META LEARNING WITH MINIMAX REGULARIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Even though meta-learning has attracted wide attention in recent years, the generalization problem of meta-learning is still not well addressed. Existing works focus on meta-generalization to unseen tasks at the meta-level, while ignoring that adapted-models may not be generalized to the tasks domain at the adaptation-level, which can not be solved trivially. To this end, we propose a new regularization mechanism for meta-learning – Minimax-Meta Regularization. Especially, we maximize the regularizer in the inner-loop to encourage the adapted-model to be more sensitive to the new task, and minimize the regularizer in the outer-loop to resist overfitting of the meta-model. This adversarial regularization forces the meta-algorithm to maintain generality at the meta-level while it is easy to learn specific assumptions at the task-specific level, thereby improving the generalization of meta-learning. We conduct extensive experiments on the representative meta-learning scenarios to verify our proposed method, including few-shot learning and robust reweighting. The results show that our method consistently improves the performance of the meta-learning algorithms and demonstrates the effectiveness of Minimax-Meta Regularization.

1 INTRODUCTION

Meta-learning has been proven to be a powerful paradigm for extracting well-generalized knowledge from data and accelerating the learning process for new tasks (Thrun & Pratt, 2012). It simulates the machine learning process by a bi-level objective (Finn et al., 2017), evaluating the query (meta-validation) set with an adapted-model learned from the meta-model by the support (meta-training) set. Meta-learning has received increasing attention in many machine learning settings such as few-shot learning (Sung et al., 2018; Sun et al., 2019; Wang et al., 2020) and robust learning (Ren et al., 2018; Shu et al., 2019; Li et al., 2019), and can be deployed in many practical applications (Kang et al., 2019; Dou et al., 2019; Yu et al., 2018; Madotto et al., 2019). Despite the success, the additional level of learning creates another potentially overfiting (Rajendran et al., 2020b), which significantly challenges the generalization of meta-learning algorithms. Specifically, the meta-model should be generalized to unseen tasks (meta-generalization). In the meanwhile, the adapted-model should be generalized to the domain of a specific task, which we called *adaptation-generalization* (Figure 1). A key challenge is how to *regularize* the meta-algorithms to ensure this *two-levels generalization*.

The deep neural networks tend to overfit the sampling bias due to its representation power, leading to poor generalization (Song et al., 2020). Regularizations such as weight decay (Krogh & Hertz, 1992), dropout (Gal & Ghahramani, 2016), and incorporating noise (Tishby & Zaslavsky, 2015; Alemi et al., 2016; Achille & Soatto, 2018), can effectively present the model from the overfitting and enhance the generalization. However, direct applying the regularization to the networks limited the flexibility of fast adaptation in the inner loop (meta-training) of meta-learning (Yao et al., 2021). Recent works aim to address the meta-generalization problem by meta-regularizations, such as constraining the meta-initialization space (Yin et al., 2019), enforcing the similarity of the performance of the meta-model on different tasks (Jamal & Qi, 2019), and augmenting meta-training data (Rajendran et al., 2020b; Ni et al., 2021; Yao et al., 2021). These methods significantly enhance the generalization for unseen tasks. However, they ignore the adaptation-generalization to the data distribution of the meta-testing tasks (Figure 1), which is not negligible.

The work takes the first step further to optimize both meta-generalization and adaptation-generalization for meta-learning. However, the adaptation-generalization is significant challenging

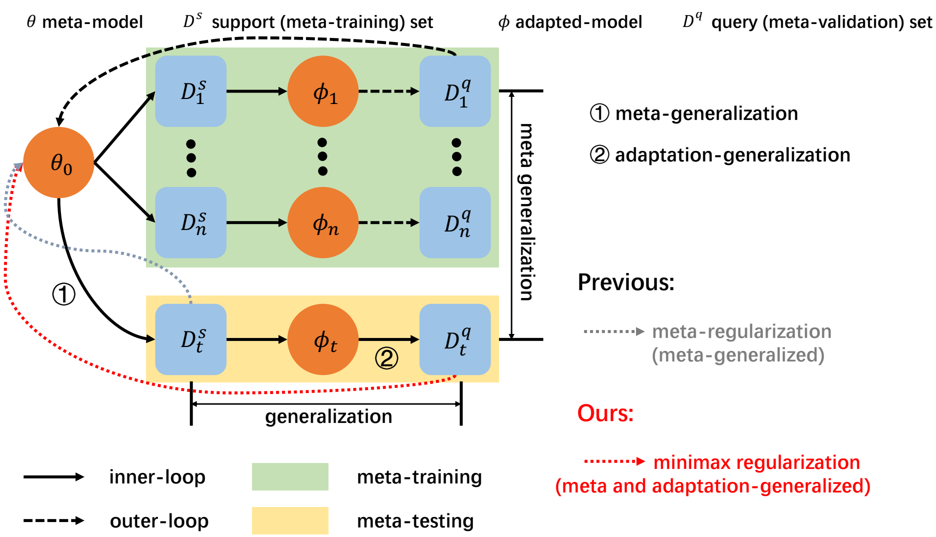

Figure 1: Illustration of our proposal and previous works. A robust meta-learning should eschew both *meta overfitting*, where the meta-model overfits to meta-training tasks due to the limited number of tasks, and *adaption overfitting*, where the adapted model overfits to few-shot samples from meta-testing due to the few-shot number. While previous works only consider the meta-generalization with meta-regularization to alleviate the meta overfitting and ignore the potential adaption overfitting. Our Minimax-Meta Regularization enables the meta-learning algorithm to be generalized at both meta-level and adaptation-level, which can significantly enhance the generality.

for meta-learning, where we meet a *dilemma between fast adaptation and generality*: 1) regularizing the model during meta-testing time can enhance the generalization to the task domain, however, limits the fast adaptation that is the goal of meta-learning; 2) exacerbating the overfitting to the few-shot samples from meta-testing can enhance the fast adaptation, however, limits the generality to the task domain.

To address the challenge, we consider learning a meta-model resistant to the adapted-model overfitting during the meta-testing time. To achieve this, we design a well-general mechanism called *Minimax-Meta Regularization* for meta-learning. During the meta-training, we enforce the adapted-model to be more overfitting to the support data by adding a *inverse (negative) regularization in the inner loop*, and enforce the meta-model to be more generalized on the test samples by adding a *positive regularization in the outer loop*. By doing so, the learned meta-model can be meta-generalized, making adapted-models perform well on the query (meta-validation) set, even when the adapted-models are prone to overfit to the support (meta-training) set. Therefore during the meta testing, the adapted-model can still be generalized to the task domain, even though they are overfitting to few-shot samples. In particular, the Minimax-Meta Regularization is well general to be implemented in all bi-level optimization frameworks without additional computational cost.

To verify the above intuition, we conduct experiments of the basic MAML (Finn et al., 2017) framework. Surprisingly, we find that both positively regularizing the outer loop meta-training and negatively regularizing the inner loop adaption can significantly enhance the few-shot classification. Another interesting finding is that adding positive regularization in the inner loop impairs the performance, which indirectly proves the efficacy of our proposal. We conduct extensive experiments on few-shot regression, few-shot classification, and robust reweighting (Ren et al., 2018). The experimental results show that Minimax-Meta Regularization generally improves the performance of bi-level meta-learning algorithms and is compatible with common methodologies for enhancing meta-learning. Moreover, Minimax-Meta Regularization shows the capability to improve the generalization of meta-learning algorithms and help address meta-overfitting problems to a certain extent.

Our Contributions. 1) we propose a limitation of previous works on meta-generalization that ignore the adaptation-generalization; **2)** we design a general mechanism named Minimax-Meta Regularization for meta-learning, which aims to capture a meta-model that is both meta-generalized and resistant to the adaptation overfitting; **3)** we empirically verify the intuition of Minimax-Meta Regularization and give possible reasons; **4)** we conduct three different bi-level optimization tasks to show the efficacy of the proposed method.

2 PRELIMINARY

We first give a brief introduction and notation of meta-learning. In the meta-learning problem setting that we consider, the goal is to learn a generalized initialization model for better adapting to new tasks from only a few samples. To achieve this, it requires a set of support (meta-training) data $\left\{\mathcal{D}_i^s = \{x_{i,j}^s, y_{i,j}^s\}_{j=1}^k\right\}_{i=1}^n$ and query (meta-testing) data $\left\{\mathcal{D}_i^q = \{x_{i,j}^q, y_{i,j}^q\}_{j=1}^m\right\}_{i=1}^n$ sampled from tasks $\{\mathcal{T}_i\}_{i=1}^n$ drawn from distribution $p(\mathcal{T})$, where k and m denote the number of data samples from support and query data, and n is the number of tasks. Denote $\mathcal{L}$ and μ to be the loss function and inner-loop learning rate.

Meta-learning (Finn et al., 2017) simulates the adaptation and evaluation procedure of machine learning, and aims to learn a well-generalized model f parameterized by θ^* by the following bi-level optimization

$$\theta^* = \arg\min_{\theta} \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m}\sum_{z\in\mathcal{D}_i^q}\mathcal{L}(\phi_i(\theta,\mathcal{D}_i^s),z), \quad s.t. \phi_i(\theta,\mathcal{D}_i^s) = \theta - \mu\nabla_\theta \sum_{z\in\mathcal{D}_i^s}\mathcal{L}(\theta,z) \tag{1}$$

where z represents the data sample (x, y). The outer loop (represents the meta-validation phase) measures the generalization performance of the adapted-model ϕ_i by the query data $\mathcal{D}_i^q$. The inner loop (represents the meta-training phase) defines that the adapted-model ϕ_i is finetuned from initialization θ by multiple steps gradient descent with the support data $\mathcal{D}_i^s$. Note that gradient steps can be more than one, the formulation 1 is written for shortness.

3 META LEARNING WITH MINIMAX-META REGULARIZATION

We aim to learn a well-generalized meta-initialization that can fast adapt to new tasks with robust performance. To achieve this, the meta-learner should be meta-generalized, i.e. learn a meta-model θ that is robust to tasks distribution $p(\mathcal{T})$, and adaptation-generalized, i.e. the adapted model $\phi(\theta,\mathcal{D}^s), \mathcal{D}^s \sim \mathcal{T}$ should be robust to the data distribution of the task domain $\mathcal{T}$. The meta-generalization problem has been studied in many previous works (Yin et al., 2019; Collins et al., 2020; Yao et al., 2021; Ni et al., 2021), and can be addressed by designing regularization in the outer loop. However, due to the limited number of samples of $\mathcal{D}^s, \mathcal{D}^q$, the adaptation-generalization problem is significantly challenging for meta-learning.

To address this, we propose a novel and well-general regularization framework for meta-learning – Minimax-Meta Regularization. In this section, we first present the training objective for minimax-regularized meta-learning while giving the intuition behind the design, and run a simulation to verify the high-level insight.

3.1 TRAINING OBJECTIVE

Based on the formulation 1 for meta-learning, we present the minimax-regularized meta-learning training objective as follows, where we add a positive regularization in the outer loop to achieve meta-generalization, and an inverse (negative) regularization in the inner loop to achieve adaptation-generalization.

$$\theta^* = \arg\min_{\theta} \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m}\sum_{z\in\mathcal{D}_i^q}\mathcal{L}(\phi_i(\theta,\mathcal{D}_i^s),z) + \lambda^{out}\frac{1}{n}\sum_{i=1}^{n}\mathbf{Reg}^{out}(\phi_i(\theta,\mathcal{D}_i^s)), \tag{2}$$

$$s.t. \quad \phi_i(\theta,\mathcal{D}_i^s) = \arg\min_{\phi}\langle\mu\nabla_\theta \sum_{z\in\mathcal{D}_i^s}\mathcal{L}(\theta,z),\phi\rangle + \frac{1}{2}\|\phi-\theta\|^2 - \lambda^{in}\frac{1}{n}\sum_{i=1}^{n}\mathbf{Reg}^{in}(\phi), \tag{3}$$

where $\mathbf{Reg}^{out}$ and $\mathbf{Reg}^{in}$ are the regularizations in the outer and inner loop, while $\lambda^{out} \geq 0$ and $\lambda^{in} \geq 0$ are the coefficients respectively. Note the the formulation $\phi_i(\theta,\mathcal{D}_i^s) =$

Table 1: Simulation Verification.

Mini-Imagenet 5-way Few-Shot Classification for MAML				
Regularization Type	Outer Reg	Inner Reg	1-Shot	5-Shot
no regularization	0	0	48.75±0.44%	64.37±0.30%
regularize the loss fuction	+	+	49.15±0.50%	66.00±0.14%
regularize the outer-loop	+	0	50.31±0.48%	68.79±0.15%
regularize the inner-loop	0	+	46.89±0.42%	64.03±0.36%
inverse regularize the inner-loop	0	-	49.92±0.23%	65.61±0.38%
minimax-meta regularization (ours)	+	-	**50.84±0.26%**	**69.40±0.17%**

$\arg\min_\phi \langle \mu \nabla_\theta \sum_{z \in \mathcal{D}_i^s} \mathcal{L}(\theta, z), \phi \rangle + \frac{1}{2}\|\phi - \theta\|^2$ in the inner loop is the equivalent mirror descent (Beck & Teboulle, 2003) version of the gradient descent. We next introduce the intuition behind this design.

Outer positive regularization. As defined in Eq 2, we add a positive regularization $\mathbf{Reg}^{out}(\phi_i(\theta, \mathcal{D}_i^s))$ that regularizes the model-overfitting of the adapted-model $\phi_i(\theta, \mathcal{D}_i^s)$. By doing so, the meta learner is enforced to learn a generalized meta-model θ^* such that the adapted-model $\phi_i(\theta^*, \mathcal{D}_i^s)$ on each tasks is not overfitting and generalized to query data. This idea has been studied in previous works (Yin et al., 2019; Collins et al., 2020; Yao et al., 2021; Ni et al., 2021), and has been shown to significantly enhance the meta-generalization.

Inner inverse regularization. The generalization performance depends not only on the complexity of the adapted-model $\phi_i(\theta, \mathcal{D}_i^s)$, but also the adaptation rule, i.e. the formulation of the inner loop function. As defined in Eq 3, we add a inverse regularization $\mathbf{Reg}^{in}(\phi)$ that negatively regularizes the model-complexity of the adapted-model $\phi_i(\theta, \mathcal{D}_i^s)$. By doing so, the inner loop function simulates the adaptation overfitting during meta-testing by enforcing the adapted-model to be overfitting during meta-training. Therefore, learning from the minimax regularized meta-learning, the learned meta-model θ^* can be resistant to adaptation overfitting.

From the above discussion, the Minimax-Meta Regularization enables the meta-learning to capture a meta-model that is *both meta-generalized and adaptation-generalized*. This framework is *computational efficient* without additional computational cost. In addition, the Minimax-Meta Regularization is *general* to all bi-level optimization formulation, thus can be directly applied to different meta-algorithms on different bi-level learning problems.

3.2 Empirical Verification.

We next verify the design by a simulation test by conducting the basic MAML framework (Finn et al., 2017) with different regularization types. As illustrated in Table 1, we make the following observations:

Outer positive regularization enhances the generalization performance. Compare the results from "no regularization" and "regularize the outer-loop", we observe that adding outer regularization can get 1.56% and 4.42% accuracy improvements in 1-shot and 5-shot experiments, which verifies the efficacy of the outer regularization. This is aligned with the intuition that outer regularization enhances the meta-generalization, leading to better performance.

Inner negative regularization enhances the generalization performance. Compare the results from "no regularization" and "inverse regularize the inner-loop", we observe that adding inner inverse regularization can get 1.17% and 1.24% accuracy improvements in 1-shot and 5-shot experiments, which verifies the efficacy of the inner inverse regularization. This is aligned with the intuition that inner reverse regularization enhances adaptation-generalization, thus improving performance.

The outer regularization and inner inverse regularization are compatible. Compare the results from "Minimax-Meta Regularization", "regularize the outer-loop", and "inverse regularize the inner-loop", we observe that Minimax-Meta Regularization can get 0.52% (1-shot)/0.61% (5-shot) and 0.92% (1-shot)/3.79% (5-shot) accuracy improvements than solely regularizing the outer-loop and inverse regularizing the inner-loop, which verifies the compatibility of the inner inverse regularization. This is aligned with the intuition that meta-generalization and adaptation-generalization are not in conflict.

Inner positive regularization impairs the generalization performance. Compare the results from "no regularization" and "regularize the inner-loop", we observe that adding inner positive regularization suffers from -1.86% and -0.34% accuracy impairments in 1-shot and 5-shot experiments, which aligns with the intuition that positive regularization that limits the adaptation in the inner-loop impairs the adaptation-generalization.

4 Related Work

Meta-learning. A line of meta-learning methods has sought to train recurrent neural networks that ingest entire datasets Santoro et al. (2016); Duan et al. (2016). However, they need to place constraints on the model architecture. Another line aims to learn a transferable metric space between samples from previous tasks (Vinyals et al., 2016; Snell et al., 2017; Mishra et al., 2018; Oreshkin et al., 2018). However, it is limited to classification problems. In this paper, we focus on gradient-based meta-learning methods that learn a meta-initialization (Finn et al., 2017; 2018; Li et al., 2017; Finn & Levine, 2018; Grant et al., 2018; Lee & Choi, 2018; Park & Oliva, 2019; Flennerhag et al., 2020), which is a well-generalized for meta-training tasks, being agnostic to both model architecture and problems. However, these approaches are shown to be overfitting the meta-training tasks and generalizing poorly to meta-testing tasks (Yoon et al., 2018; Collins et al., 2020; Rothfuss et al., 2021; Yao et al., 2021).

Meta-Regularization. Standard regularizations such as weight decay (Krogh & Hertz, 1992), dropout (Gal & Ghahramani, 2016), and incorporating noise (Tishby & Zaslavsky, 2015; Alemi et al., 2016; Achille & Soatto, 2018), which can significantly enhance the generality of single-loop machine learning. However, the straightforward method that regularizes the neural networks limits the flexibility of fast adaptation in the inner loop (Yao et al., 2021). Recently, a few works were proposed to design the meta-regularization to improve meta-generalization. MR-MAML (Yin et al., 2019) constrains the search space of the meta-model, and allows the adaptation to be sufficient in the inner loop. Jamal & Qi (2019) proposed TAML to enforce the meta-model to perform similarly across tasks. Rajendran et al. (2020a) explored an information-theoretic framework of meta-augmentation by adding randomness to labels of both support and query sets. Yao et al. (2021) proposed two task augmentation methods – MetaMix and Channel Shuffle, which is theoretically proved to be generalized to unseen tasks. Ni et al. (2021) investigated the distinct ways where data augmentation can be integrated at both the image and class levels. Rothfuss et al. (2021) addressed the meta-generalization problem using the PAC-Bayesian framework, and proposed PACOH that is PAC-optimal with Gaussian processes. However, these works focus only on the meta-generalization, i.e., generalize to the unseen tasks, while the adaptation-generalization that measures how the adapted-model generalizes to the task domain is merely considered.

This paper proposes the Minimax-Meta Regularization for meta-learning, implementing a positive regularization in the outer-loop and a negative regularization in the inner-loop. The framework can enhance both meta-generalization and adaptation-generalization, and thus improve the performance.

5 Experiments

In this section, we conduct extensive experiments on three types of classical meta-learning tasks including, few-shot classification, few-shot regression, and robust reweighting with meta-learning, to demonstrate the efficacy of our proposed methods. With these experiments, we demonstrate that our methods i) outperform previous meta-learning algorithms in terms of predictive accuracy; ii) mitigate the meta-overfitting effectively. We will introduce the experimental setup, results, and analysis in the following subsections.

5.1 Few-shot Classification

We first carry out experiments on the few-shot classification task, one of the most popular tasks to evaluate meta-learning algorithms. To verify the effectiveness of our approach, we adapt Minimax-

Table 2: Omniglot 20-way 1-shot experiment. We report the test accuracy in the last epoch with 95% confidence interval for the mean over 3 runs.

Omniglot 20-way 1-Shot Classification	
	Accuracy
Siamese Nets	88.2%
Matching Nets	93.8%
Neural Statistician	93.2%
Memory Mod.	95.0%
Meta-SGD	95.93±0.38%
Meta-Networks	97.00%
MAML	94.20± 0.941%
Minimax-MAML (ours)	95.76±0.32%
MAML++	97.21±0.51%
Minimax-MAML++ (ours)	**97.77±0.05%**

Table 3: Mini-Imagenet 5-way few-shot experiment. We report the test accuracy in the last epoch with 95% confidence interval for the mean over 3 runs.

Mini-Imagenet 5-way Few-Shot Classification		
Approach	Accuracy	
	1-Shot	5-Shot
Matching Nets	43.56%	55.31%
Meta-SGD	50.47±1.87%	64.03±0.94%
Meta-Networks	49.21%	-
MAML	48.75±0.44%	64.37±0.30%
Minimax-MAML (ours)	50.84±0.26%	69.40±0.17%
MAML++	52.25±0.19%	68.38±0.46%
Minimax-MAML++ (ours)	**52.80±0.06%**	**71.43±0.38%**

Meta Regularization into bi-level optimization meta-learning algorithms and make a benchmark to compare with other methods.

5.1.1 EXPERIMENTAL SETUP

Datasets. For the few-shot classification task, we experiment on the public released datasets Mini-Imagenet (Ravi & Larochelle, 2017; Vinyals et al., 2016) and Omniglot (Lake et al., 2015), following the few-shot benchmark setting provided in (Antoniou et al., 2018). The Omniglot dataset is a collection of 1623 character classes with different alphabets. Each class in the dataset contains 20 instances. In the experiment, all the character classes are shuffled, and then the shuffled classes are divided into the training set, validation set, and test set, with 1150, 50, and 423 instances respectively. Rotation augmentation is applied to the images with 90-degree increments to create new classes. The second dataset used in the few-shot classification experiment is Mini-Imagenet (Ravi & Larochelle, 2017), which is sampled from ImageNet with 600 instances of 100 classes. Each image is resized into 84×84. Following the work (Ravi & Larochelle, 2017), we split the Mini-Imagenet dataset into 64 classes for training, 12 classes for validation, and 24 classes for testing.

Experimental details. We select MAML (Finn et al., 2017) as the representative bi-level optimization meta-leanring model. To evaluate the effectiveness of Minimax-Meta Regularization, we first begin the experiment with the baseline MAML on the 5-way 1/5-shot Mini-Imagenet setting. Then, on top of the original MAML, we implement Minimax-MAML by adding Minimax-Meta Regularization. We then compare Minimax-MAML with original MAML and other meta-learning baselines on the 5-way 1/5-shot Mini-Imagenet setting and the the 20-way 1-shot Omniglot setting. The compared baselines include *Matching Networks* (Vinyals et al., 2016), *Meta-SGD* (Li et al., 2017), *Meta-Networks* (Munkhdalai & Yu, 2017), *Siamese Nets* (Koch et al., 2015), *Neural Statistician* (Edwards & Storkey, 2016), and *Memory Module* (Kaiser et al., 2017). Here we also include MAML++ (Antoniou et al., 2018) in the experiment and further implement Minimax-MAML++ for comparison. MAML++ is an improved version of MAML, with 6 specific methodologies added together for the performance improvement of MAML. We include MAML++ in our experiment for studying two questions: i) By comparing Minimax-MAML with MAML++, we want to analyze if Minimax-Meta Regularization, as a general improving mechanism, has the potential to outperform algorithm-specific methodologies. ii) By comparing Minimax-MAML++ with MAML++, we want to evaluate if Minimax-Meta Regularization is compatible with complicated model-specific improving methodologies in bi-level optimization models. Note regularization is only added during the training phase. All the MAML/MAML++ experiments involving regularization share the same form of regularization objective. The regularization is achieved by combining the l2-norm regularization and output entropy regularization. More detailed experiment setting information could be find in Appendix B.

Table 4: Test MSE for the sinusoid regression under the non-mutually-exclusive setting. We implemented Minimax-Meta Regularization for MAML and MR-MAML, and compare them with the original methods. Each experiment is repeated for 5 times, we report the mean test MSE and std in parentheses.

Methods	MAML	Minimax-MAML(ours)	MR-MAML(A)	Minimax-MR-MAML(A)(ours)	MR-MAML(W)	Minimax-MR-MAML(W)(ours)
5-shot	0.686(0.080)	0.461(0.036)	0.229(0.045)	0.209(0.048)	0.179(0.050)	**0.136(0.005)**
10-shot	0.153(0.008)	0.125 (0.005)	0.121(0.017)	0.112(0.028)	0.065 (0.018)	**0.059(0.005)**

5.1.2 Results and Analysis

The baseline comparison results under Omniglot and Mini-Imagenet settings are shown in Table 2 and Table 3. Minimax-Meta Regularization are shown to improve both the original MAML and the MAML++ frameworks. In the Omniglot 20-way 1-shot classification experiment, the mean accuracy of MAML and MAML++ are improved from 94.20% and 97.21% to 95.76% and 97.77% respectively. Both the methods had unstable results in these experiments. After adopting Minimax-Meta Regularization, the std values of the final accuracy of these two methods have been significantly reduced, indicating better performance stability. The Minimax-MAML++ reached the best performance in this setting compared to other baselines with good stability. Significant improvements from Minimax-Meta Regularization are also shown in Mini-Imagenet 5-way 1/5-shot classification experiments. In the 1-shot experiments, the original MAML cannot outperform Meta-SGD and Meta-Networks baselines. The Minimax-Meta Regularization improves the accuracy of MAML from the average of 48.75% to 50.84%, which enables MAML to outperform other baselines. In the 5-shot experiments, Minimax-MAML could outperform MAML++ by 1.02%. Considering that MAML++ adopts 6 individual techniques specifically designed for MAML, Minimax-Meta Regularization shows strong effectiveness in this outperform as a general methodology.

5.2 Few-shot Regression

5.2.1 Experimental Setup

Datasets. For the few-shot regression task, we consider a non-mutually-exclusive regression problem based on the Sinusoids synthetic dataset. Each task of Sinusoids regression involves the regressing from the input to the output of a generated sine wave, where the amplitudes of the sinusoids are different among tasks. In our experiment, we follow the setting provided by (Yin et al., 2019). The Sinusoids data is created in the following way: the amplitude A of the sinusoid is uniformly sampled from a set of 20 scalars $\{0.1, 0.3, \cdots, 4\}$; u is sampled uniformly from $[-5, 5]$; and y is sampled from $\mathcal{N}(Asin(u), 0.1^2)$.

Experimental details. During the training, both u and A are provided as input of models, i.e. $x = (u, A)$. During the test time, we expand the range of the tasks by randomly sampling the amplitude A uniformly from [0.1, 4] and use a random one-hot vector as the input of the network. The meta-training tasks are a proper subset of the meta-test tasks. Under this setting, the amplitude input at the training phase makes this regression problem non-mutually-exclusive, which makes the meta-learning model prone to the memorization problem(Yin et al., 2019) during training. In the experiments, we compare with the representative bi-level optimization meta-learning baseline MAML (Finn et al., 2017), and the meta-regularized MAML (MR-MAML) (Yin et al., 2019) where the regularization is either on the activations (MR-MAML(A)) or the weights (MR-MAML(W)). Both MR-MAML(A) and MR-MAML(W) are initially designed for solving the memorization problem. The Minimax-Meta Regularization is implemented for all the above 3 methods with l2-norm as the regularization objectives for both the inner-loop and outer-loop.

5.2.2 Results and Analysis

Original MAML was shown to be capable of solving normal sinusoid few-shot regression problem(Finn et al., 2017). However, the results of non-mutually-exclusive sinusoid regression 5.2.2 suggest that added amplitude input makes MAML suffer from memorization problem and give poor test result. From the experiment result5.2.2, we could observe that Minimax-Meta Regularization improves the performance of MAML on both 5-shot and 10-shot tasks. In the 10-shot task, the

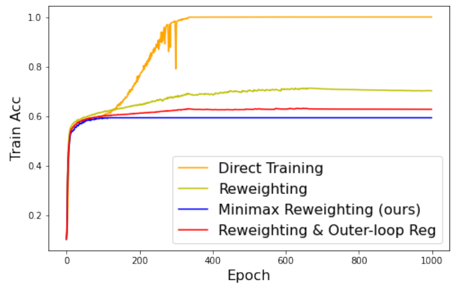

Figure 2: Training accuracy curve. Since 40% of training samples are with the wrong label, the model maintains train-acc around 60% would be considered resistant to overfitting on the training.

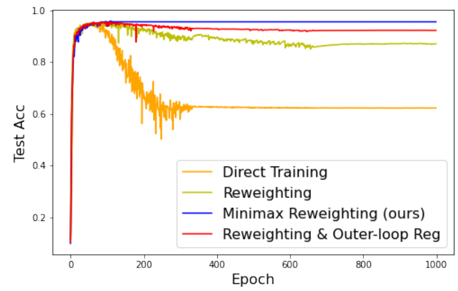

Figure 3: Test accuracy curve. Since the test dataset is clean, models that can maintain high test-acc show better learning robustness (less affected by training noise).

test MSE of MAML improved from 0.153 to 0.125 with Minimax-Meta Regularization, which is close to the MR-MAML(A). This observation suggests that the minimax-regularization could help the meta-learning model be more resistant to the memorization problem to some extent. Moreover, by comparing the results of MR-MAML methods with Minimax-Meta Regularization and the original MR-MAML models, we could find that both MAML(A) and MAML(W) gained performance improvements with added Minimax-Meta Regularization on both 5-shot and 10-shot tasks. And the smaller std values indicate the promotion of stability. This shows that minimax could be compatible with methods specifically designed for addressing memorization problem and further improve the performance.

5.3 Robust Reweighting with Meta-Learning

5.3.1 Experimental Setup

To verify the general effectiveness of our proposed methods, we further conduct the experiments on the task of robust reweighting with meta-learning. For this experiment, we compare the performance of our method and baselines on the noisy MNIST dataset, which is created by randomly flipping the labels of 40% training images. Each image has a dimension of 28×28. The task is to classify each image into 0 to 9 handwritten numbers, where the 10000 training images have 40% noisy labeled data. The validation set consists of 100 correctly-labeled images that are randomly selected from the correctly-labeled samples in the training set to ensure that the reweight method does not have the privilege of training on more data. We use the LeNet-5 as the backbone model and train the model for 1000 epochs. The learning rates for the first 1/3, the middle 1/3, and the last 1/3 training epochs are set to be 1e-2, 1e-3, and 1e-4 respectively. The basic meta-learning baselines we evaluate here is Meta-Reweighting introduced by the work (Ren et al., 2018). The Meta-Reweighitng algorithm learns to assign weights to training examples for robust learning. To determine the example weights, Meta-Reweighting performs a meta gradient descent step on the mini-batch example weights (which are initialized from zero) to minimize the loss on a clean unbiased validation set. Our method adds the Minimax-Meta Regularization on top of Meta-Reweighting. We add regularization on the outer-loop, where the optimal weights are calculated and adopted for meta-update. The inverted regularization is added on the inner-loop, where the weighted inner-model fits the clean unbiased validation set for optimal weight calculation. Intuitively, such a regularization method makes the model becomes more conservative when updating based on noise train data in the outer loop and values the diversity of predictions more, thereby resisting overfit. At the same time, the inner model was encouraged to make sharper predictions on the clean validation set by the inverted regularization, so that the potential of the clean data set can be more fully utilized. The regularization objective used in our method is maximizing output entropy (minimizing output entropy in the outer-loop). We call our method Minimax Reweighting. Detailed information of the implementation of Minimax Reweighting is provided in Appendix C.

Table 5: MNIST noisy label experiment quantitative results. We report the test and train accuracy in the last epoch with 95% confidence interval for the mean over 5 runs.

	Train Accuracy	Test Accuracy
Direct Training	99.99%±0.01%	62.3±0.20%
Meta-Reweighting	70.38±0.34%	87.45±0.34%
Meta-Reweighting & Outer-loop Regularization	**59.870±0.22%**	93.90±0.25%
Minimax Reweighting	**59.33±0.25%**	**95.38±0.12%**

5.3.2 Results and Analysis

Under this setting, models experienced large epoch training with the big initial learning rate. Models are extremely prone to overfit the training dataset during the training phase. To understand the performance of the models under a robust learn setting, we could first look at the training curve of the models (Figure 2,3). Since the training set is noisy, models overfitted to the train set would show significant performance deduction on the clean test set. From the perspective of robust learning, the direct training model sets the lower performance bound to some extent. Since it does not have any denoising ability, it quickly overfits the training set during the training. It reaches peak accuracy on the clean test set around the 80th epoch and the overfitting begins after that epoch. We could identify the overfitting characteristic from the training and testing accuracy curve. Since 40% of the labels in the training set are wrong, once the model starts to predict the training data with accuracy larger than 60%, it's fitting the distribution of the noise training data instead of the ground truth distribution. At the same time, the performance deduction on the clean test set would begin. Finally, we could observe the training accuracy and testing accuracy of directly trained model converged to nearly 100% and 60% respectively, which indicates a complete overfit. On the contrary, the model with optimal learning robustness should never overfit the train set, which would maintain a train accuracy value close to 60%(since only 60% of the train labels are correct) and keep optimal performance on the clean test set. Compared to direct training, the training curve of Meta-Reweighting baseline (Ren et al., 2018) shows a significant improvement in the learning robustness. However, it still suffers from overfitting. It neither completely overfits the training dataset nor ignores all the noises, its training accuracy converges to around 70%. Meta-Reweighting model could finally maintain test accuracy at around 87.5%, experienced continual test accuracy deduction after around 100th epoch. Minimax-Reweighting nearly reached the optimal learning robustness under this setting. The training accuracy of Minimax-Reweighting stuck at around 60% with rarely any change throughout the training phase. And the testing accuracy maintained peak value around 95.5% without observable deduction. To further evaluate the effectiveness of Minimax-Reweighting, we implemented the outer-loop-only regularization on top of the Meta-Reweighting algorithm to make comparisons. The results intend that only regularizing the outer loop at the meta-level cannot reach the performance of Minimax-Meta Regularization. Quantitative results of final accuracy are shown in Table 5. As for train accuracy, the original Meta-Reweighting algorithm reached 70.38% accuracy, which indicates a certain overfit. On the contrary, after adding regularization, both Minimax Reweighting and outer-loop regularized Meta-Reweighting could preserve a training accuracy of around 60%, which represents the resistance to training set overfit. However, Minimax Reweighting outperforms outer-loop regularized Meta Reweighting on the clean test set accuracy

6 Conclusion

This paper studies the generalization problem of meta-learning. In this paper, we go one step deeper and propose a new regularization mechanism for meta-learning – Minimax-Meta Regularization. Specifically, we maximize the regularizer in the inner loop to encourage the adapted-model to fit an "aggressive, more specific, prone to overfitting" hypothesis and minimize the regularizer in the outer loop to fit a "conservative, more general, resistant to overfitting" hypothesis. Such adversarial regularization forces the meta-model to maintain generality at the meta-level even when it is easy to learn specific assumptions at the task-specific level, thereby improving the robustness of the meta-model. In the experiment, representative meta-learning scenarios, including few-shot learning, robust learning, and reinforcement learning, are conducted to verify our method. The results show that our method consistently improves the meta-learning algorithms' performance and demonstrates the advantage of Minimax-Meta Regularization.

REFERENCES

Alessandro Achille and Stefano Soatto. Information dropout: Learning optimal representations through noisy computation. *IEEE transactions on pattern analysis and machine intelligence*, 40 (12):2897–2905, 2018.

Alexander A Alemi, Ian Fischer, Joshua V Dillon, and Kevin Murphy. Deep variational information bottleneck. *arXiv preprint arXiv:1612.00410*, 2016.

Antreas Antoniou, Harrison Edwards, and Amos Storkey. How to train your maml. *arXiv preprint arXiv:1810.09502*, 2018.

Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.

Liam Collins, Aryan Mokhtari, and Sanjay Shakkottai. Task-robust model-agnostic meta-learning. *arXiv preprint arXiv:2002.04766*, 2020.

Zi-Yi Dou, Keyi Yu, and Antonios Anastasopoulos. Investigating meta-learning algorithms for low-resource natural language understanding tasks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 1192–1197, 2019.

Yan Duan, John Schulman, Xi Chen, Peter L Bartlett, Ilya Sutskever, and Pieter Abbeel. Rl^2: Fast reinforcement learning via slow reinforcement learning. *arXiv preprint arXiv:1611.02779*, 2016.

Harrison Edwards and Amos Storkey. Towards a neural statistician. *arXiv preprint arXiv:1606.02185*, 2016.

Chelsea Finn and Sergey Levine. Meta-learning and universality: Deep representations and gradient descent can approximate any learning algorithm. In *International Conference on Learning Representations*, 2018.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *International Conference on Machine Learning*, pp. 1126–1135. PMLR, 2017.

Chelsea Finn, Kelvin Xu, and Sergey Levine. Probabilistic model-agnostic meta-learning. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, pp. 9537–9548, 2018.

Sebastian Flennerhag, Andrei Rusu, Razvan Pascanu, Francisco Visin, Hujun Yin, and Raia Hadsell. Meta-learning with warped gradient descent. In *International Conference on Learning Representations 2020*, 2020.

Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pp. 1050–1059. PMLR, 2016.

Erin Grant, Chelsea Finn, Sergey Levine, Trevor Darrell, and Thomas Griffiths. Recasting gradient-based meta-learning as hierarchical bayes. In *International Conference on Learning Representations*, 2018.

Muhammad Abdullah Jamal and Guo-Jun Qi. Task agnostic meta-learning for few-shot learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11719–11727, 2019.

Łukasz Kaiser, Ofir Nachum, Aurko Roy, and Samy Bengio. Learning to remember rare events. *arXiv preprint arXiv:1703.03129*, 2017.

Bingyi Kang, Zhuang Liu, Xin Wang, Fisher Yu, Jiashi Feng, and Trevor Darrell. Few-shot object detection via feature reweighting. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8420–8429, 2019.

Gregory Koch, Richard Zemel, Ruslan Salakhutdinov, et al. Siamese neural networks for one-shot image recognition. In *ICML deep learning workshop*, volume 2. Lille, 2015.

Anders Krogh and John A Hertz. A simple weight decay can improve generalization. In *Advances in neural information processing systems*, pp. 950–957, 1992.

Brenden M Lake, Ruslan Salakhutdinov, and Joshua B Tenenbaum. Human-level concept learning through probabilistic program induction. *Science*, 350(6266):1332–1338, 2015.

Yoonho Lee and Seungjin Choi. Gradient-based meta-learning with learned layerwise metric and subspace. In *International Conference on Machine Learning*, pp. 2927–2936. PMLR, 2018.

Junnan Li, Yongkang Wong, Qi Zhao, and Mohan S Kankanhalli. Learning to learn from noisy labeled data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5051–5059, 2019.

Zhenguo Li, Fengwei Zhou, Fei Chen, and Hang Li. Meta-sgd: Learning to learn quickly for few-shot learning. *arXiv preprint arXiv:1707.09835*, 2017.

Andrea Madotto, Zhaojiang Lin, Chien-Sheng Wu, and Pascale Fung. Personalizing dialogue agents via meta-learning. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 5454–5459, 2019.

Nikhil Mishra, Mostafa Rohaninejad, Xi Chen, and Pieter Abbeel. A simple neural attentive meta-learner. In *International Conference on Learning Representations*, 2018.

Tsendsuren Munkhdalai and Hong Yu. Meta networks. In *International Conference on Machine Learning*, pp. 2554–2563. PMLR, 2017.

Renkun Ni, Micah Goldblum, Amr Sharaf, Kezhi Kong, and Tom Goldstein. Data augmentation for meta-learning. In *International Conference on Machine Learning*, pp. 8152–8161. PMLR, 2021.

Boris N Oreshkin, Pau Rodriguez, and Alexandre Lacoste. Tadam: task dependent adaptive metric for improved few-shot learning. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, pp. 719–729, 2018.

Eunbyung Park and Junier B Oliva. Meta-curvature. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, pp. 3314–3324, 2019.

Janarthanan Rajendran, Alex Irpan, and Eric Jang. Meta-learning requires meta-augmentation. *arXiv preprint arXiv:2007.05549*, 2020a.

Janarthanan Rajendran, Alexander Irpan, and Eric Jang. Meta-learning requires meta-augmentation. In *Advances in Neural Information Processing Systems*, pp. 5705–5715, 2020b.

Sachin Ravi and Hugo Larochelle. Optimization as a model for few-shot learning. In *ICLR*, 2017.

Mengye Ren, Wenyuan Zeng, Bin Yang, and Raquel Urtasun. Learning to reweight examples for robust deep learning. In *International Conference on Machine Learning*, pp. 4334–4343. PMLR, 2018.

Jonas Rothfuss, Vincent Fortuin, Martin Josifoski, and Andreas Krause. Pacoh: Bayes-optimal meta-learning with pac-guarantees. In *International Conference on Machine Learning*, pp. 9116–9126. PMLR, 2021.

Adam Santoro, Sergey Bartunov, Matthew Botvinick, Daan Wierstra, and Timothy Lillicrap. Meta-learning with memory-augmented neural networks. In *International conference on machine learning*, pp. 1842–1850. PMLR, 2016.

Jun Shu, Qi Xie, Lixuan Yi, Qian Zhao, Sanping Zhou, Zongben Xu, and Deyu Meng. Meta-weight-net: Learning an explicit mapping for sample weighting. *Advances in Neural Information Processing Systems*, 32:1919–1930, 2019.

Jake Snell, Kevin Swersky, and Richard Zemel. Prototypical networks for few-shot learning. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 4080–4090, 2017.

Hwanjun Song, Minseok Kim, Dongmin Park, Yooju Shin, and Jae-Gil Lee. Learning from noisy labels with deep neural networks: A survey. *arXiv preprint arXiv:2007.08199*, 2020.

Qianru Sun, Yaoyao Liu, Tat-Seng Chua, and Bernt Schiele. Meta-transfer learning for few-shot learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 403–412, 2019.

Flood Sung, Yongxin Yang, Li Zhang, Tao Xiang, Philip HS Torr, and Timothy M Hospedales. Learning to compare: Relation network for few-shot learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1199–1208, 2018.

Sebastian Thrun and Lorien Pratt. *Learning to learn*. Springer Science & Business Media, 2012.

Naftali Tishby and Noga Zaslavsky. Deep learning and the information bottleneck principle. In *2015 IEEE Information Theory Workshop (ITW)*, pp. 1–5. IEEE, 2015.

Oriol Vinyals, Charles Blundell, Timothy Lillicrap, Daan Wierstra, et al. Matching networks for one shot learning. *Advances in neural information processing systems*, 29:3630–3638, 2016.

Yaqing Wang, Quanming Yao, James T Kwok, and Lionel M Ni. Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys (CSUR)*, 53(3):1–34, 2020.

Huaxiu Yao, Long-Kai Huang, Linjun Zhang, Ying Wei, Li Tian, James Zou, Junzhou Huang, et al. Improving generalization in meta-learning via task augmentation. In *International Conference on Machine Learning*, pp. 11887–11897. PMLR, 2021.

Mingzhang Yin, George Tucker, Mingyuan Zhou, Sergey Levine, and Chelsea Finn. Meta-learning without memorization. *arXiv preprint arXiv:1912.03820*, 2019.

Jaesik Yoon, Taesup Kim, Ousmane Dia, Sungwoong Kim, Yoshua Bengio, and Sungjin Ahn. Bayesian model-agnostic meta-learning. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, pp. 7343–7353, 2018.

Tianhe Yu, Chelsea Finn, Annie Xie, Sudeep Dasari, Tianhao Zhang, Pieter Abbeel, and Sergey Levine. One-shot imitation from observing humans via domain-adaptive meta-learning. *arXiv preprint arXiv:1802.01557*, 2018.

APPENDIX

A GENERAL FORM OF MINIMAX-META REGULARIZATION IN META-LEARNING

Algorithm 1 General Form of Minimax-Meta Regularization in Meta-Learning

Require: Meta-training set $\mathscr{D}_{meta-train}$, Learner M with parameters ϕ
Require: Meta-Learner R with parameters θ.
Ensure: ϕ_T

```
randomly initialize φ
for d = 1, n do
    D_support, D_query ← random dataset from 𝒟_meta−train
    φ_0 ← c_0
    for t=1, T do
        X_t, Y_t ← random batch from D_support
        L_t ←
        L(M(X_t; φ_{t−1}), Y_t) + InverseRegObjective(M(X_t; φ_{t−1}), Y_t, φ_{t−1})
        c_t ← R((∇_{φ_{t−1}} L_t, L_t); θ_{d−1})
        φ_t ← c_t
    end for
    X, Y ← D_query
    L_test ← L(M(X; φ_T), Y) + RegObjective(M(X; φ_T), Y, θ_{d−1})
    Update θ_d using ∇_{φ_T} L_test
end for
```

B DETAILS OF FEW-SHOT CLASSIFICATION EXPERIMENT

B.1 IMPLEMENTATION OF MINIMAX-MAML

Algorithm 2 Minimax-MAML

Require: $p(\mathcal{T})$: distribution over tasks
Require: α, β : step size hyperparameters γ_e, γ_n: reg-rate hyperparameter of Information Entropy, L2_Norm
Ensure: θ_T

```
randomly initialize θ
while not done do
    for all T_i do
        Evaluate ∇_θ L_{T_i}(f_θ) with respect to K examples
        Compute adapted parameters with gradient
        descent: θ'_i = θ − α∇_θ (L_{T_i}(f_θ) + γ_e Entropy_{T_i}(f_θ) − γ_n L2_Norm(θ))
    end for
    Update θ ← θ − β∇_θ Σ_{T_i∼p(T)} (L_{T_i}(f_{θ'_i}) − γ_e Entropy_{T_i}(f_{θ'_i}) + γ_n L2_Norm(θ_i′)
end while
```

Pseudo code is shown in Algorithm 2.

For all few-shot classification experiments, we use a γ_e = 2 and γ_n = 5e-5.

All the MAML/MAML++ experiments involving regularization share the same form of regularization objective. The regularization is achieved by combining the l2-norm regularization and output entropy regularization. The bi-level optimization objective could be written as:

$$\theta^* = \arg\min_{\theta} \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m}\sum_{z\in\mathcal{D}_i^q}\mathcal{L}(\phi_i(\theta,\mathcal{D}_i^s),z) + \delta^{out}\cdot(\gamma^n\cdot 0.5\|\phi_i(\theta,\mathcal{D}_i^s)\|^2 - \gamma^e\mathcal{H}(\phi_i(\theta,\mathcal{D}_i^s),z))), \tag{4}$$

$$s.t. \quad \phi_i(\theta,\mathcal{D}_i^s) = \theta - \mu\nabla_\theta\sum_{z\in\mathcal{D}_i^s}\mathcal{L}(\theta,z) + \delta^{in}\cdot(\gamma^n\cdot 0.5\|\theta\|^2 - \gamma^e\mathcal{H}(\theta,z))), \tag{5}$$

Where $\mathcal{H}(\theta, z)$ denotes the information entropy of prediction of z using θ as model parameter. Here δ^{in} and δ^{out} respectively determine the type of regularization for the inner-loop and outer-loop. Their values of δ^{in} and δ^{out} can be 1, 0 or -1, corresponding to normal regularization, none regularization, and inverse regularization respectively. Original MAML has δ^{in} = 0 and δ^{out} = 0. MAML becomes Minimax-MAML while δ^{in} and δ^{out} are set by -1 and 1. The selection for δ^{in} and δ^{out} values for other experiment could be found in 1. γ^n and γ^e are hyper-parameters controlling the regularization rate. We use γ^n = 0.0005 and γ^e=2 for all the experiments. All the MAML experiments take 5 inner-steps. In one experiment, the training takes 100 epochs, and each epoch consists of 500 iterations. After each epoch, the performance of the model is evaluated on the validation set. When the training is complete, a prediction of the test set is made by the ensemble of the top 5 performing models on the validation set. Each experiment is repeated 3 times. The Adam optimizer was adopted for the model training, with a learning rate of 0.001, $\beta_1 = 0.9$ and $\beta_2 = 0.99$. Task batch size for all Omniglot experiments is 16. Mini-Imagenet experiments use task batch sizes of 4 and 2 for 1-shot and 5-shot experiments respectively.

As for empirical verification experiment, on top of the original MAML, we implement different individual regularization methods and run experiments for each one separately. The regularization methods include *outer-loop-only regularization*, *inner-loop-only regularize* , *inverse inner-loop regularization*, *loss function regularization*, and *Minimax-Meta Regularization*. This stage of experiments complete the empirical verification of method discussed in 3.2.

C Implementation Detail of Minimax Meta-Reweighting

Pseudo code is shown in Algorithm 3. In our experiment, we use a γ_{in} = 0.25 and γ_{out} = 2

Algorithm 3 Weighted Minimax Meta-Reweighting.

Require: $model\ \theta_0, train\ D_f, valid\ D_g, n, m, \gamma_{in}, \gamma_{out}$
Ensure: θ_T

for $t = 0 \ldots T-1$ **do**
  $\{X_f, y_f\} \leftarrow SampleMiniBatch\left(\mathcal{D}_f, n\right)$
  $\{X_g, y_g\} \leftarrow SampleMiniBatch\left(\mathcal{D}_g, m\right)$
  $\hat{y}_f \leftarrow$ Forward $\left(X_f, \theta_t\right)$
  $\epsilon \leftarrow 0; l_f \leftarrow \sum_{i=1}^{n}\epsilon_i C\left(y_{f,i}, \hat{y}_{f,i}\right)$
  $\nabla\theta_t \leftarrow$ BackwardAD $\left(l_f, \theta_t\right)$
  $\hat{\theta}_t \leftarrow \theta_t - \alpha\nabla\theta_t$
  $\hat{y}_g \leftarrow$ Forward $\left(X_g, \hat{\theta}_t\right)$
  $l_g \leftarrow \frac{1}{m}\sum_{i=1}^{m}\left(C\left(y_{g,i}, \hat{y}_{g,i}\right) + \gamma_{in}Entropy(\hat{y}_{g,i})\right)$
  $\nabla\epsilon \leftarrow BackwardAD\left(l_g, \epsilon\right)$
  $\tilde{w} \leftarrow \max(-\nabla\epsilon, 0); w \leftarrow \frac{\tilde{w}}{\sum_j \tilde{w} + \delta\left(\sum_j \tilde{w}\right)}$
  $\hat{l}_f \leftarrow \sum_{i=1}^{n} w_i\left(C\left(y_{f,i}, \hat{y}_{f,i}\right) - \gamma_{out}Entropy(\hat{y}_{f,i})\right)$
  $\nabla\theta_t \leftarrow BackwardAD\left(\hat{l}_f, \theta_t\right)$
  $\theta_{t+1} \leftarrow$ OptimizerStep $\left(\theta_t, \nabla\theta_t\right)$
end for