RESEARCH TRACK

POND: Multi-Source Time Series Domain Adaptation with Information-Aware Prompt Tuning

The 30th SIGKDD Conference on Knowledge Discovery and Data mining

Presenter: Wei Cheng

Junxiang Wang (NEC Labs America);
Guangji Bai (Emory University);
Wei Cheng (NEC Labs America);
ZhengZhang Chen (NEC Labs America);
Liang Zhao (Emory University);
Haifeng Chen (NEC Labs America);

KDD2024 BARCELONA, SPAIN

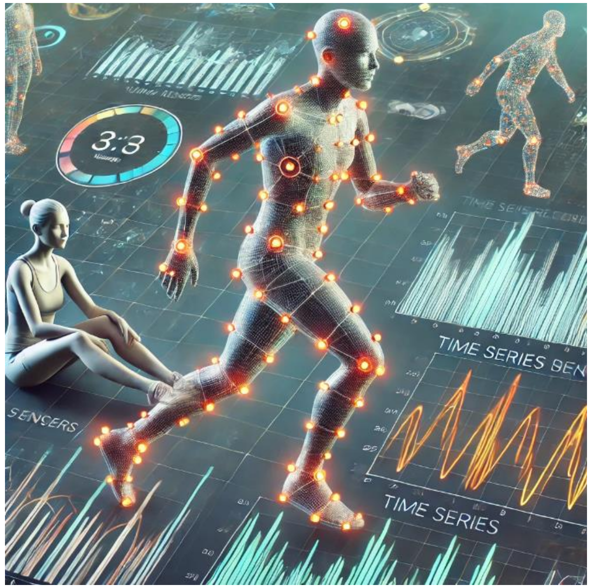
Outline

- Background
- Proposed Method - POND
- Experiment
- Conclusion

KDD2024
BARCELONA, SPAIN

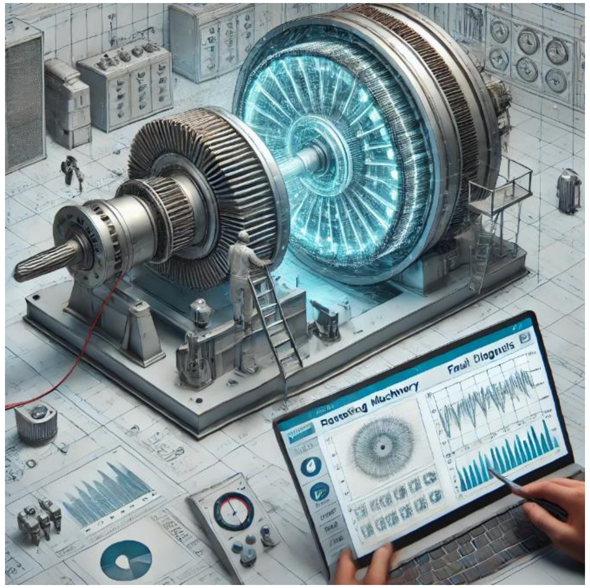
Time Series Domain Adaptation and Its Applications

Due to the prevalence of time series sensor data, time series domain adaptation has found applications in various real-world scenarios.



Human Activity Recognition

Sleep Stage Classification

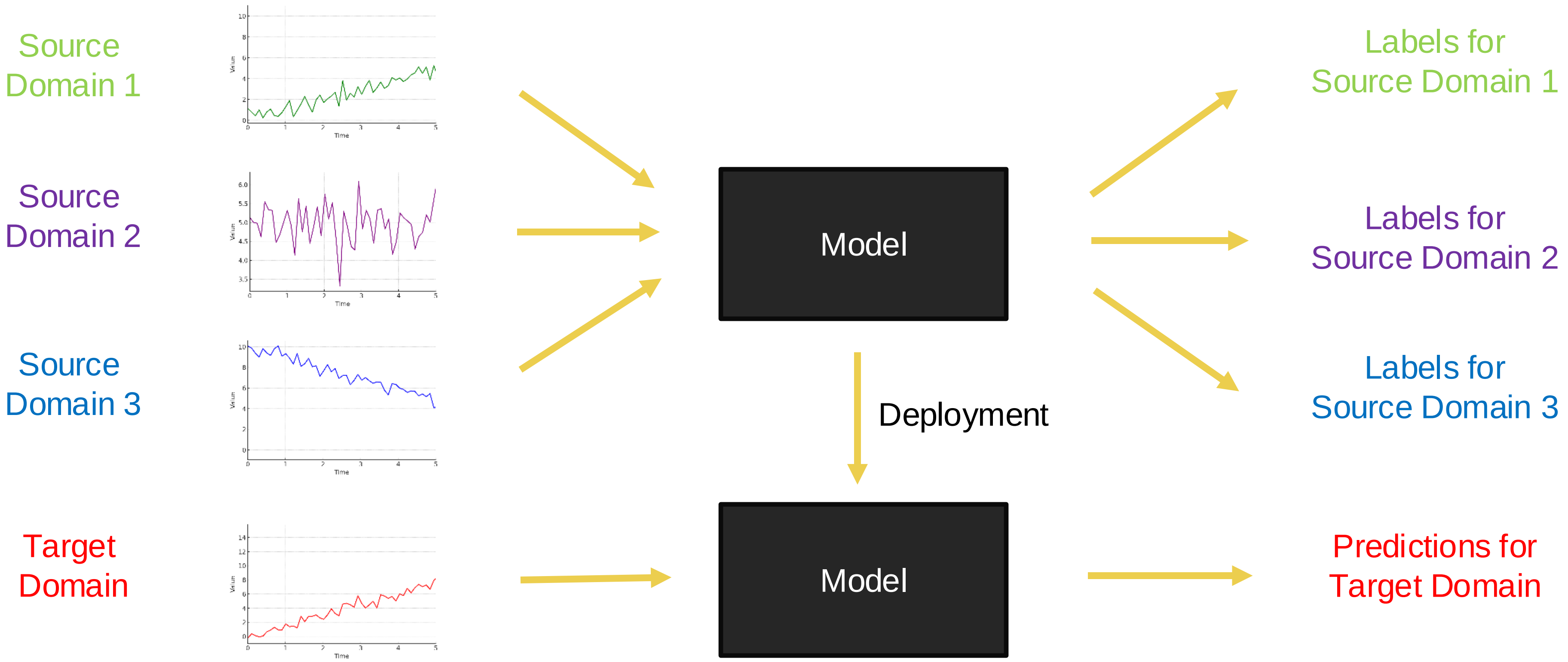


Machine Fault Diagnosis

KDD2024
BARCELONA, SPAIN

Multi-source Time Series Domain Adaptation Problem

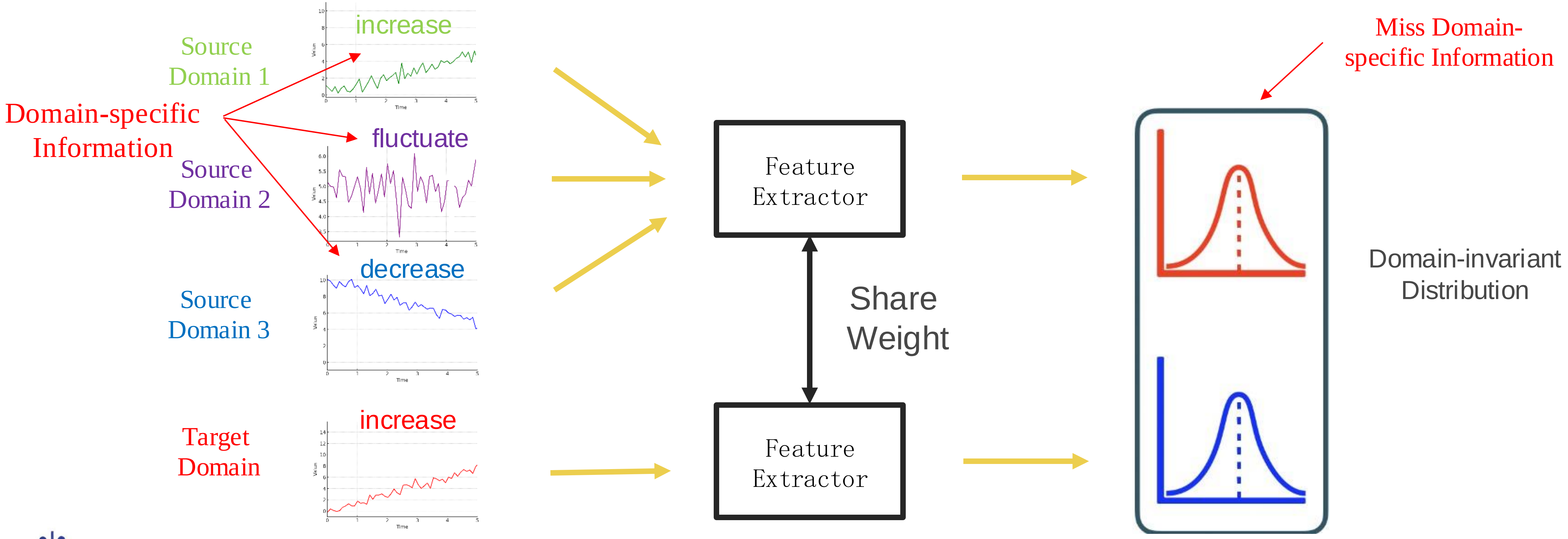
The problem leverages labeled data from multiple domains (i.e., source domains) to infer labels for unlabeled data in other domains (i.e., target domains).



KDD2024
BARCELONA, SPAIN

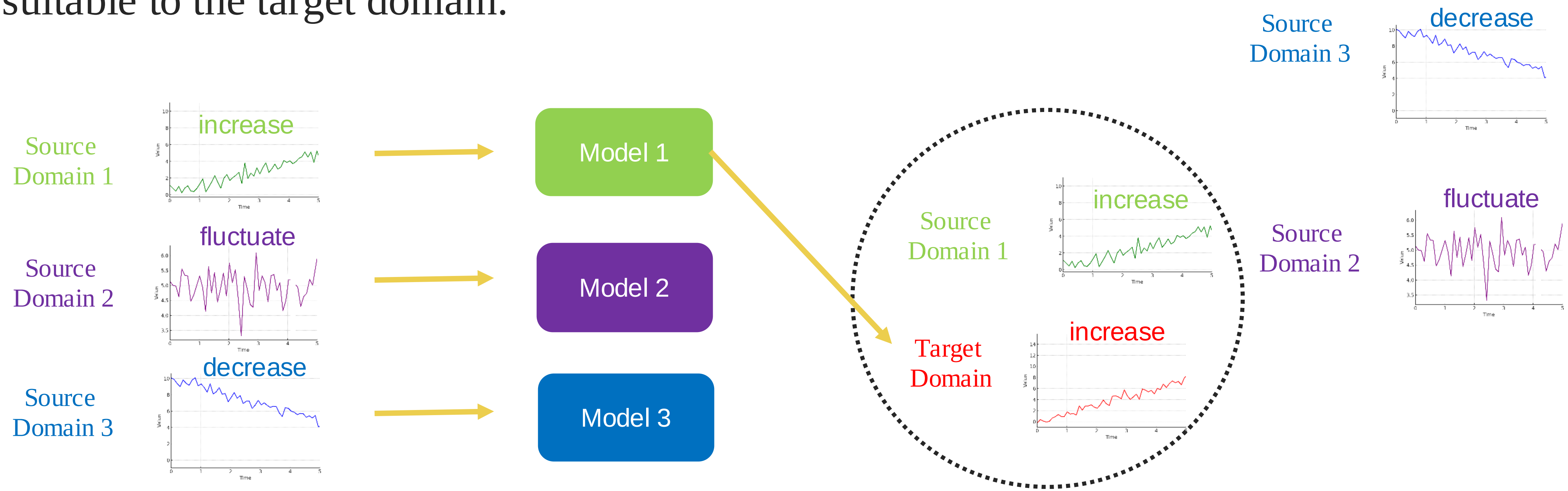
Limitations of Existing Works

Ideas of existing papers propose various methods are mainly to learn domain-invariant time series representations, but they ignore domain-specific information.



Learning Domain-Specific Information

Domain-specific information provides knowledge to determine which source domain is the most suitable to the target domain.



The target domain is the most similar to source domain 1,which both show the increase trend. Therefore, model 1 is used for prediction in the target domain.

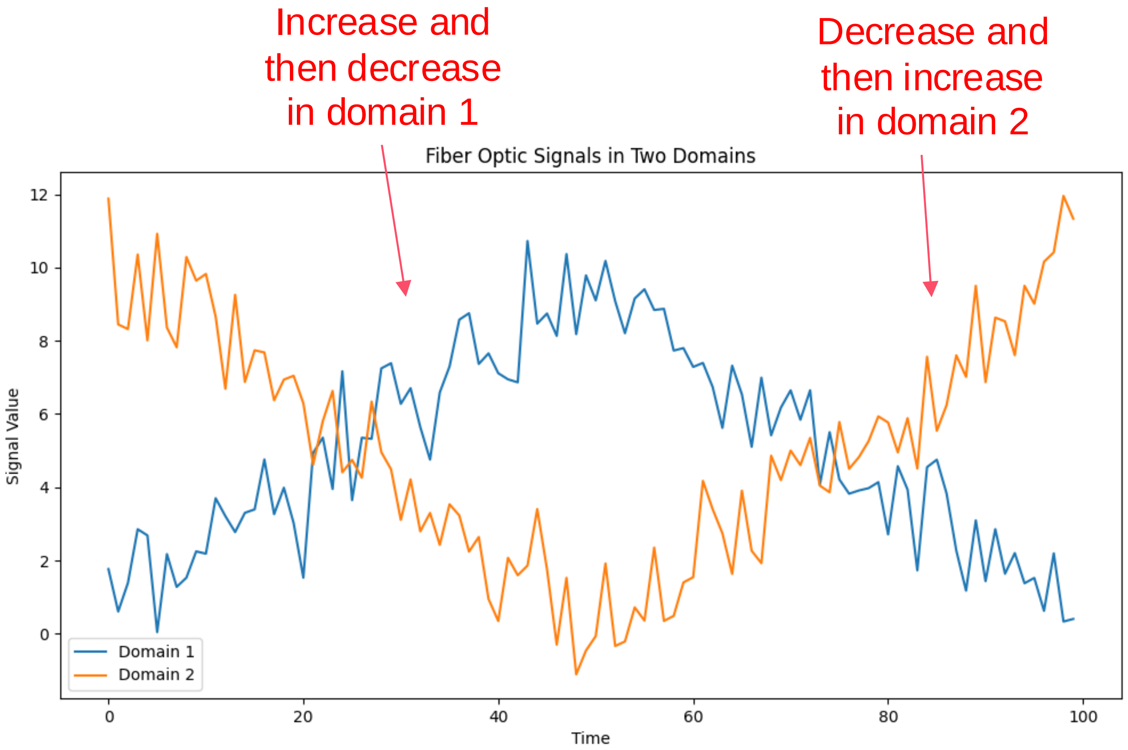
KDD2024 BARCELONA, SPAIN

What are challenges of learning domain-specific information?

Challenge 1: Domain-Specific Information Changes Over time

- Domain-specific information can be dynamically changing, which is extremely difficult to capture.

- For example, the right figure shows the distributions of fiber-optic signals shift over time in two domains.



KDD2024
BARCELONA, SPAIN

Challenge 2: Difficulty to Evaluate Learned Domain-Specific Information

- It is unclear whether the learned domain-specific information accurately reflects the true.

- Domain-specific information is often associated with unique but inexplicable underlying patterns.

- Unlike images and languages with human-recognizable features, such time series patterns are difficult for humans to understand.

KDD2024
BARCELONA, SPAIN

Outline

Background

Proposed Method - POND

Experiment

Conclusion

KDD2024
BARCELONA, SPAIN

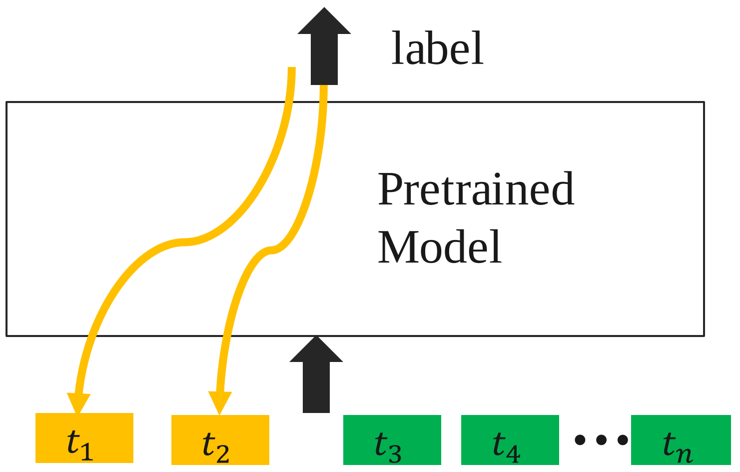
Key Idea: Prompts to Learn Domain-Specific Information

- Prompts are extra time segments to input time-series and use the small task-specific data (with labels) to learn prompt values.

- In the prompt tunning phase, the pretrained time-series model is frozen, and

➢ Prompt values are updated by the gradient based on the labels in training data.

➢ Because of only a tiny number of prompt tokens, we only need a very small number of labeled samples to learn prompt values.



KDD2024 BARCELONA, SPAIN

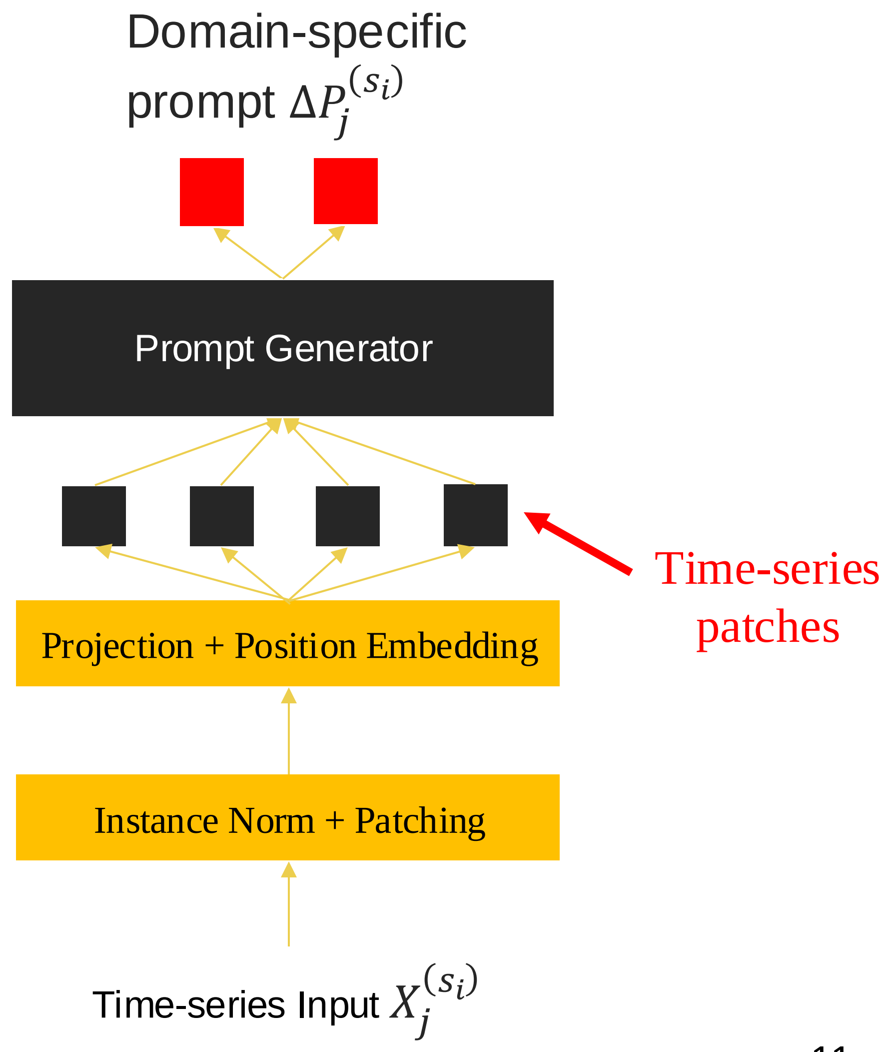
Prompt Generator to Address Challenge 1

- We decompose a prompt $P^{(s_i)}$ from the source domain S_i into the common prompt P and the domain-specific prompt $\Delta P^{(s_i)}$: $P^{(s_i)} = P + \Delta P^{(s_i)}$.

- While P is fixed, $\Delta P^{(s_i)}$ needs to capture dynamic domain-specific information. The time series input usually contains such information (e.g., time series distributions and trends).

- To achieve this, we propose a conditional prompt generator $g^{(S_i)}$ to generate instance-level prompts based on time series instances: $\Delta P_j^{(s_i)} = g^{(S_i)}(X_j^{(s_i)}; \zeta)$.

- The domain-level $\Delta P^{(s_i)}$ is an aggregation of instance-level prompts $\Delta P_j^{(s_i)}$: $\Delta P^{(s_i)} = \frac{1}{|S_i|} \sum_{j=1}^{|S_i|} \Delta P_j^{(s_i)}$.



KDD2024
BARCELONA, SPAIN

Two Important Criteria to Address Challenge 2

To evaluate learned prompts, we proposes two important criteria:

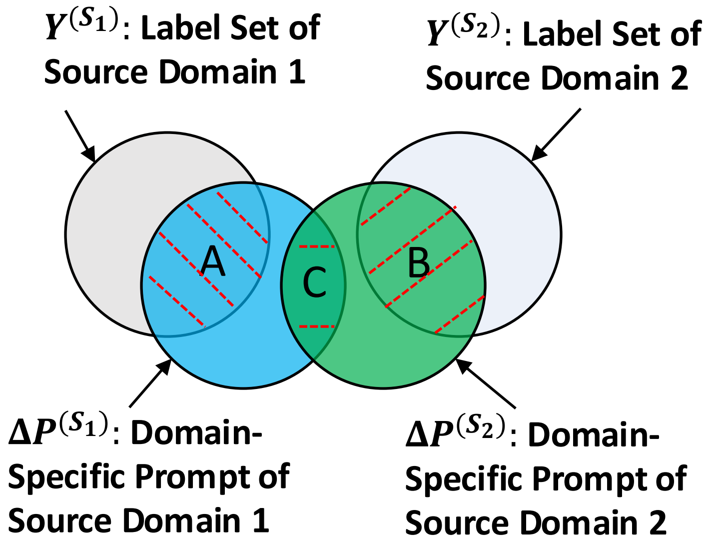
- High Fidelity: the motivation is that the generated prompt $\Delta P_j^{(s_i)}$ preserves the domain-specific information of the source domain S_i. This can be achieved by the following:

$$\max \sum MI(\Delta P_j^{(s_i)}, Y_j^{(s_i)})$$

- High Distinction: the generated domain-specific prompt $P^{(s_i)}$ distinguishes the unique information of the source domain S_i from other source domains. This can be achieved by the following:

$$min \sum MI(\Delta P^{(s_{i_1})}, \Delta P^{(s_{i_2})})$$

Where $MI(\cdot)$ is the mutual information operator, and $Y_j^{(s_i)}$ is the instance-level label.



Fidelity: A+B

Distinction: C

KDD2024
BARCELONA, SPAIN

Learning Objective and Optimization

$$min_{P,g^{(S_i)}}\, G\left(P, g^{(S_i)}\right) = \underbrace{\sum R\left(f\left(P + \Delta P_j^{(s_i)}, X_j^{(s_i)}\right), Y_j^{(s_i)}\right)}_{\text{Training Loss}} + \lambda_1 \underbrace{\sum MI\left(\Delta P^{(s_{i_1})}, \Delta P^{(s_{i_2})}\right)}_{\text{Discrimination Loss}} - \lambda_2 \underbrace{\sum MI\left(\Delta P_j^{(s_i)}, Y_j^{(s_i)}\right)}_{\text{Fidelity Loss}}$$

where $R(\cdot)$ is a risk function, and $f(\cdot)$ is our proposed POND model.

To optimize the objective, we propose a simple, yet effective learning algorithm based on the classic Reptile meta-learning framework, which randomly picks a source domain each time and conducts standard steps of gradient descent, as shown in Algorithm 1.

Algorithm 1 Reptile-based meta-learning for Prompt Tuning

Require: $(X_j^{(S_i)}, Y_j^{(S_i)})$, the global learning rate $\delta \in (0, 1]$, the local learning rate $\eta > 0$, the number of global steps N.

Ensure: the common prompt P, the prompt generator $g^{(S_i)}$.

for $i = 1$ to N **do**

  Randomly pick a source time series domain S_τ.

  $g^{(S_\tau)} \leftarrow g^{(S_\tau)} - \eta \nabla_{g^{(S_\tau)}} G$.

  $Q \leftarrow P - \eta \nabla_P \ell_T$.

  $P \leftarrow P + \delta(Q - P)$.

end for

Target Domain Transfer

- After the common prompt P and the prompt generators $g^{(S_i)}$ are optimized, the prompt generator $g^{(T)}$ of the target domain T is optimized using few labeled data in the target domain as follows:

$$\min_{g^{(T)}} \sum R(f\left(P + \Delta P_j^{(T)}, X_j^{(T)}\right), Y_j^{(T)})$$

where $\Delta P_j^{(T)} = g^{(T)}(X_j^{(T)}; \zeta)$ is the instance-level domain-specific prompt of the time series input $X_j^{(T)}$.

- The domain-level $\Delta P^{(T)}$ is an aggregation of instance-level prompts $\Delta P_j^{(T)}$: $\Delta P^{(T)} = \frac{1}{|T|} \sum_{j=1}^{|T|} \Delta P_j^{(T)}$.

- $\Delta P^{(T)}$ is utilized as a heuristic to find the most similar source domain by the simple nearest neighbor rule (i.e. prompt adaptation):

$$S_i = \mathrm{argmax}_{S_i}\, sim(\Delta P^{(S_i)}, \Delta\, P^{(T)})$$

where $sim(\cdot)$ is the similarity function (e.g. cosine similarity), and the target domain will use the prompt generator of the most similar source domain for prediction.

KDD2024
BARCELONA, SPAIN

Discussion

Model Architecture:

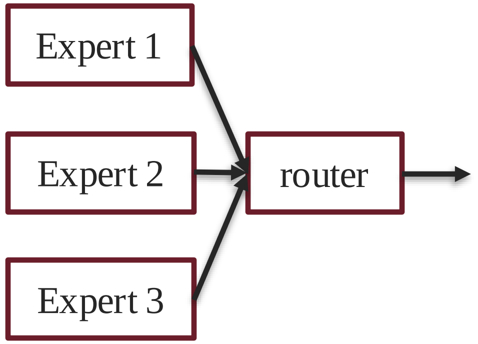
- For the model architecture of our proposed POND model, we employ the popular Mixture of Expert (MoE) technique to enhance performance: each expert makes an independent prediction, and the router is responsible for learning probability distributions over all predictions. The overall output of our POND model is a linear combination of all predictions.

- For the architecture of a single expert, the time series input is fed into “a patching layer” (i.e., splitting a timeseries input into subseries-level patches), a projection layer, a position embedding layer, a transformer layer, and a linear head sequentially.

Theoretical Analysis: we prove that our proposed POND model shares the universal approximation with the traditional prompt tuning, and we also prove that our proposed POND model overcomes the limitation of the traditional prompt tuning.

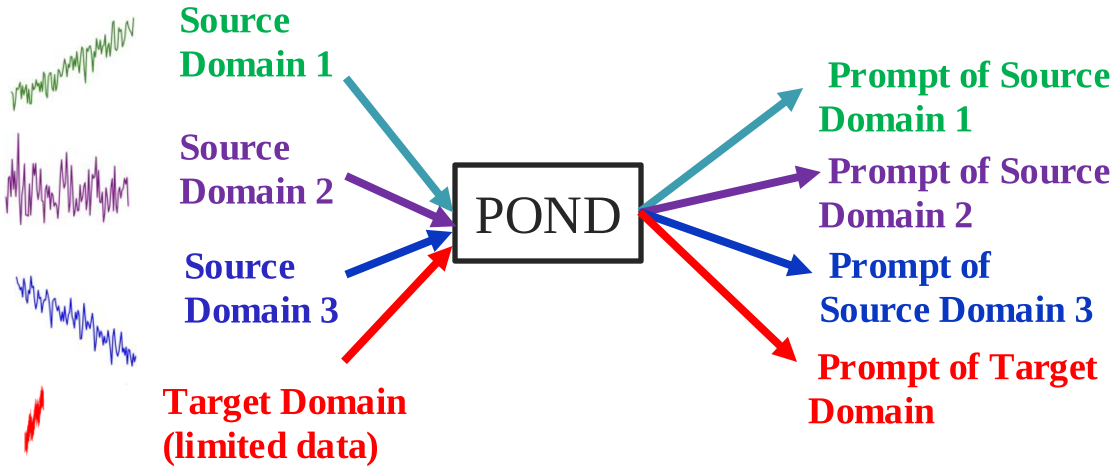
POND Model Implementation

Step 1. Model Pretraining

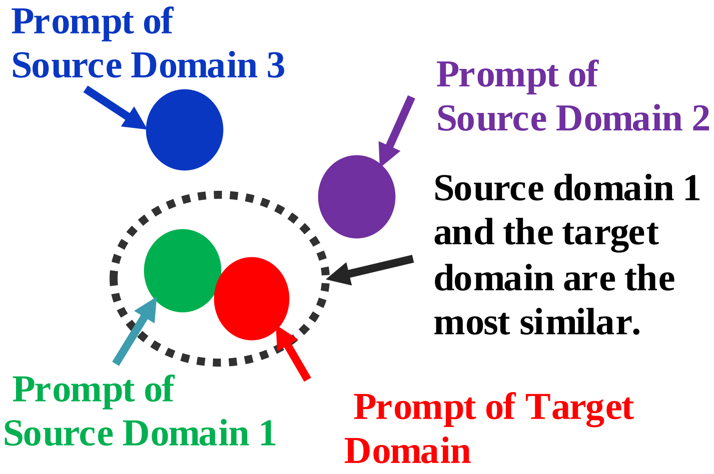


POND Model

Step 2. Prompt Tuning



Step 3. Prompt Adaptation



Outline

Background

Proposed Method - POND

Experiment

Conclusion

KDD2024
BARCELONA, SPAIN

Experiment Setup

Training Protocols:

- The ratio of training, validation, and test is 60%, 20% and 20%.
- Ten samples were used in the target domain for domain transfer.
- All scenario-target scenarios were selected randomly.

Evaluation metric: Macro-F1 score and Accuracy.

Baselines:

- Single-source methods: Raincoat [1], Deep Coral [2], MMDA [3], DIRT-T [4], DSAN [5].
- Multi-source method: CoDATs [6].

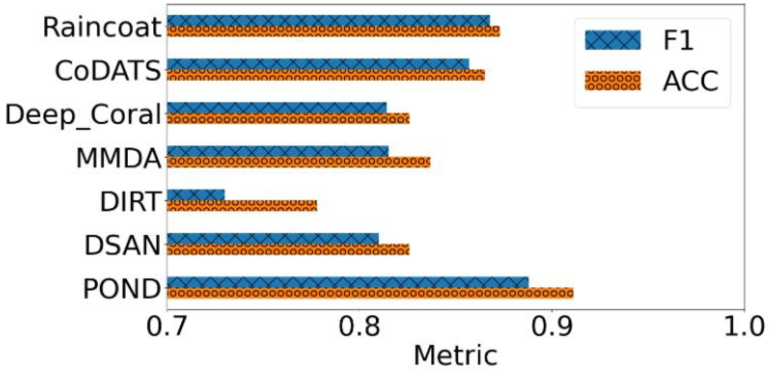
Dataset	# Domain	# Channel	# Class	Seq Len	# Train	# Test
HAR	30	9	6	128	2300	990
WISDM	36	3	6	128	1350	720
HHAR	9	3	6	128	12716	5218
SSC	20	1	5	3000	14280	6130

KDD2024
BARCELONA, SPAIN

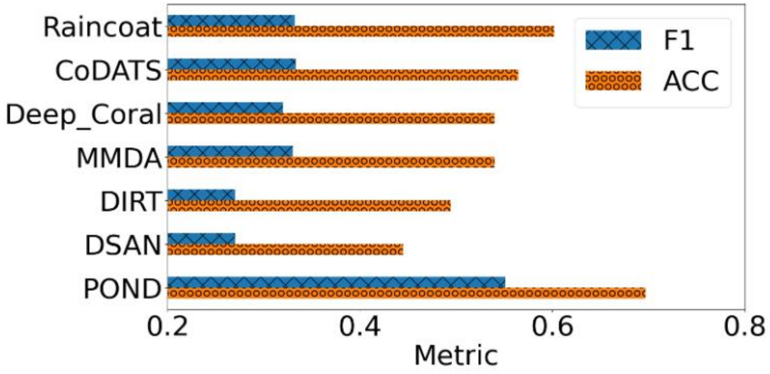
[1]. He, Huan, et al. "Domain adaptation for time series under feature and label shifts." International Conference on Machine Learning. PMLR, 2023.
[2]. Sun, Baochen, and Kate Saenko. "Deep coral: Correlation alignment for deep domain adaptation." Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part III 14. Springer International Publishing, 2016.
[3]. Rahman, et al. "On minimum discrepancy estimation for deep domain adaptation." Domain Adaptation for Visual Understanding (2020): 81-94.
[4]. Shu, Rui, et al. "A DIRT-T Approach to Unsupervised Domain Adaptation." International Conference on Learning Representations. 2018.
[5]. Zhu, Yongchun, et al. "Deep subdomain adaptation network for image classification." IEEE transactions on neural networks and learning systems 32.4 (2020): 1713-1722.
[6]. Wilson, Garrett, Janardhan Rao Doppa, and Diane J. Cook. "Multi-source deep domain adaptation with weak supervision for time-series sensor data." Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining. 2020.

Experiments: Overall Performance

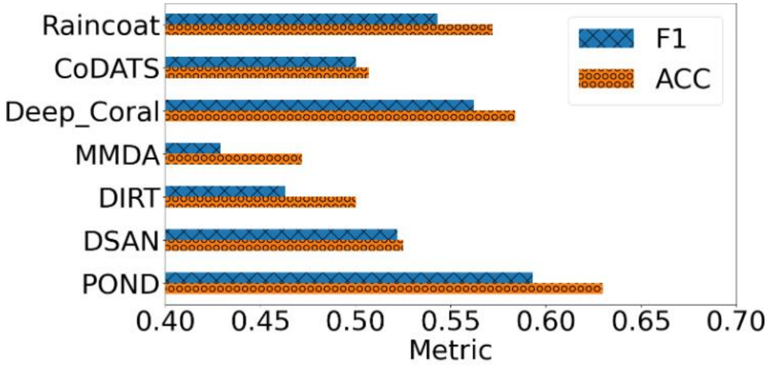
- We conducted comprehensive evaluation of 50 scenarios on four datasets.
- Our proposed POND method consistently outperforms others across all four datasets, with the largest gap on the WISDM dataset.
- Among the comparison methods, Raincoat emerges as the best overall.



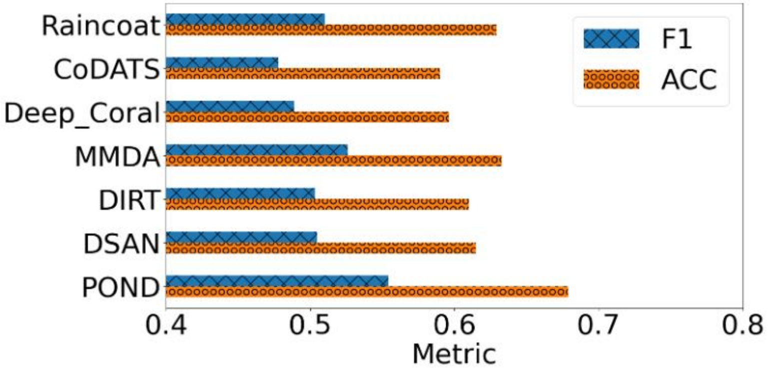
(a). HAR.



(b). WISDM.



(c). HHAR.



(d). SSC.

KDD2024
BARCELONA, SPAIN

Experiments: Scenario Performance

- In terms of scenario performance, all results include means and standard deviations from ten implementations.
- Our proposed POND model excels in others on the F1-score up to **66%**, and is nearly close to target only performance in some cases.
- Our proposed POND model is more stable than all comprise methods.

Scenario	Raincoat	CoDATs	Deep_Coral	MMDA	DIRT	DSAN	POND	Target Only
HAR 1-15 → 16	0.823 ± 0.094	0.767 ± 0.093	0.773 ± 0.082	0.679 ± 0.084	0.612 ± 0.135	0.738 ± 0.095	**0.849 ± 0.021**	0.856 ± 0.027
HAR 1-15 → 20	0.872 ± 0.142	0.932 ± 0.025	0.923 ± 0.023	0.921 ± 0.034	0.848 ± 0.101	0.929 ± 0.033	**0.968 ± 0.021**	0.983 ± 0.018
HAR 1-15 → 21	0.867 ± 0.141	0.903 ± 0.070	0.882 ± 0.028	**0.974 ± 0.039**	0.921 ± 0.090	0.909 ± 0.110	0.972 ± 0.021	1.000 ± 0.000
HAR 1-15 → 28	0.766 ± 0.107	0.775 ± 0.166	**0.852 ± 0.044**	0.778 ± 0.085	0.671 ± 0.175	0.783 ± 0.046	0.829 ± 0.018	0.853 ± 0.019
HAR 16-20 → 1	0.792 ± 0.072	0.744 ± 0.053	0.667 ± 0.077	0.654 ± 0.074	0.546 ± 0.060	0.698 ± 0.037	**0.883 ± 0.017**	0.986 ± 0.010
HAR 16-20 → 2	0.825 ± 0.048	0.821 ± 0.151	0.796 ± 0.055	0.651 ± 0.045	0.509 ± 0.050	0.652 ± 0.057	**0.936 ± 0.017**	0.943 ± 0.024
HAR 16-20 → 3	0.814 ± 0.028	0.746 ± 0.078	0.741 ± 0.058	0.657 ± 0.033	0.605 ± 0.056	0.565 ± 0.043	**0.878 ± 0.018**	0.978 ±0.013
HAR 16-20 → 4	0.679 ± 0.084	0.605 ± 0.082	0.479 ± 0.110	0.513 ± 0.058	0.336 ± 0.110	0.436 ± 0.032	**0.754 ± 0.033**	0.921 ± 0.018
WISDM 0-17 → 18	0.379 ± 0.061	0.384 ± 0.049	0.346 ± 0.023	0.297 ± 0.016	0.300 ± 0.041	0.287 ± 0.045	**0.606 ± 0.020**	0.705 ± 0.046
WISDM 0-17 → 20	0.354 ± 0.040	0.368 ± 0.039	0.376 ± 0.031	0.452 ± 0.098	0.347 ± 0.071	0.269 ± 0.064	**0.570 ± 0.023**	0.704 ± 0.051
WISDM 0-17 → 21	0.355 ± 0.057	0.310 ± 0.088	0.259 ± 0.018	0.250 ± 0.000	0.276 ± 0.055	0.245 ± 0.046	**0.450 ± 0.026**	0.636 ± 0.095
WISDM 0-17 → 23	0.306 ± 0.015	0.327 ± 0.075	0.318 ± 0.031	0.327 ± 0.023	0.271 ± 0.016	0.277 ± 0.044	**0.482 ± 0.017**	0.538 ± 0.034
WISDM 0-17 → 25	0.365 ± 0.030	0.540 ± 0.125	0.435 ± 0.043	0.436 ± 0.094	0.314 ± 0.107	0.353 ± 0.120	**0.559 ± 0.050**	0.672 ± 0.039
WISDM 0-17 → 28	0.399 ± 0.028	0.431 ± 0.033	0.418 ± 0.032	0.454 ± 0.064	0.304 ± 0.044	0.339 ± 0.030	**0.656 ± 0.046**	0.689 ± 0.048
WISDM 0-17 → 30	0.314 ± 0.020	0.305 ± 0.028	0.298 ± 0.023	0.359 ± 0.072	0.266 ± 0.035	0.246 ± 0.076	**0.670 ± 0.039**	0.791 ± 0.028
WISDM 18-23 → 5	0.648 ± 0.001	0.558 ± 0.129	0.534 ± 0.102	0.510 ± 0.020	0.549 ± 0.097	0.484 ± 0.055	**0.652 ± 0.035**	0.734 ± 0.095
WISDM 18-23 → 6	0.544 ± 0.074	0.565 ± 0.143	0.437 ± 0.078	0.543 ± 0.160	0.405 ± 0.089	0.454 ± 0.112	**0.628 ± 0.033**	0.872 ± 0.049
WISDM 18-23 → 7	0.588 ± 0.070	0.404 ± 0.117	0.530 ± 0.094	0.477 ± 0.060	0.518 ± 0.120	0.476 ± 0.127	**0.672 ± 0.029**	0.888 ± 0.035
HHAR 0-6 → 7	0.765 ± 0.142	0.652 ± 0.108	0.815 ± 0.105	0.641 ± 0.050	0.649 ± 0.005	0.730 ± 0.164	**0.834 ± 0.014**	0.861 ± 0.016
HHAR 5-8 → 2	0.321 ± 0.023	0.347 ± 0.082	0.309 ± 0.032	0.216 ± 0.032	0.276 ± 0.021	0.314 ± 0.095	**0.352 ± 0.014**	0.881 ± 0.018
SSC 0-9 → 16	0.578 ± 0.028	0.510 ± 0.044	0.537 ± 0.024	0.559 ± 0.027	0.523 ± 0.019	0.515 ± 0.044	**0.568 ± 0.012**	0.601 ± 0.018
SSC 0-9 → 17	0.511 ± 0.024	0.413 ± 0.118	0.452 ± 0.077	0.504 ± 0.060	0.530 ± 0.053	0.463 ± 0.081	**0.559 ± 0.006**	0.602 ± 0.014
SSC 0-9 → 18	**0.605 ± 0.016**	0.548 ± 0.037	0.544 ± 0.046	0.597 ± 0.032	0.574 ± 0.021	0.569 ± 0.046	0.604 ± 0.014	0.602 ± 0.013
SSC 0-9 → 19	0.562 ± 0.024	0.540 ± 0.052	0.531 ± 0.055	**0.570 ± 0.044**	0.565 ± 0.028	0.568 ± 0.080	**0.570 ± 0.010**	0.613 ± 0.019
SSC 10-12 → 8	0.294 ± 0.028	0.380 ± 0.066	0.379 ± 0.076	0.398 ± 0.060	0.322 ± 0.048	0.411 ± 0.046	**0.470 ± 0.010**	0.531 ± 0.019

KDD2024
BARCELONA, SPAIN

Experiments: Scenario Performance

Nearly close to target-only performance

- In terms of scenario performance, all results include means and standard deviations from ten implementations.
- Our proposed POND model excels in others on the F1-score up to **66%**, and is nearly close to target only performance in some cases.
- Our proposed POND model is more stable than all comprise methods.

Scenario	Raincoat	CoDATs	Deep_Coral	MMDA	DIRT	DSAN	POND	Target Only
HAR 1-15 → 16	0.823 ± 0.094	0.767 ± 0.093	0.773 ± 0.082	0.679 ± 0.084	0.612 ± 0.135	0.738 ± 0.095	**0.849 ± 0.021**	0.856 ± 0.027
HAR 1-15 → 20	0.872 ± 0.142	0.932 ± 0.025	0.923 ± 0.023	0.921 ± 0.034	0.848 ± 0.101	0.929 ± 0.033	**0.968 ± 0.021**	0.983 ± 0.018
HAR 1-15 → 21	0.867 ± 0.141	0.903 ± 0.070	0.882 ± 0.028	**0.974 ± 0.039**	0.921 ± 0.090	0.909 ± 0.110	0.972 ± 0.021	1.000 ± 0.000
HAR 1-15 → 28	0.766 ± 0.107	0.775 ± 0.166	**0.852 ± 0.044**	0.778 ± 0.085	0.671 ± 0.175	0.783 ± 0.046	0.829 ± 0.018	0.853 ± 0.019
HAR 16-20 → 1	0.792 ± 0.072	0.744 ± 0.053	0.667 ± 0.077	0.654 ± 0.074	0.546 ± 0.060	0.698 ± 0.037	**0.883 ± 0.017**	0.986 ± 0.010
HAR 16-20 → 2	0.825 ± 0.048	0.821 ± 0.151	0.796 ± 0.055	0.651 ± 0.045	0.509 ± 0.050	0.652 ± 0.057	**0.936 ± 0.017**	0.943 ± 0.024
HAR 16-20 → 3	0.814 ± 0.028	0.746 ± 0.078	0.741 ± 0.058	0.657 ± 0.033	0.605 ± 0.056	0.565 ± 0.043	**0.878 ± 0.018**	0.978 ±0.013
HAR 16-20 → 4	0.679 ± 0.084	0.605 ± 0.082	0.479 ± 0.110	0.513 ± 0.058	0.336 ± 0.110	0.436 ± 0.032	**0.754 ± 0.033**	0.921 ± 0.018
WISDM 0-17 → 18	0.379 ± 0.061	0.384 ± 0.049	0.346 ± 0.023	0.297 ± 0.016	0.300 ± 0.041	0.287 ± 0.045	**0.606 ± 0.020**	0.705 ± 0.046
WISDM 0-17 → 20	0.354 ± 0.040	0.368 ± 0.039	0.376 ± 0.031	0.452 ± 0.098	0.347 ± 0.071	0.269 ± 0.064	**0.570 ± 0.023**	0.704 ± 0.051
WISDM 0-17 → 21	0.355 ± 0.057	0.310 ± 0.088	0.259 ± 0.018	0.250 ± 0.000	0.276 ± 0.055	0.245 ± 0.046	**0.450 ± 0.026**	0.636 ± 0.095
WISDM 0-17 → 23	0.306 ± 0.015	0.327 ± 0.075	0.318 ± 0.031	0.327 ± 0.023	0.271 ± 0.016	0.277 ± 0.044	**0.482 ± 0.017**	0.538 ± 0.034
WISDM 0-17 → 25	0.365 ± 0.030	0.540 ± 0.125	0.435 ± 0.043	0.436 ± 0.094	0.314 ± 0.107	0.353 ± 0.120	**0.559 ± 0.050**	0.672 ± 0.039
WISDM 0-17 → 28	0.399 ± 0.028	0.431 ± 0.033	0.418 ± 0.032	0.454 ± 0.064	0.304 ± 0.044	0.339 ± 0.030	**0.656 ± 0.046**	0.689 ± 0.048
WISDM 0-17 → 30	0.314 ± 0.020	0.305 ± 0.028	0.298 ± 0.023	0.359 ± 0.072	0.266 ± 0.035	0.246 ± 0.076	**0.670 ± 0.039**	0.791 ± 0.028
WISDM 18-23 → 5	0.648 ± 0.001	0.558 ± 0.129	0.534 ± 0.102	0.510 ± 0.020	0.549 ± 0.097	0.484 ± 0.055	**0.652 ± 0.035**	0.734 ± 0.095
WISDM 18-23 → 6	0.544 ± 0.074	0.565 ± 0.143	0.437 ± 0.078	0.543 ± 0.160	0.405 ± 0.089	0.454 ± 0.112	**0.628 ± 0.033**	0.872 ± 0.049
WISDM 18-23 → 7	0.588 ± 0.070	0.404 ± 0.117	0.530 ± 0.094	0.477 ± 0.060	0.518 ± 0.120	0.476 ± 0.127	**0.672 ± 0.029**	0.888 ± 0.035
HHAR 0-6 → 7	0.765 ± 0.142	0.652 ± 0.108	0.815 ± 0.105	0.641 ± 0.050	0.649 ± 0.005	0.730 ± 0.164	**0.834 ± 0.014**	0.861 ± 0.016
HHAR 5-8 → 2	0.321 ± 0.023	0.347 ± 0.082	0.309 ± 0.032	0.216 ± 0.032	0.276 ± 0.021	0.314 ± 0.095	**0.352 ± 0.014**	0.881 ± 0.018
SSC 0-9 → 16	0.578 ± 0.028	0.510 ± 0.044	0.537 ± 0.024	0.559 ± 0.027	0.523 ± 0.019	0.515 ± 0.044	**0.568 ± 0.012**	0.601 ± 0.018
SSC 0-9 → 17	0.511 ± 0.024	0.413 ± 0.118	0.452 ± 0.077	0.504 ± 0.060	0.530 ± 0.053	0.463 ± 0.081	**0.559 ± 0.006**	0.602 ± 0.014
SSC 0-9 → 18	**0.605 ± 0.016**	0.548 ± 0.037	0.544 ± 0.046	0.597 ± 0.032	0.574 ± 0.021	0.569 ± 0.046	0.604 ± 0.014	0.602 ± 0.013
SSC 0-9 → 19	0.562 ± 0.024	0.540 ± 0.052	0.531 ± 0.055	**0.570 ± 0.044**	0.565 ± 0.028	0.568 ± 0.080	**0.570 ± 0.010**	0.613 ± 0.019
SSC 10-12 → 8	0.294 ± 0.028	0.380 ± 0.066	0.379 ± 0.076	0.398 ± 0.060	0.322 ± 0.048	0.411 ± 0.046	**0.470 ± 0.010**	0.531 ± 0.019

KDD2024
BARCELONA, SPAIN

Experiments: Ablation Study

- In the ablation study, we explore the necessity of the MoE technique, common prompt, and prompt generator.
In the following table, the first three columns show whether three components are included in the model. Other columns show different scenarios on the WISDM dataset.

- Experimental results show that they all contributes to the outstanding performance of our POND model. Any missing component in the model will lead to significant performance drop.

MoE	Common Prompt	Prompt Generator	0-17→ 22	0-17→ 23	0-17→ 24	0-17→ 25	18-23→ 5	18-23→ 6	Overall
✗	✓	✗	0.622±0.057	0.415±0.015	0.510±0.030	0.581±0.036	0.623±0.058	0.516±0.038	0.545±0.039
✗	✗	✓	0.646±0.064	0.396±0.048	0.527±0.030	0.573±0.034	0.628±0.051	0.512±0.057	0.547±0.047
✗	✓	✓	0.632±0.069	0.384±0.041	0.498±0.032	0.572±0.045	0.611±0.055	0.514±0.025	0.535±0.045
✓	✗	✓	0.575±0.043	0.349±0.029	0.517±0.032	0.584±0.030	0.621±0.056	0.578±0.035	0.537±0.038
✓	✓	✗	0.719±0.062	0.405±0.052	0.529±0.042	0.588±0.034	0.616±0.050	0.565±0.049	0.570±0.048
✓	✓	✓	**0.725±0.031**	**0.482±0.017**	**0.559±0.050**	**0.695±0.035**	**0.652±0.035**	**0.628±0.033**	**0.624±0.034**

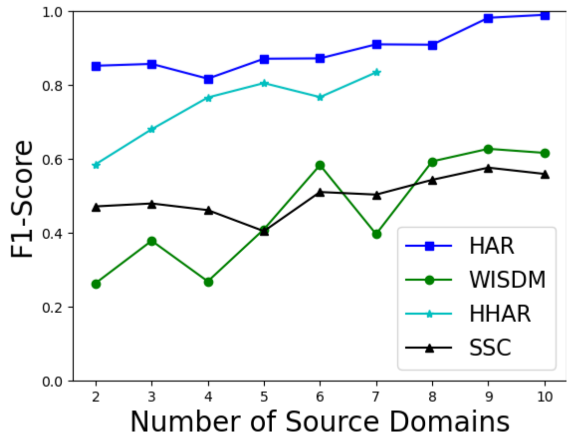
KDD2024
BARCELONA, SPAIN

Experiments: Sensitivity Analysis

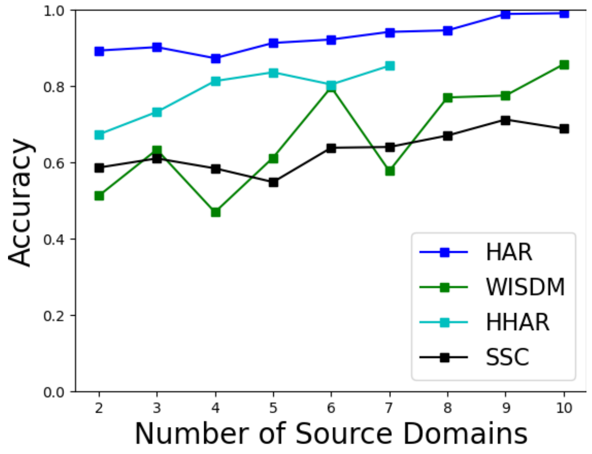
- We explore how source domains influence performance on the target domain, averaged by ten implementations. Figures (a) and (b) demonstrate F1-score and accuracy versus the number of source domains, respectively. The HHAR dataset has less than 10 domains.

- Generally, our proposed POND model demonstrates improved performance with an increasing number of source domains.

- However, in some cases, the performance drops with the increasing number of source domains.



(a). F1-score versus source domains



(b). Accuracy versus source domains

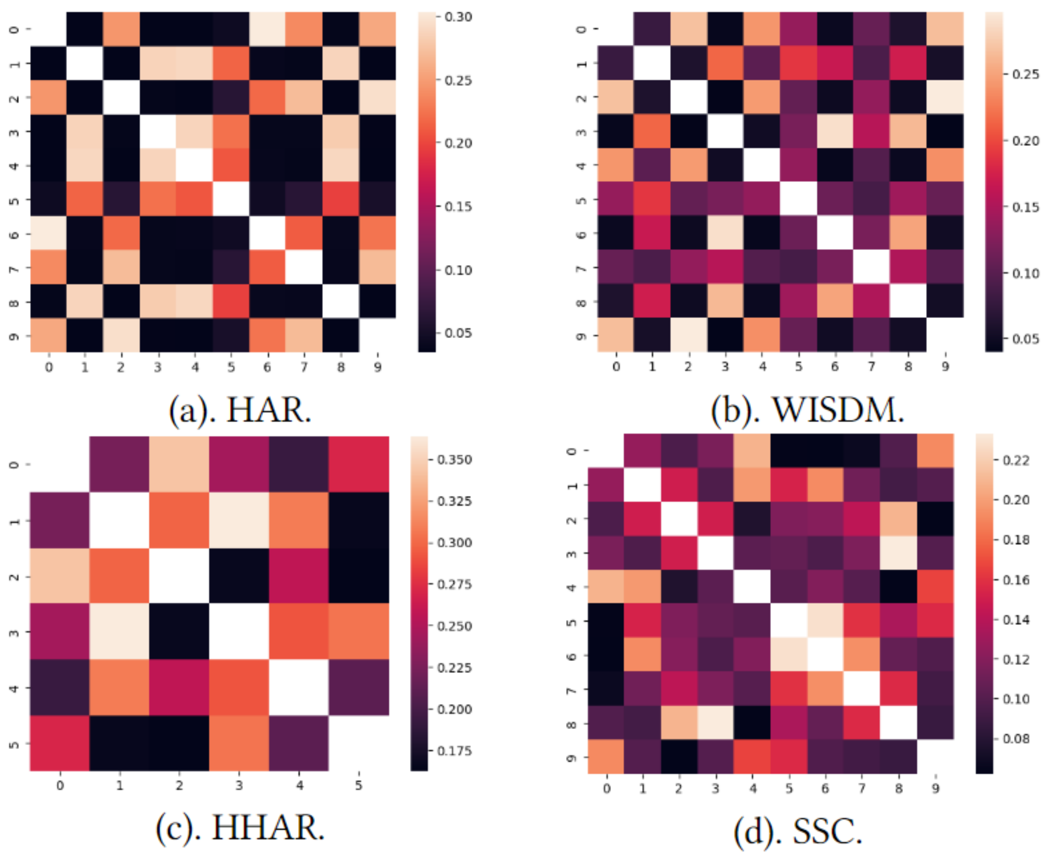
KDD2024
BARCELONA, SPAIN

Experiments: Discrimination Loss

- Finally, we present a visualization of the discrimination loss for pairwise source domains. The right figure illustrates the exponents of discrimination losses for all pairs of source domains across four datasets. Both the X-axis and Y-axis represent the indexes of source domains.

- Darker colors indicate smaller discrimination losses, reflecting better domain discrimination. The diagonals are left blank.

- Overall, our proposed POND model effectively discriminate most source domains, as evidenced by the predominance of dark squares.



(a). HAR. (b). WISDM.

(c). HHAR. (d). SSC.

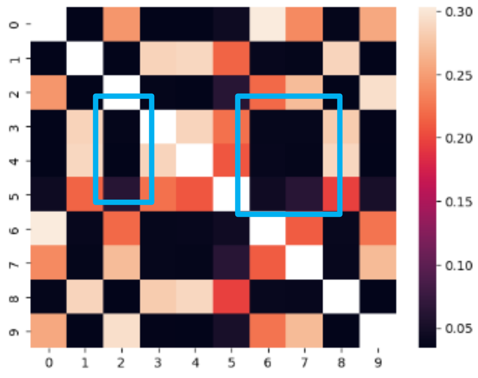
KDD2024 BARCELONA, SPAIN

Experiments: Discrimination Loss

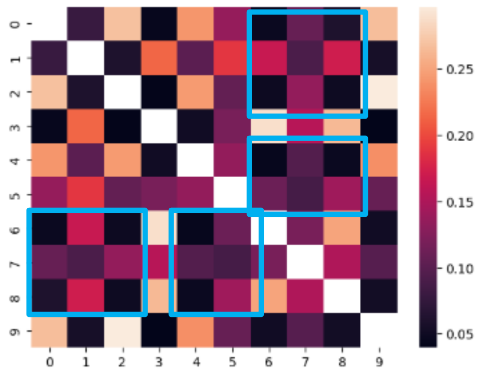
- Finally, we present a visualization of the discrimination loss for pairwise source domains. The right figure illustrates the exponents of discrimination losses for all pairs of source domains across four datasets. Both the X-axis and Y-axis represent the indexes of source domains.

- Darker colors indicate smaller discrimination losses, reflecting better domain discrimination. The diagonals are left blank.

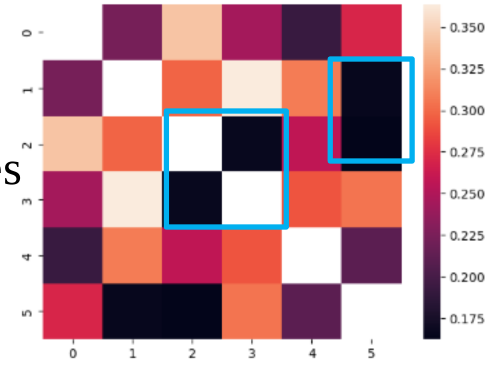
- Overall, our proposed POND model effectively discriminates most source domains, as evidenced by the predominance of dark squares.



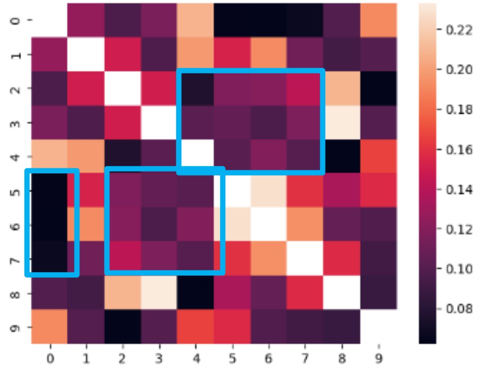
(a). HAR.



(b). WISDM.



(c). HHAR.



(d). SSC.

KDD2024
BARCELONA, SPAIN

Outline

Background

Proposed Method - POND

Experiment

Conclusion

KDD2024
BARCELONA, SPAIN

Key Takeaways

- Multi-source time series domain adaptation is a challenging task. Instead of learning domain-invariant representations, we capture domain-specific information by prompts for domain adaptation.

- We propose a prompt generator to capture the dynamic aspect of domain-specific information.

- Two important criteria fidelity and distinction are proposed to select good prompts for the evaluation of learned domain-specific information.

- Extensive experiments on 50 scenarios on four benchmark datasets demonstrate the effectiveness and robustness of our proposed POND model.

KDD2024
BARCELONA, SPAIN

THANK YOU

KDD2024
BARCELONA, SPAIN