Contribution ID: **91** Type: **not specified**

Shapelet learning via sparse feature selection for interpretable time series classification

Time-series classification faces recurring challenges, including high dimensionality, autocorrelation, and the difficulty of identifying features that capture essential dynamics across temporal scales and phase shifts. We address these issues through shapelet decomposition, a technique that extracts shape-based features from time series while preserving both temporal and frequency information. The approach represents a dataset through minimal distances to representative patterns, thereby improving interpretability by highlighting the patterns most relevant to classification decisions.

In this framework, selecting the most informative shapelets, or equivalently an effective projection basis, is critical for both performance and interpretability. Building on prior work that optimizes shapelet values to maximize classifier accuracy, we propose a formulation based on sparse feature selection. Rather than jointly optimizing shapelets as classifier parameters, our approach seeks a data transformation that produces covariates that are highly predictive of the class label while being weakly dependent, thereby reducing redundancy and improving classifier stability and interpretability.

To this end, shapelets are determined by optimizing a HSIC-Lasso criterion on the transformed dataset. HSIC-Lasso is a nonlinear feature-selection method for high-dimensional settings that identifies a sparse subset of features that are highly informative about the response while exhibiting very limited or null redundancy. To make the optimization tractable, the non-smooth minimum-distance operator between shapelets and time series is replaced by a soft-minimum approximation, yielding a differentiable loss function.

The proposed method is validated on both synthetic and real-world datasets, demonstrating improvements in classification performance, scalability, and interpretability.

Special/ Invited session

FrENBIS

Classification

Both methodology and application

Keywords

interpretable machine learning, time series classification, shapelet transformation, shapelet learning, sparse feature selection

Primary authors: PELAMATTI, Julien (EDF R&D); BOUSQUET, Nicolas (EDF); Mr SEYDI, Ibrahim (EDF R&D)

Presenter: PELAMATTI, Julien (EDF R&D)

Track Classification: Other/special session/invited session