The ROCKET Algorithm: From Classification to Prediction

Lila ROIG, Eva ROBILLARD

MVA 2024-2025
Machine Learning for Time Series
Final Project Presentation

école normale supérieure paris–saclay

MATHÉMATIQUES VISION APPRENTISSAGE

Introduction

ROCKET (for RandOm Convolutional KErnel Transform) [1]

↪ Method for univariate time series classification.

↪ Provides a fast and accurate solution to computationally expensive classification methods (i.e., InceptionTime)

Objectives of this project

- **Verification:** Evaluate ROCKET's performance on its original task – classification – using a new dataset.
- **Adaptation:** Extend ROCKET to:
  - Multivariate data.
  - A novel task – **time series prediction** – and evaluate its performance.

[1] Dempster, Petitjean, and Webb 2020

The ROCKET Algorithm for Classification

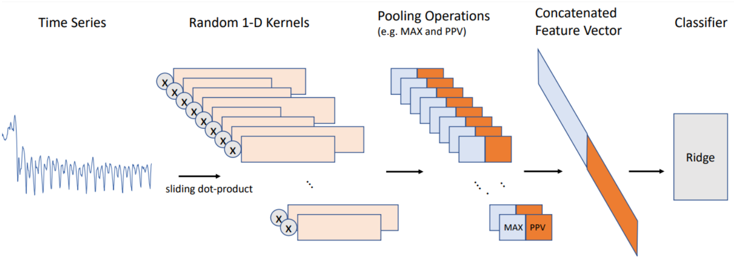


Figure: The steps of the ROCKET algorithm

Results on the OliveOil Dataset Classification

↪ OliveOil dataset: 60 time series, all univariate, each serie contains 570 time points.

↪ Validation of ROCKET: accuracy of **88.889%**.

Adapting ROCKET to Multivariate Time Series

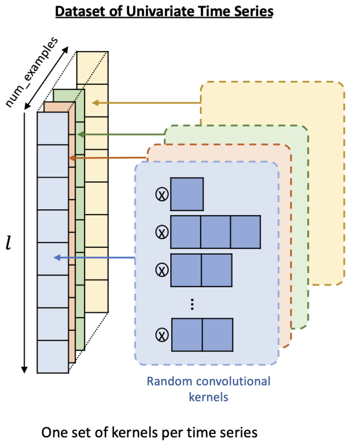


Figure: ROCKET applied to univariate time series for forcasting

Adapting ROCKET to Multivariate Time Series

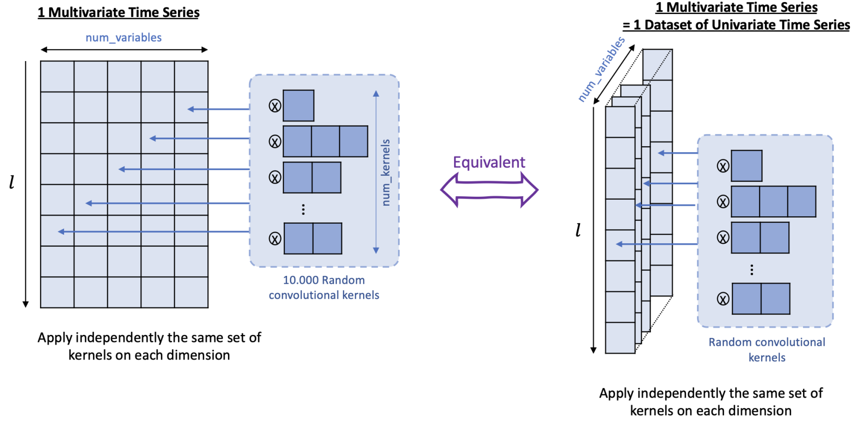


Figure: ROCKET applied to multivariate time series for forecasting

Adapting ROCKET to Time Series Forecasting

- Adapting Rocket for forecasting: replace the **classifier** by a **regressor** that takes the computed features as input and predicts values for the chosen **forecast horizon** H.
- **Step 1** : Sliding Windows

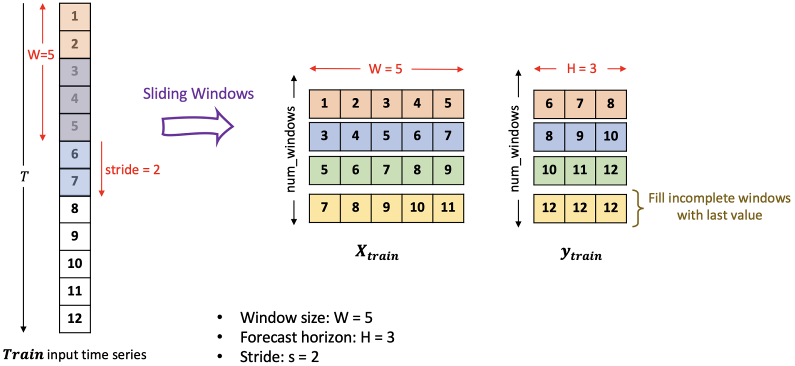


Figure: Sliding windows mechanism

Adapting ROCKET to Time Series Forecasting

- Apply sliding windows to each input time series

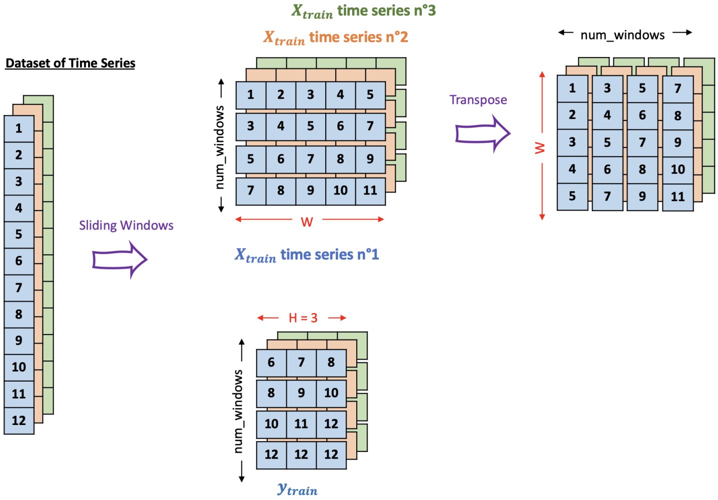
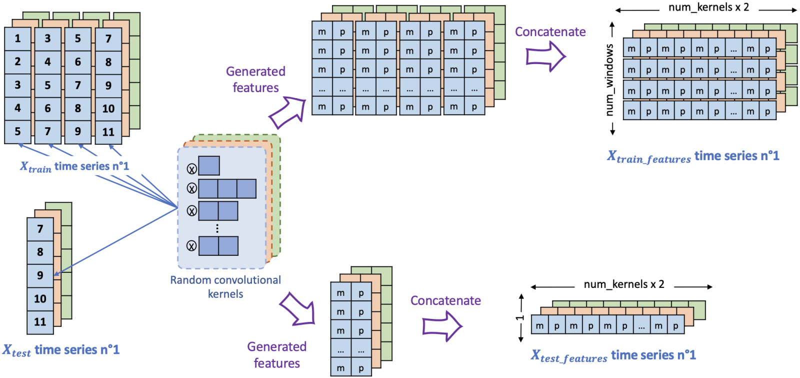


Figure: Apply sliding windows to the time series dataset

Adapting ROCKET to Time Series Forecasting

- **Step 2** : Feature extraction with ROCKET
  - Generate random convolutional kernels and apply them to X_{train}.
  - Each kernel outputs *Max* and *PPV*.
  - Use the last window of X_{train} as X_{test}.
  - Apply the same kernels to X_{test}.



Adapting ROCKET to Time Series Forecasting

- **Step 3** : Train a regressor on the computed features
  - Train a regression model on $X_{\text{train\_features}}$ and y_{train}.
  - Apply the trained model to $X_{\text{test\_features}}$.
  - Generate predictions for the specified forecast horizon H.

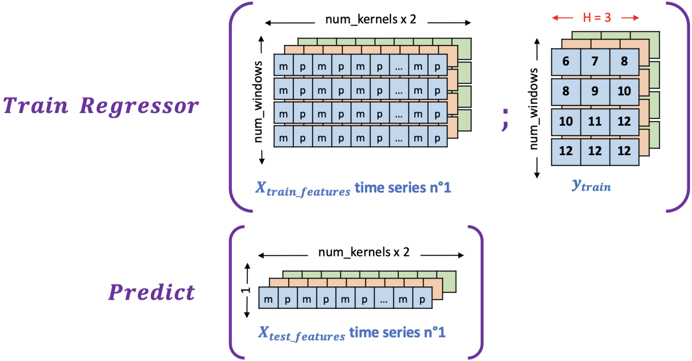


Figure: Apply Regressor and make prediction

Tested regressor: RidgeCV, XGBoost, Random Forest

Datasets for prediction: the Monash Archive

Overview of the Monash Archive

58 datasets (univariate and multivariate) spanning various sampling rates (yearly, quarterly, monthly, etc.). Some series have missing values. 30 datasets were used.

Dataset	**Domain**	**Nb.Series**	**Min.Len.**	**Max.Len.**	**Freq.**	**Multivariate**
M1	Multiple	1001	15	150	3	No
M4	Multiple	100000	19	9933	6	No
Bitcoin	Economy	18	2659	4581	1	No
Elec.	Energy	321	26304	26304	2	Yes

Table: Sample of datasets in the Monash Time Series Forecasting Archive

How are the datasets used for prediction ?

Forecast Horizon (H):

- Determined based on the frequency of each time series (as per Monash paper recommendations [2]).

Sliding Windows (W):

- Size of sliding windows corresponds to the **number of lags (past values)**.
- Determination of lags:
  - ↪ Based on seasonality **multiplied by 1.25**.
  - ↪ For short time series: $lags = forecast\ horizon \times 1.25$.
  - ↪ Adjustments for specific datasets, e.g., **hourly data** (using daily seasonality to reduce memory requirements).
- Minimum number of lags = forecast horizon (ensures sufficient past data).

[2] Godahewa et al. 2021

Results on forecasting: metrics

Metrics used

$\hookrightarrow$ **MASE (Mean Absolute Scaled Error):** Accounts for scale invariance and seasonality.

$\hookrightarrow$ **MSE (Mean Squared Error):** Penalizes large deviations.

$\hookrightarrow$ **RMSE (Root Mean Squared Error):** Provides interpretability in the original scale.

$\hookrightarrow$ **sMAPE (Symmetric Mean Absolute Percentage Error):** Measures relative error.

Results on forecasting: by frequency.

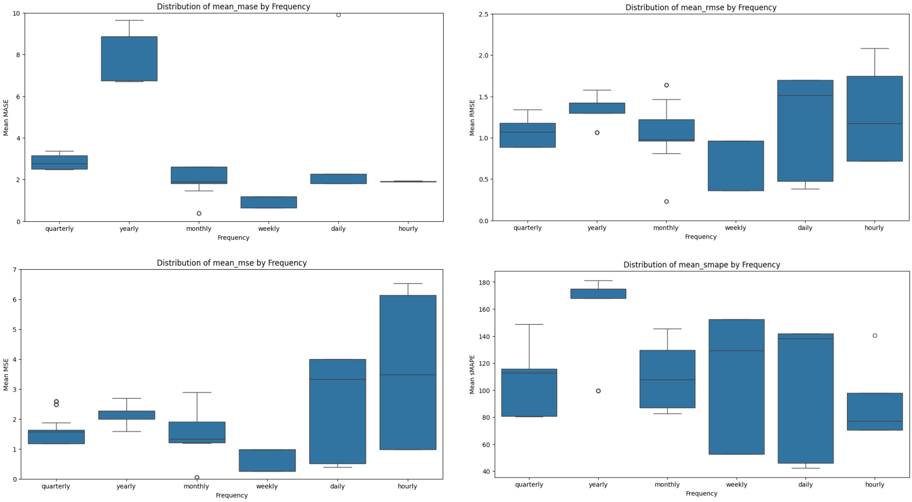


Figure: Metrics distribution by frequency

Results on forecasting: by regressor.

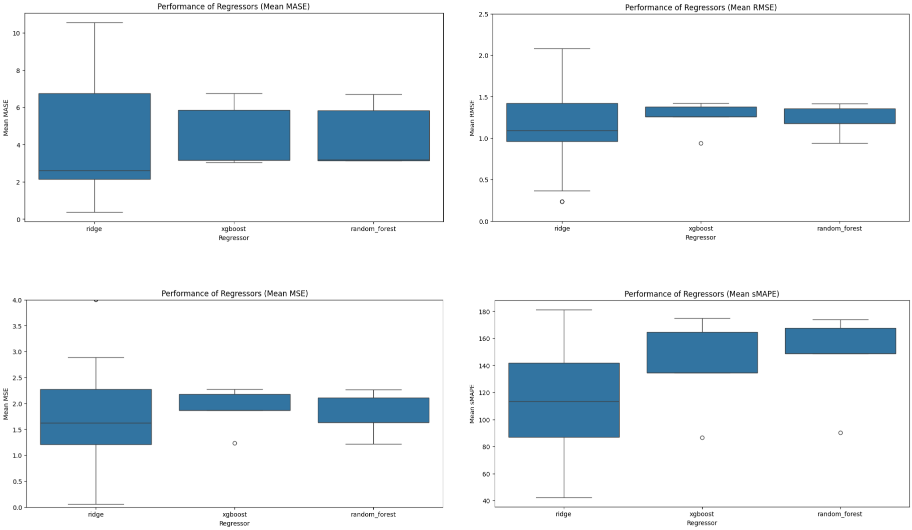


Figure: Metrics distribution by regressor

Results on forecasting: by dataset type.

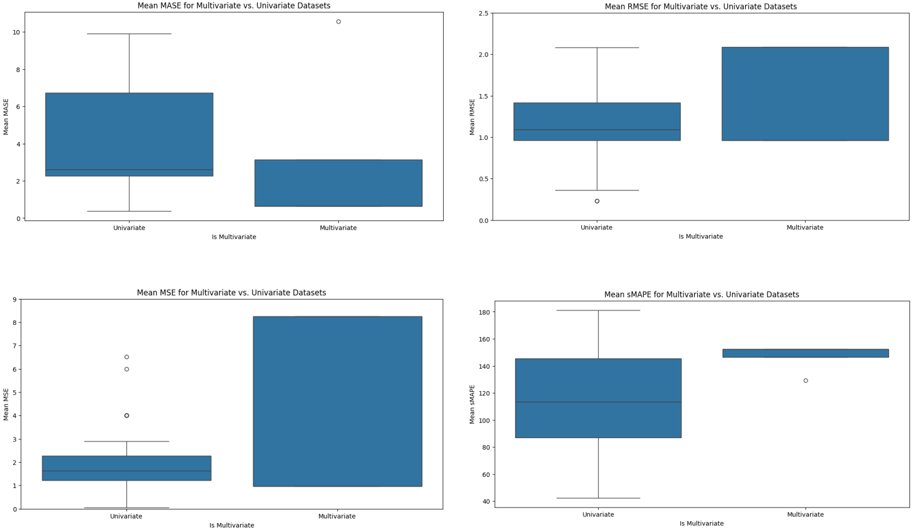


Figure: Metrics distribution for univariate and multivariate datasets

Thank you for your attention!

Bibliography I

Dempster, Angus, François Petitjean, and Geoffrey I Webb (2020). "ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels". In: *Data Mining and Knowledge Discovery* 34.5, pp. 1454–1495.

Godahewa, Rakshitha et al. (2021). "Monash Time Series Forecasting Archive". In: *Neural Information Processing Systems Track on Datasets and Benchmarks*.