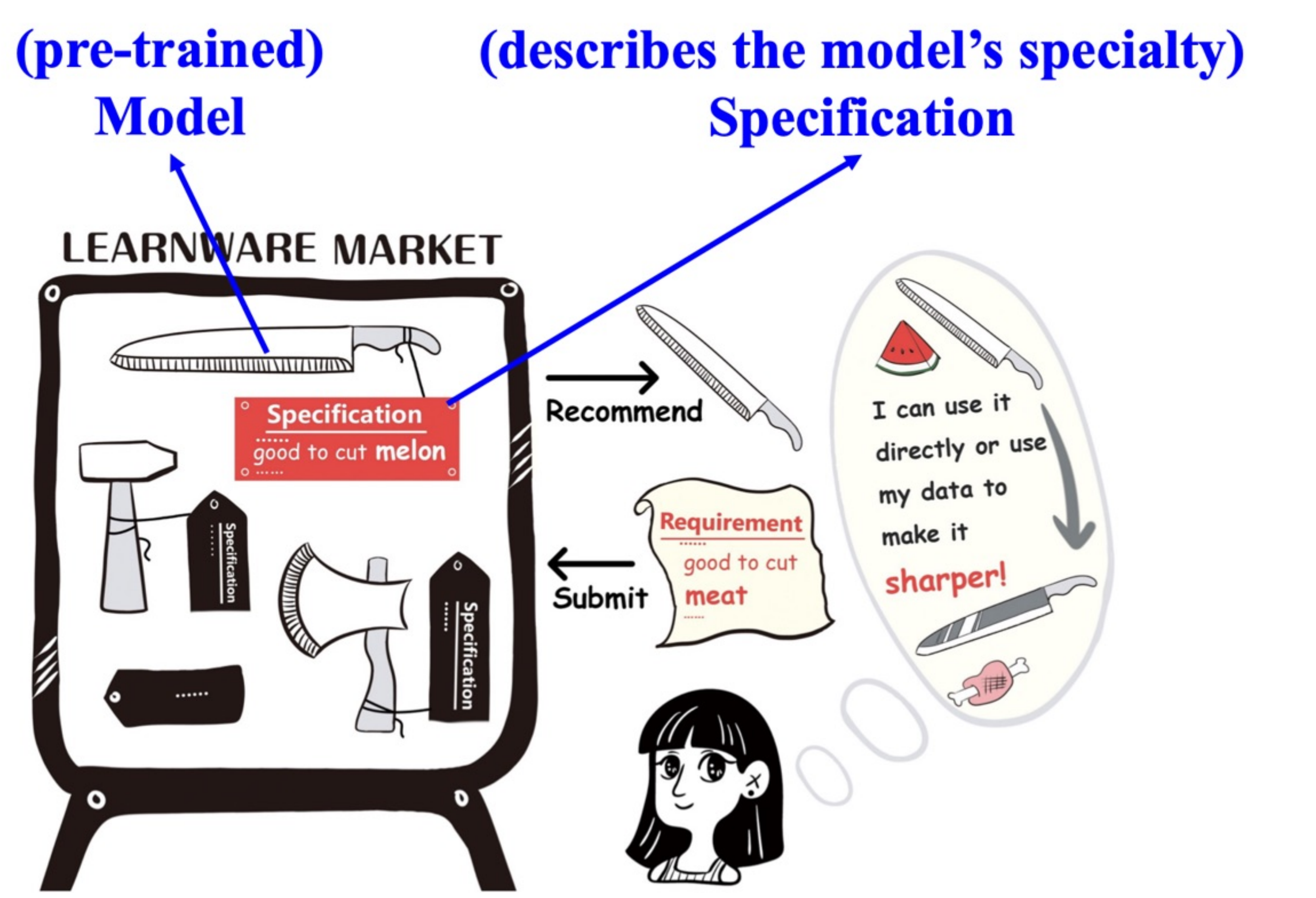
Identifying Useful Learnwares for Heterogeneous Label Spaces

Lan-Zhe Guo, Zhi Zhou, Yu-Feng Li, Zhi-Hua Zhou
National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China

ICML International Conference On Machine Learning

LAMDA Learning And Mining from DatA

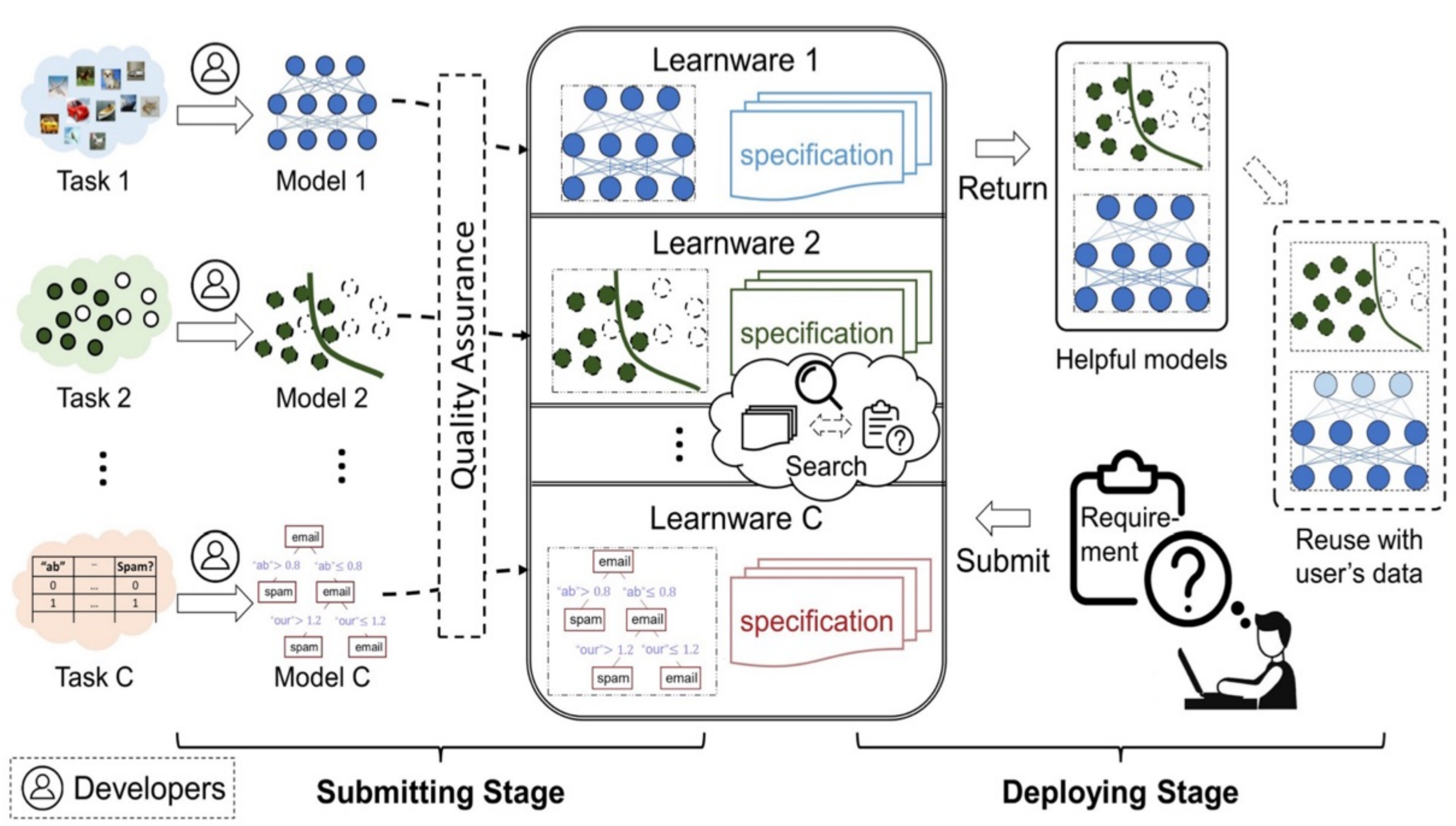
Learnware = Model + Specification



[Zhou & Tan, SCIS 2023]

The specification enables the model to be identified to reuse according to the requirement of future users

Learnware Paradigm



Note

The learnware market is different from model pool like HuggingFace, where models are used "**as-what-was-submitted**", whereas learnware enables models to be used "**beyond-what-was submitted**"

Our Method

In the submitting stage

Problem: How to assign specifications to the submitted model?

Basic Idea: The model's specialty should be related to the training data distribution and the model's functionality

Our Solution

✓Describe the data distribution via Reduced Kernel Mean Embedding

$$\min_{\beta,U} \left\| \frac{1}{n}\sum_{i=1}^{n} k(Z_i,\cdot) - \sum_{j=1}^{m} \beta_j k(U_j,\cdot) \right\|_{\mathcal{H}}^2$$

$k(\cdot,\cdot)$ is the kernel function
U is the reduced set
β is the coefficient

✓Describe the model's functionality with a linear proxy model

$$\min_{W} \mathcal{L}_{CE}(ZW,Y) + \mathcal{L}_{KL}(ZW,f(X))$$

Specification: $\{\Phi = \{\beta, U\}, W\}$

In the deploying stage

Problem: How to identify useful learnwares given a user's requirements?

Basic Idea: Matching the user's requirements with the model's specification by considering the class-specific specification explicitly

Our Solution

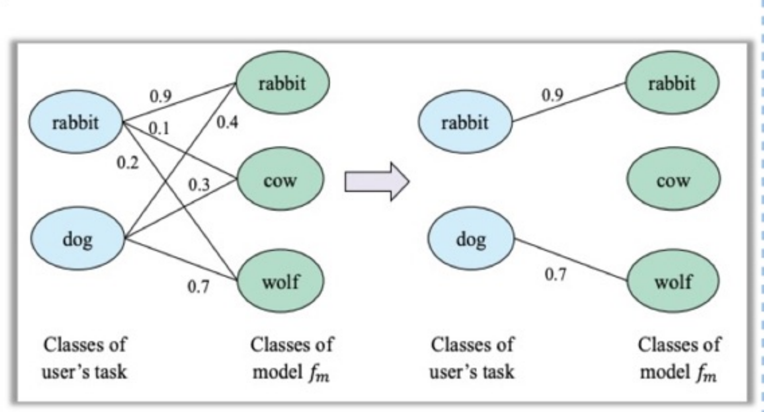
✓Generate RKME and the proxy model's parameter with the user's data

$$\Phi_T = \frac{1}{N_T}\sum_{i=1}^{N_T} k(Z_{T,i},\cdot), \qquad \min_{W_T} \mathcal{L}_{CE}(Z_T W_T, Y_T)$$

✓Identify multiple candidate learnwares via Φ_T

✓For class k in the user's task, select a learnware via $W_{T,k}$

✓Compute the matching score for each learnware



What Can Learnware do?

Learnware offers the possibility of addressing issues:

✓ **Lack of training data**
Only small data are needed for learnware search and adaptation

✓ **Lack of training skills**
No need to train a model from scratch

✓ **Catastrophic forgetting**
The model always be accommodated in the market

✓ **Continual learning**
New knowledge with the new submission

✓ **Data privacy**
Only submit models without sharing data

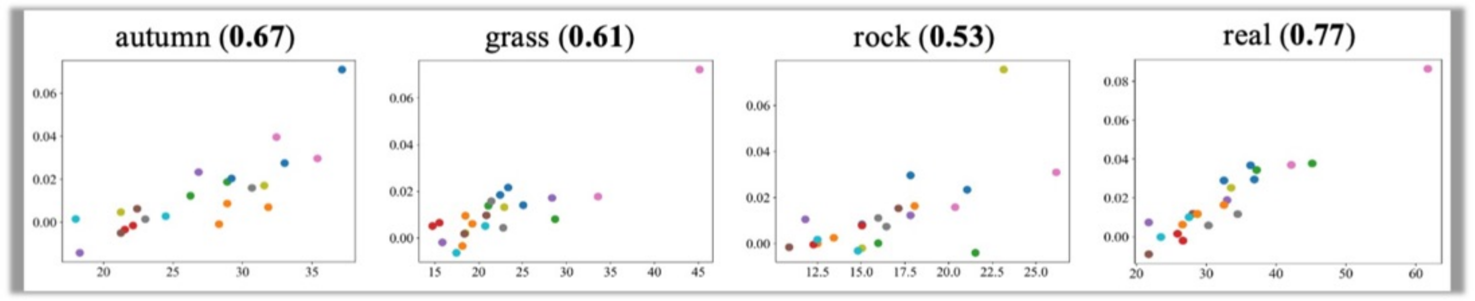
Experiments

We manually constructed a learnware market with 21 learnwares

➢Can the most useful model be identified via the specification?

Settings	Methods	Pre@1	Pre@2	Pre@3
Homo-direct use	RKME-basic	54.54	81.82	95.45
	Ours	**95.15**	**100.0**	**100.0**
Homo-fine-tuning	RKME-basic	40.90	68.18	81.82
	Ours	**81.82**	**90.91**	**90.91**
Hetero-fine-tuning	RKME-basic	36.36	45.45	50.00
	Ours	**59.09**	**63.63**	**68.18**

➢How about the correlation between model reuse performance and specification similarity



Kendall's coefficient between performance (X-axis) and similarities (Y-axis)

Take Home Messages

➢**Learnware provides a promising way to connect ML models to accomplish different AI tasks**

➢**Specification plays a pivotal role in the learnware paradigm**

➢**This paper provides a powerful specification considering both data distribution and model functionality**

Mail

waitting for your feedback