

# Meta-learning for Few-shot Natural Language Processing: A Survey ([Link](#))

Unlike traditional NLP tasks, few-shot Natural Language Processing tasks consists of training the model on a very small set of labeled data. The idea of Meta-Learning is given data/experience on previous tasks, model learns new task quickly and efficiently. In order words, the goal is to train a model on a variety of tasks with rich annotations, so that it can solve a new task with very few labelled data. Model's initial parameters are trained in such a way that it has maximal performance on a new task after the parameters have been updated through zero or a couple of gradient steps. There are lot of surveys for meta-learning, but this paper focuses on applying meta learning to few-shot NLP.

The paper gives a detailed explanation of what meta-learning is, followed by the difference between multi-task, meta-learning and transfer-learning.

## Metric-based meta-learning

Metric-based method learns a distance function between data points to classify test instances by comparing them to K-labeled examples. Essentially, metric-learning is a pretrained nearest-neighbor algorithm as explained in the following proposed works- Siamese Network (Koch et al, 2015) and Matching Network (Vinyals et al, 2016), Prototypical Network (Snell et al, 2017), Relation Network (Sung et al, 2018).

## Optimization-based meta-learning

This includes parameter initialization for a neural model with few steps of gradient descent. The state-of-the-art models for this are MAML (Model Agnostic Meta-Learning), FOMAML- First Order MAML (Finn et al, 2017) and Reptile (Nichole et al, 2018).

The paper then elaborates on different types of NLP problems which can be solved using meta-learning. A class is a task and Domain is a task. To target few-shot task, problems having rich-annotation datasets can be used, but is more challenging and practically useful.

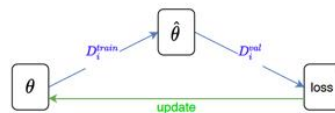


Figure 2: MAML meta-learning

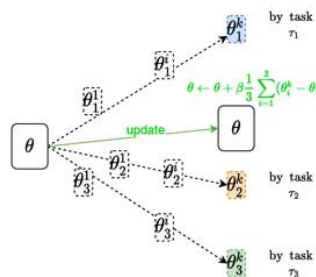


Fig. MAML and Reptile meta-learning.

## Datasets for few-shot NLP

**FewRel**- It's a relation classification dataset, having 100 relations each with 700 labeled sentences.

**SNIPS**- Intent classification dataset with only seven intent types.

**CLINC150**- It has 23,700 instances in which 22,500 examples covers 150 intents, and 1,200 instances are out-of-scope.

**ARSC**- is a sentiment classification dataset. It is comprised of Amazon reviews for 23 types of products.