

Enhancing Time Series Classification with Diversity-Driven Neural Network Ensembles

Javidan Abdullayev, Maxime Devanne, Cyril Meyer, Ali Ismail-Fawaz,
Jonathan Weber, Germain Forestier

IRIMAS, Université de Haute-Alsace

July 2, 2025



① Introduction and Literature Review

② Proposed Method

③ Experimental Evaluation

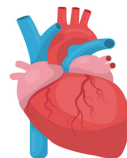
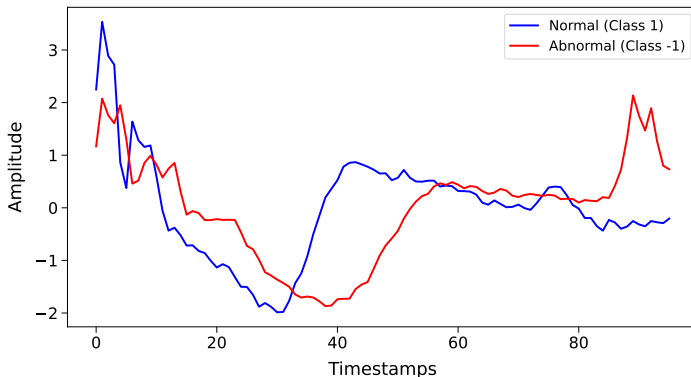
1 Introduction and Literature Review

2 Proposed Method

3 Experimental Evaluation

Time Series Data

Examples on ECG200 Dataset

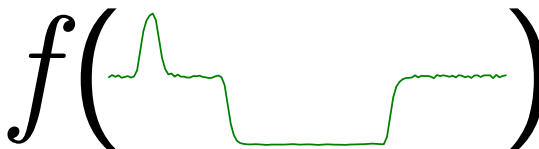


1

¹Dau, Hoang Anh, et al. "The UCR time series archive." IEEE/CAA Journal of Automatica Sinica 6.6 (2019): 1293-1305.

Time Series Classification (TSC)

Time Series Classification Task



- ▷ f is a classifier
- ▷ f assigns a label y to an input sample
- ▷ $y \in \{0, \dots, C - 1\}$, with the C number of classes

Deep Learning approaches for TSC

▷ Deep Learning (DL) for TSC

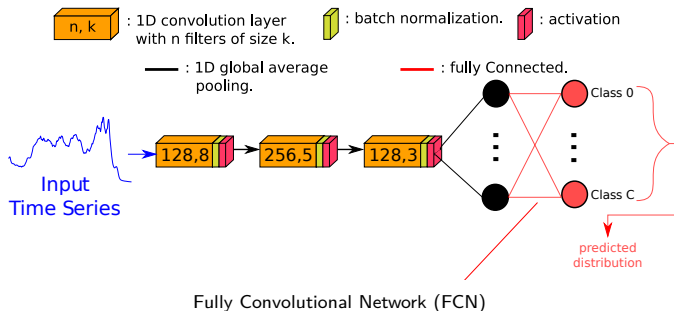
Best performing DL approaches for TSC based on Convolutional Neural Networks

Reference: Middlehurst, Matthew, Patrick Schäfer, and Anthony Bagnall. "Bake off redux: a review and experimental evaluation of recent time series classification algorithms." *Data Mining and Knowledge Discovery* 38.4 (2024): 1958-2031.

Deep Learning Approaches for TSC

▷ Deep Learning (DL) for TSC

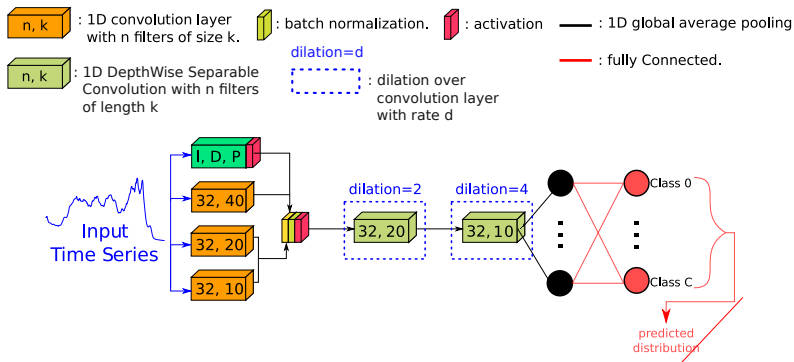
Best performing DL approaches for TSC based on Convolutional Neural Networks



Reference: Ismail Fawaz, Hassan, et al. (2019). "Deep Learning for Time Series Classification: A Review." Data Mining and Knowledge Discovery, 33(4), 917–963.

- ▷ Deep Learning (DL) for TSC

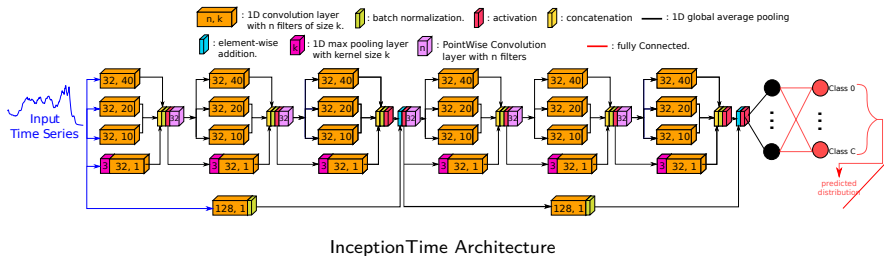
Best performing DL approaches for TSC based on Convolutional Neural Networks



Lightweight InceptionTime Ensemble (LITE)

Reference: Ismail Fawaz, Hassan, et al. (2019). "Deep Learning for Time Series Classification: A Review." *Data Mining and Knowledge Discovery*, 33(4), 917-963.

Best performing DL approaches for TSC based on Convolutional Neural Networks



Reference: Ismail Fawaz, Hassan, et al. (2019). "Deep Learning for Time Series Classification: A Review." *Data Mining and Knowledge Discovery*, 33(4), 917-963.

Motivation

- ▷ Current state-of-the-art deep learning-based TSC models are ensemble-based:
 - ▷ Inception**Time**
 - ▷ H-Inception**Time**
 - ▷ LIT**Time**

Motivation

- ▷ Current state-of-the-art deep learning-based TSC models are ensemble-based:
 - ▷ Inception**Time**
 - ▷ H-Inception**Time**
 - ▷ LIT**Time**

- ▷ Why do ensembles often perform better:
 - ▷ Feature diversity
 - ▷ Filter diversity

Motivation

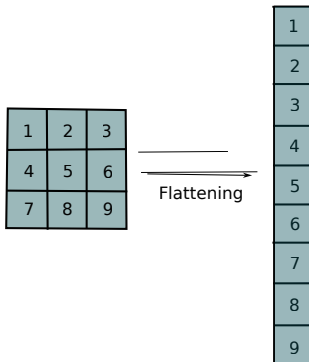
- ▷ Current state-of-the-art deep learning-based TSC models are ensemble-based:
 - ▷ Inception**Time**
 - ▷ H-Inception**Time**
 - ▷ LIT**ETime**

- ▷ Why do ensembles often perform better:
 - ▷ Feature diversity
 - ▷ Filter diversity

- ▷ **Can we improve the efficiency of ensemble models while enhancing diversity among their members?**

Orthogonality in Vision Models

- ▷ Some vision models apply orthogonality constraints on filter weights (Yang et al.; Ayinde et al.; Wang et al.).

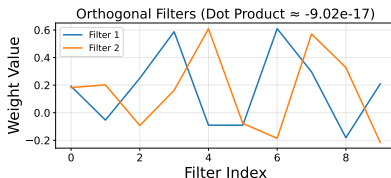


References:

- Yang et al. *Orthogonality loss: Learning discriminative representations for face recognition* IEEE Trans. Circuits Syst. Video Technol., 31(6), 2301–2314, 2020.
- Ayinde et al. *Regularizing deep neural networks by enhancing diversity in feature extraction* IEEE Trans. Neural Netw. Learn. Syst., 30(9), 2650–2661, 2019.
- Wang et al. *Orthogonal convolutional neural networks*. CVF, 2020.

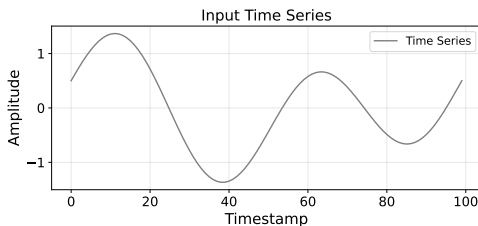
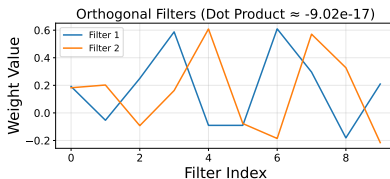
Rethinking Diversity

- ▷ By shifting filter position it is possible to achieve orthogonality



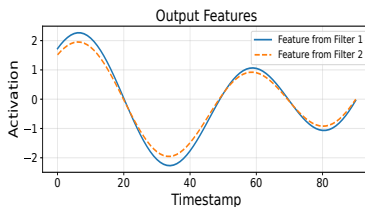
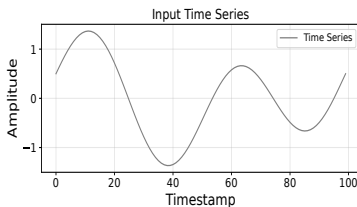
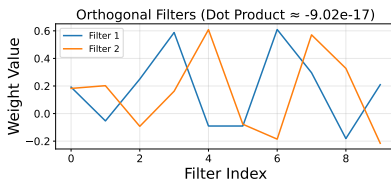
Rethinking Diversity

- ▷ By shifting filter position it is possible to achieve orthogonality



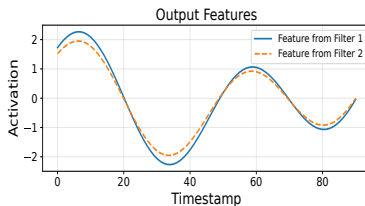
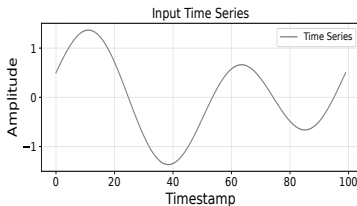
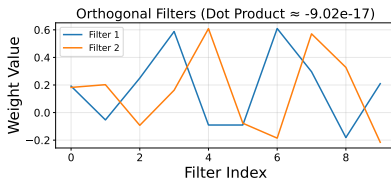
Rethinking Diversity

- ▷ By shifting filter position it is possible to achieve orthogonality



Rethinking Diversity

- ▷ By shifting filter position it is possible to achieve orthogonality



- ▷ In this work, we impose diversity constraints on features.

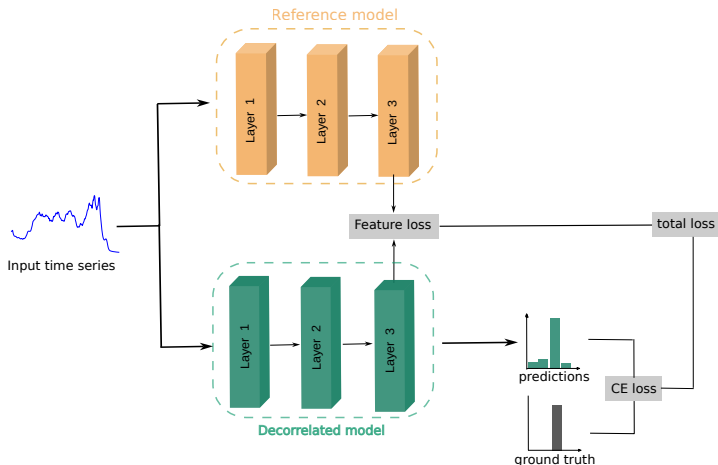
① Introduction and Literature Review

② Proposed Method

③ Experimental Evaluation

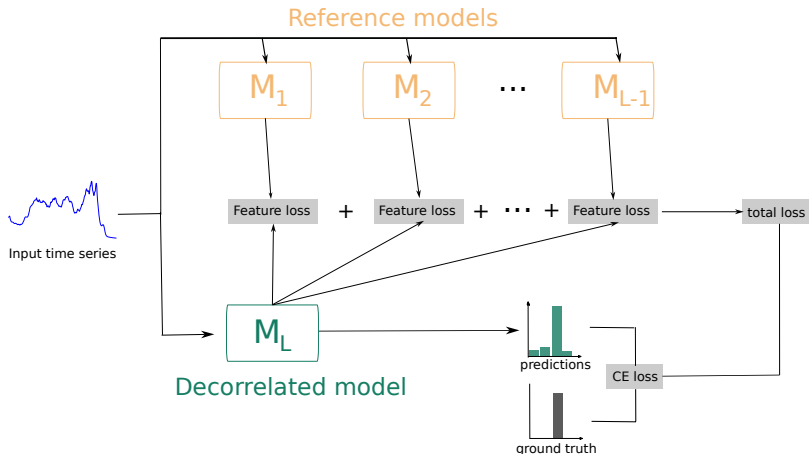
Proposed Approach

Overview of our Diversity Driven Decorrelated Learning Approach



Proposed Approach

Overview of our Diversity Driven Decorrelated Learning Approach



Diversity Loss Formulation

- ▷ We propose a diversity loss that encourages orthogonality between feature maps:

$$L_{orth} = \sum_{i \neq j} \left| \frac{\mathbf{F}_{deco,i} \cdot \mathbf{F}_{base,j}^T}{|\mathbf{F}_{deco,i}| |\mathbf{F}_{base,j}|} \right| \quad (1)$$

Diversity Loss Formulation

- ▷ We propose a diversity loss that encourages orthogonality between feature maps:

$$L_{orth} = \sum_{i \neq j} \left| \frac{\mathbf{F}_{deco,i} \cdot \mathbf{F}_{base,j}^T}{|\mathbf{F}_{deco,i}| |\mathbf{F}_{base,j}|} \right| \quad (1)$$

- ▷ Where:

$\mathbf{F}_{deco,i}$ is feature from decorrelated network

$\mathbf{F}_{base,j}$ is the feature from base network

The loss minimizes cosine similarity between two networks' features

Diversity Loss Formulation

- ▷ We propose a diversity loss that encourages orthogonality between feature maps:

$$L_{orth} = \sum_{i \neq j} \left| \frac{\mathbf{F}_{deco,i} \cdot \mathbf{F}_{base,j}^T}{\|\mathbf{F}_{deco,i}\| \|\mathbf{F}_{base,j}\|} \right| \quad (1)$$

- ▷ Where:

$\mathbf{F}_{deco,i}$ is feature from decorrelated network

$\mathbf{F}_{base,j}$ is the feature from base network

The loss minimizes cosine similarity between two networks' features

- ▷ Combined with standard classification loss:

$$L_{total} = \alpha L_{class} + (1 - \alpha) L_{orth} \quad (2)$$

where α controls the balance

① Introduction and Literature Review

② Proposed Method

③ Experimental Evaluation

Experimental Setup

- ▷ **Baseline model:** LITETime

Reference:

Dau et al., *The UCR Time Series Classification Archive*, 2018.

Experimental Setup

- ▷ **Baseline model:** LITETime
- ▷ **Datasets:** University California Riverside (UCR) archive 2018
 - ▷ We use 128 univariate time series datasets
 - ▷ All datasets already split into train and test sets

Reference:

Dau et al., *The UCR Time Series Classification Archive*, 2018.

Experimental Setup

- ▷ **Baseline model:** LITETime

- ▷ **Datasets:** University California Riverside (UCR) archive 2018
 - ▷ We use 128 univariate time series datasets
 - ▷ All datasets already split into train and test sets

- ▷ **Performance Comparison**
 - ▷ We compared standard LITETime ensembles with their decorrelated counterparts (Deco-LITETime), focusing on ensemble configurations ranging from 2 to 5 models.

Reference:

Dau et al., *The UCR Time Series Classification Archive*, 2018.

Experimental Setup

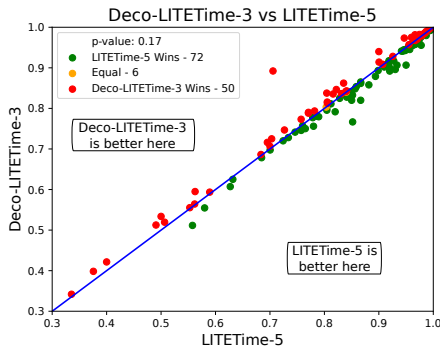
- ▷ **Baseline model:** LITETime
- ▷ **Datasets:** University California Riverside (UCR) archive 2018
 - ▷ We use 128 univariate time series datasets
 - ▷ All datasets already split into train and test sets
- ▷ **Performance Comparison**
 - ▷ We compared standard LITETime ensembles with their decorrelated counterparts (Deco-LITETime), focusing on ensemble configurations ranging from 2 to 5 models.
- ▷ **Metrics**
 - ▷ Ensemble accuracy of models are compared for each dataset from the UCR archive
 - ▷ Number of wins, ties and losses are computed for comparing models

Reference:

Dau et al., *The UCR Time Series Classification Archive*, 2018.

Performance Comparison

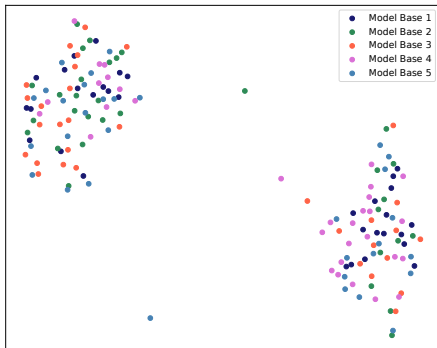
- ▷ 40% smaller *Deco-LITETime-3* achieves state-of-the-art performance as *LITETime-5*



Qualitative Diversity Analysis

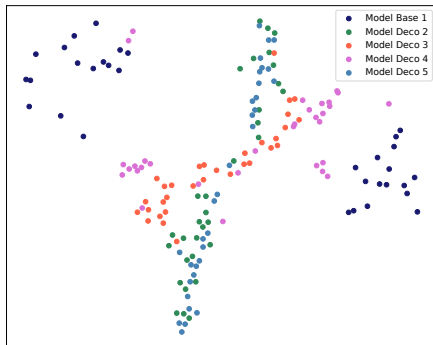
- Visualized convolutional filters from *base* and *decorrelated* models
- The t-SNE technique was employed to reduce the dimensionality of the filters for visualization

Convolutional Filter Visualization - Base Models



(a) Filters from base models

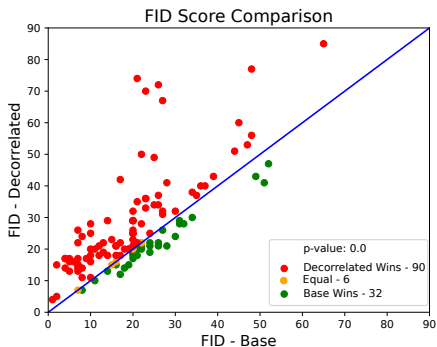
Convolutional Filter Visualization - Deco Models



(b) Filters from decorrelated models

Quantitative Diversity Analysis

- ▷ We use the Fréchet Inception Distance (FID) to quantitatively assess the diversity between feature outputs.
- ▷ The FID score compares:
 - x-axis:** Between two base models (Base_1 vs Base_2)
 - y-axis:** Between a base model and a decorrelated model (Base_1 vs Decorrelated_2)



FID comparison plot showing the effect of decorrelation on feature diversity.

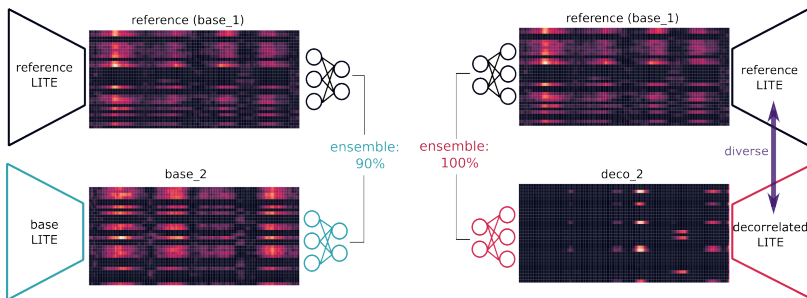
Takeaway

- ▷ We proposed Diversity Driven Ensemble Learning approach for TSC tasks
- ▷ Our approach, achieves state-of-the-art performance with 40% less parameters than the baseline LITETime model
- ▷ A current limitation is that the ensemble learning process is sequential
- ▷ As a perspective, in future work we plan to explore parallelizing the ensemble training

Code available here: <https://github.com/MSD-IRIMAS/decorrelated-learning>

Contact: <https://javidanabdullayev.github.io/>

Feature Diversity

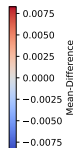


Comparison of ensemble model performances and feature maps on the BirdChicken dataset, from standard and decorrelated training.

Performance Comparison

The Multi-Comparison Matrix illustrates performance of each decorrelated and base ensemble variants in one-vs-one comparisons.

	Deco-LITime-4 Mean-Accuracy 0.8496	Deco-LITime-5 0.8494	LITime-5 0.8472	Deco-LITime-3 0.8469	LITime-4 0.8463	LITime-3 0.8455	LITime-2 0.8431	Deco-LITime-2 0.8427
Deco-LITime-4 0.8496	Mean-Difference $r > c / r = c / r < c$ Wilcoxon p-value	0.0002 44 / 20 / 64 0.2862	0.0025 59 / 8 / 61 0.5721	0.0028 81 / 13 / 34 $\leq 1e-04$	0.0033 68 / 9 / 51 0.0492	0.0041 73 / 13 / 42 0.0012	0.0066 85 / 10 / 33 $\leq 1e-04$	0.0069 94 / 11 / 23 $\leq 1e-04$
Deco-LITime-5 0.8494	-0.0002 64 / 20 / 44 0.2862	-	0.0023 62 / 9 / 57 0.2409	0.0026 83 / 13 / 32 $\leq 1e-04$	0.0031 71 / 8 / 49 0.0136	0.0039 75 / 12 / 41 0.0005	0.0064 85 / 10 / 33 $\leq 1e-04$	0.0067 94 / 10 / 24 $\leq 1e-04$
LITime-5 0.8472	-0.0025 61 / 8 / 59 0.5721	-0.0023 57 / 9 / 62 0.2409	-	0.0003 72 / 6 / 50 0.1718	0.0009 68 / 27 / 33 0.0016	0.0017 78 / 14 / 36 0.0002	0.0041 88 / 13 / 27 $\leq 1e-04$	0.0045 89 / 6 / 33 $\leq 1e-04$
Deco-LITime-3 0.8469	-0.0028 34 / 13 / 81 $\leq 1e-04$	-0.0026 32 / 13 / 83 $\leq 1e-04$	-0.0003 50 / 6 / 72 0.1718	-	0.0006 54 / 6 / 68 0.5387	0.0014 60 / 7 / 61 0.6419	0.0038 74 / 11 / 43 0.0012	0.0042 89 / 8 / 31 $\leq 1e-04$
Deco-LITime-2 0.8427	-0.0069 23 / 11 / 94 $\leq 1e-04$	-0.0067 24 / 10 / 94 $\leq 1e-04$	-0.0045 33 / 6 / 89 $\leq 1e-04$	-0.0042 31 / 8 / 89 $\leq 1e-04$	-0.0036 39 / 7 / 82 0.0002	-0.0028 40 / 6 / 82 0.0007	-0.0004 58 / 6 / 64 0.4763	If in bold, then p-value < 0.05



Performance Comparison

The Multi-Comparison Matrix illustrates performance of each decorrelated and base ensemble variants in one-vs-one comparisons.

Mean-Accuracy	Deco-LITime-4 0.8496	Deco-LITime-5 0.8494	LITime-5 0.8472	Deco-LITime-3 0.8469	LITime-4 0.8463	LITime-3 0.8455	LITime-2 0.8431	Deco-LITime-2 0.8427
Deco-LITime-4 0.8496	Mean-Difference $r > c / r = c / r < c$ Wilcoxon p-value	0.0002 44 / 20 / 64 0.2862	0.0025 59 / 8 / 61 0.5721	0.0028 81 / 13 / 34 $\leq 1e-04$	0.0033 68 / 9 / 51 0.0492	0.0041 73 / 13 / 42 0.0012	0.0066 85 / 10 / 33 $\leq 1e-04$	0.0069 94 / 11 / 23 $\leq 1e-04$
Deco-LITime-5 0.8494	-0.0002 64 / 20 / 44 0.2862	-	0.0023 62 / 9 / 57 0.2409	0.0026 83 / 13 / 32 $\leq 1e-04$	0.0031 71 / 8 / 49 0.0136	0.0039 75 / 12 / 41 0.0005	0.0064 85 / 10 / 33 $\leq 1e-04$	0.0067 94 / 10 / 24 $\leq 1e-04$
LITime-5 0.8472	-0.0025 61 / 8 / 59 0.5721	-0.0023 57 / 9 / 62 0.2409	-	0.0003 72 / 6 / 50 0.1718	0.0009 68 / 27 / 33 0.0016	0.0017 78 / 14 / 36 0.0002	0.0041 88 / 13 / 27 $\leq 1e-04$	0.0045 89 / 6 / 33 $\leq 1e-04$
Deco-LITime-3 0.8469	-0.0028 34 / 13 / 81 $\leq 1e-04$	-0.0026 32 / 13 / 83 $\leq 1e-04$	-0.0003 50 / 6 / 72 0.1718	-	0.0006 54 / 6 / 68 0.5387	0.0014 60 / 7 / 61 0.6419	0.0038 74 / 11 / 43 0.0012	0.0042 89 / 8 / 31 $\leq 1e-04$
Deco-LITime-2 0.8427	-0.0069 23 / 11 / 94 $\leq 1e-04$	-0.0067 24 / 10 / 94 $\leq 1e-04$	-0.0045 33 / 6 / 89 $\leq 1e-04$	-0.0042 31 / 8 / 89 $\leq 1e-04$	-0.0036 39 / 7 / 82 0.0002	-0.0028 40 / 6 / 82 0.0007	-0.0004 58 / 6 / 64 0.4763	If in bold, then p-value < 0.05

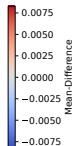
Mean-Difference



Performance Comparison

The Multi-Comparison Matrix illustrates performance of each decorrelated and base ensemble variants in one-vs-one comparisons.

	Deco-LITime-4 Mean-Accuracy 0.8496	Deco-LITime-5 0.8494	LITime-5 0.8472	Deco-LITime-3 0.8469	LITime-4 0.8463	LITime-3 0.8455	LITime-2 0.8431	Deco-LITime-2 0.8427
Deco-LITime-4 0.8496	Mean-Difference $r > c / r = c / r < c$ Wilcoxon p-value	0.0002 44 / 20 / 64 0.2862	0.0025 59 / 8 / 61 0.5721	0.0028 81 / 13 / 34 $\leq 1e-04$	0.0033 68 / 9 / 51 0.0492	0.0041 73 / 13 / 42 0.0012	0.0066 85 / 10 / 33 $\leq 1e-04$	0.0069 94 / 11 / 23 $\leq 1e-04$
Deco-LITime-5 0.8494	-0.0002 64 / 20 / 44 0.2862	-	0.0023 62 / 9 / 57 0.2409	0.0026 83 / 13 / 32 $\leq 1e-04$	0.0031 71 / 8 / 49 0.0136	0.0039 75 / 12 / 41 0.0005	0.0064 85 / 10 / 33 $\leq 1e-04$	0.0067 94 / 10 / 24 $\leq 1e-04$
LITime-5 0.8472	-0.0025 61 / 8 / 59 0.5721	-0.0023 57 / 9 / 62 0.2409	-	0.0003 72 / 6 / 50 0.1718	0.0009 68 / 27 / 33 0.0016	0.0017 78 / 14 / 36 0.0002	0.0041 88 / 13 / 27 $\leq 1e-04$	0.0045 89 / 6 / 33 $\leq 1e-04$
Deco-LITime-3 0.8469	-0.0028 34 / 13 / 81 $\leq 1e-04$	-0.0026 32 / 13 / 83 $\leq 1e-04$	-0.0003 50 / 6 / 72 0.1718	-	0.0006 54 / 6 / 68 0.5387	0.0014 60 / 7 / 61 0.6419	0.0038 74 / 11 / 43 0.0012	0.0042 89 / 8 / 31 $\leq 1e-04$
Deco-LITime-2 0.8427	-0.0069 23 / 11 / 94 $\leq 1e-04$	-0.0067 24 / 10 / 94 $\leq 1e-04$	-0.0045 33 / 6 / 89 $\leq 1e-04$	-0.0042 31 / 8 / 89 $\leq 1e-04$	-0.0036 39 / 7 / 82 0.0002	-0.0028 40 / 6 / 82 0.0007	-0.0004 58 / 6 / 64 0.4763	If in bold, then p-value < 0.05



Performance Comparison

The Multi-Comparison Matrix illustrates performance of each decorrelated and base ensemble variants in one-vs-one comparisons.

Mean-Accuracy	Deco-LITime-4 0.8496	Deco-LITime-5 0.8494	LITime-5 0.8472	Deco-LITime-3 0.8469	LITime-4 0.8463	LITime-3 0.8455	LITime-2 0.8431	Deco-LITime-2 0.8427
Deco-LITime-4 0.8496	Mean-Difference $r > c / r = c / r < c$ Wilcoxon p-value	0.0002 44 / 20 / 64 0.2862	0.0025 59 / 8 / 61 0.5721	0.0028 81 / 13 / 34 $\leq 1e-04$	0.0033 68 / 9 / 51 0.0492	0.0041 73 / 13 / 42 0.0012	0.0066 85 / 10 / 33 $\leq 1e-04$	0.0069 94 / 11 / 23 $\leq 1e-04$
Deco-LITime-5 0.8494	-0.0002 64 / 20 / 44 0.2862	-	0.0023 62 / 9 / 57 0.2409	0.0026 83 / 13 / 32 $\leq 1e-04$	0.0031 71 / 8 / 49 0.0136	0.0039 75 / 12 / 41 0.0005	0.0064 85 / 10 / 33 $\leq 1e-04$	0.0067 94 / 10 / 24 $\leq 1e-04$
LITime-5 0.8472	-0.0025 61 / 8 / 59 0.5721	-0.0023 57 / 9 / 62 0.2409	-	0.0003 72 / 6 / 50 0.1718	0.0009 68 / 27 / 33 0.0016	0.0017 78 / 14 / 36 0.0002	0.0041 88 / 13 / 27 $\leq 1e-04$	0.0045 89 / 6 / 33 $\leq 1e-04$
Deco-LITime-3 0.8469	-0.0028 34 / 13 / 81 $\leq 1e-04$	-0.0026 32 / 13 / 83 $\leq 1e-04$	-0.0003 50 / 6 / 72 0.1718	-	0.0006 54 / 6 / 68 0.5387	0.0014 60 / 7 / 61 0.6419	0.0038 74 / 11 / 43 0.0012	0.0042 89 / 8 / 31 $\leq 1e-04$
Deco-LITime-2 0.8427	-0.0069 23 / 11 / 94 $\leq 1e-04$	-0.0067 24 / 10 / 94 $\leq 1e-04$	-0.0045 33 / 6 / 89 $\leq 1e-04$	-0.0042 31 / 8 / 89 $\leq 1e-04$	-0.0036 39 / 7 / 82 0.0002	-0.0028 40 / 6 / 82 0.0007	-0.0004 58 / 6 / 64 0.4763	If in bold, then p-value < 0.05

Mean-Difference



Performance Comparison

The Multi-Comparison Matrix illustrates performance of each decorrelated and base ensemble variants in one-vs-one comparisons.

	Deco-LITime-4 Mean-Accuracy 0.8496	Deco-LITime-5 Mean-Accuracy 0.8494	LITime-5 Mean-Accuracy 0.8472	Deco-LITime-3 Mean-Accuracy 0.8469	LITime-4 Mean-Accuracy 0.8463	LITime-3 Mean-Accuracy 0.8455	LITime-2 Mean-Accuracy 0.8431	Deco-LITime-2 Mean-Accuracy 0.8427
Deco-LITime-4 0.8496	Mean-Difference $r > c / r = c / r < c$ Wilcoxon p-value	0.0002 44 / 20 / 64 0.2862	0.0025 59 / 8 / 61 0.5721	0.0028 81 / 13 / 34 $\leq 1e-04$	0.0033 68 / 9 / 51 0.0492	0.0041 73 / 13 / 42 0.0012	0.0066 85 / 10 / 33 $\leq 1e-04$	0.0069 94 / 11 / 23 $\leq 1e-04$
Deco-LITime-5 0.8494	-0.0002 64 / 20 / 44 0.2862	-	0.0023 62 / 9 / 57 0.2409	0.0026 83 / 13 / 32 $\leq 1e-04$	0.0031 71 / 8 / 49 0.0136	0.0039 75 / 12 / 41 0.0005	0.0064 85 / 10 / 33 $\leq 1e-04$	0.0067 94 / 10 / 24 $\leq 1e-04$
LITime-5 0.8472	-0.0025 61 / 8 / 59 0.5721	-0.0023 57 / 9 / 62 0.2409	-	0.0003 72 / 6 / 50 0.1718	0.0009 68 / 27 / 33 0.0016	0.0017 78 / 14 / 36 0.0002	0.0041 88 / 13 / 27 $\leq 1e-04$	0.0045 89 / 6 / 33 $\leq 1e-04$
Deco-LITime-3 0.8469	-0.0028 34 / 13 / 81 $\leq 1e-04$	-0.0026 32 / 13 / 83 $\leq 1e-04$	-0.0003 50 / 6 / 72 0.1718	-	0.0006 54 / 6 / 68 0.5387	0.0014 60 / 7 / 61 0.6419	0.0038 74 / 11 / 43 0.0012	0.0042 89 / 8 / 31 $\leq 1e-04$
Deco-LITime-2 0.8427	-0.0069 23 / 11 / 94 $\leq 1e-04$	-0.0067 24 / 10 / 94 $\leq 1e-04$	-0.0045 33 / 6 / 89 $\leq 1e-04$	-0.0042 31 / 8 / 89 $\leq 1e-04$	-0.0036 39 / 7 / 82 0.0002	-0.0028 40 / 6 / 82 0.0007	-0.0004 58 / 6 / 64 0.4763	If in bold, then p-value < 0.05

Mean-Difference



Performance Comparison

The Multi-Comparison Matrix illustrates performance of each decorrelated and base ensemble variants in one-vs-one comparisons.

Mean-Accuracy	Deco-LITime-4 0.8496	Deco-LITime-5 0.8494	LITime-5 0.8472	Deco-LITime-3 0.8469	LITime-4 0.8463	LITime-3 0.8455	LITime-2 0.8431	Deco-LITime-2 0.8427
Deco-LITime-4 0.8496	Mean-Difference $r > c / r = c / r < c$ Wilcoxon p-value	0.0002 44 / 20 / 64 0.2862	0.0025 59 / 8 / 61 0.5721	0.0028 81 / 13 / 34 $\leq 1e-04$	0.0033 68 / 9 / 51 0.0492	0.0041 73 / 13 / 42 0.0012	0.0066 85 / 10 / 33 $\leq 1e-04$	0.0069 94 / 11 / 23 $\leq 1e-04$
Deco-LITime-5 0.8494	-0.0002 64 / 20 / 44 0.2862	-	0.0023 62 / 9 / 57 0.2409	0.0026 83 / 13 / 32 $\leq 1e-04$	0.0031 71 / 8 / 49 0.0136	0.0039 75 / 12 / 41 0.0005	0.0064 85 / 10 / 33 $\leq 1e-04$	0.0067 94 / 10 / 24 $\leq 1e-04$
LITime-5 0.8472	-0.0025 61 / 8 / 59 0.5721	-0.0023 57 / 9 / 62 0.2409	-	0.0003 72 / 6 / 50 0.1718	0.0009 68 / 27 / 33 0.0016	0.0017 78 / 14 / 36 0.0002	0.0041 88 / 13 / 27 $\leq 1e-04$	0.0045 89 / 6 / 33 $\leq 1e-04$
Deco-LITime-3 0.8469	-0.0028 34 / 13 / 81 $\leq 1e-04$	-0.0026 32 / 13 / 83 $\leq 1e-04$	-0.0003 50 / 6 / 72 0.1718	-	0.0006 54 / 6 / 68 0.5387	0.0014 60 / 7 / 61 0.6419	0.0038 74 / 11 / 43 0.0012	0.0042 89 / 8 / 31 $\leq 1e-04$
Deco-LITime-2 0.8427	-0.0069 23 / 11 / 94 $\leq 1e-04$	-0.0067 24 / 10 / 94 $\leq 1e-04$	-0.0045 33 / 6 / 89 $\leq 1e-04$	-0.0042 31 / 8 / 89 $\leq 1e-04$	-0.0036 39 / 7 / 82 0.0002	-0.0028 40 / 6 / 82 0.0007	-0.0004 58 / 6 / 64 0.4763	If in bold, then p-value < 0.05

Mean-Difference

