论文汇报

《Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach》 2023 TSE(Early Access)

Ziyi Zhou, Huiqun Yu, Guisheng Fan, Zijie Huang, and Kang Yang

虞慧群
华东理工大学, 教授

HQLgen: deep learning based HQL query generation from program context

Automated Software Engineering
2022-11 | Journal article
DOI: 10.1007/s10515-022-00359-5
Part of ISSN: 0928-8910
Part of ISSN: 1573-7535
CONTRIBUTORS: Ziyi Zhou; huiqun yu; Guisheng Fan; Zijie Huang; Kang Yang; Jiayin Zhang

Show more detail

Source: Ziyi Zhou

Summarizing source code with hierarchical code representation

Information and Software Technology
2022-03 | Journal article
DOI: 10.1016/j.infsof.2021.106761
Part of ISSN: 0950-5849
CONTRIBUTORS: Ziyi Zhou; Huiqun Yu; Guisheng Fan; Zijie Huang; Xingguang Yang

Show more detail

Source: Ziyi Zhou

Adversarial training and ensemble learning for automatic code summarization

Neural Computing and Applications
2021-10-05 | Journal article
DOI: 10.1007/s00521-021-05907-w
CONTRIBUTORS: Ziyi Zhou; Huiqun Yu; Guisheng Fan

Show more detail

帮助

2023.3.9

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

Code Summarization

Retrieval-based

Learning-based

IR algorithm *template*

encoder-decoder framework

Large Corpus

Generalization

Meta-Learning

Meta-learning *is a technique that endows an agent with the ability of learning to learn. It intends to design models that can adapt to new concepts rapidly without numerous training examples*

Metric-based

Model-based

Optimization-based

Model-Agnostic Meta-Learning (MAML)

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

MLCS-Overview

Meta-Learning for Code Summarization -- MLCS

Formal Task Definition

Given a target code snippet x^i to summarize, the goal is to learn a unique summarizer for it using its similar example pairs $\mathcal{S}^i = \{\langle x^{ij}, y^{ij}\rangle | 1 \le j \le k_i\}$ as training data in a few-shot learning manner, where y^{ij} is the comment of the j-th similar code snippet x^{ij} and k_i is determined by the code-to-code retrieval process.

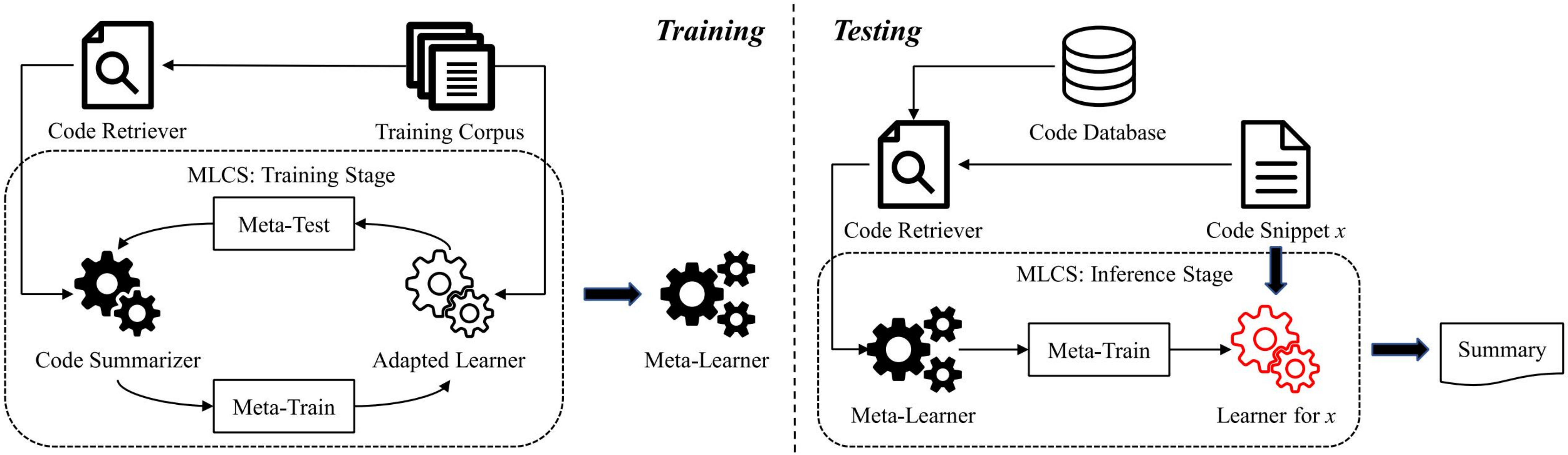


Fig. 1. Overall framework of our approach.

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

Training Stage

Optimization

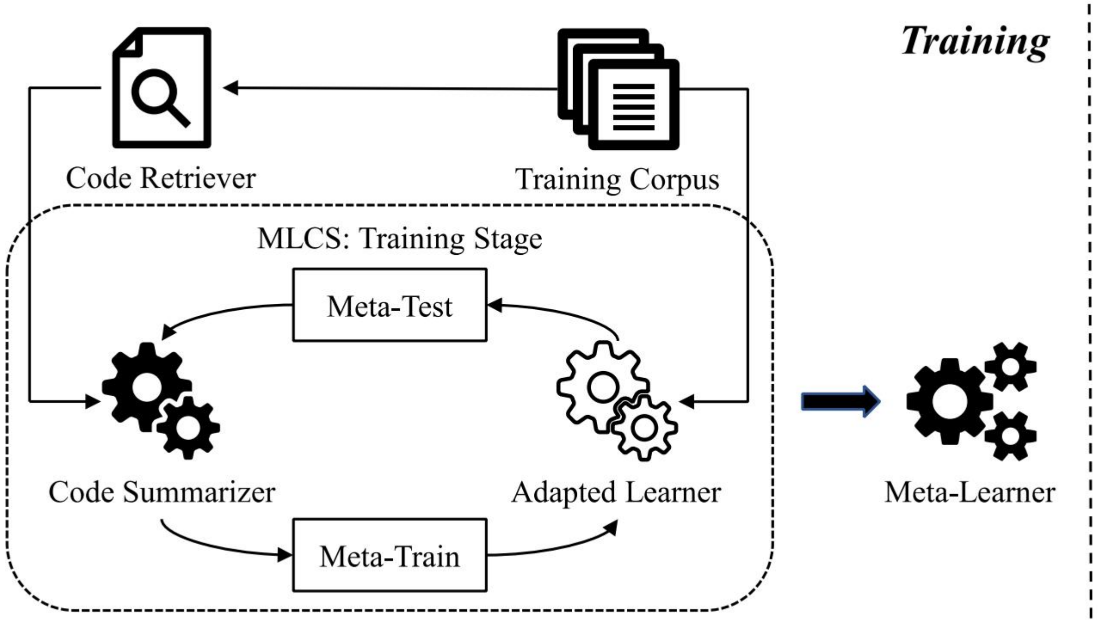


Fig. 1. Overall framework of our approach.

Algorithm 1 MLCS: Training Stage

Require: Hyper-parameters α, β, k, ε
Require: Code summarization model f_θ
Require: Training dataset $\mathcal{D}$ consisting of code-summary pairs

Randomly initialize θ
while not done **do**
    Sample batch of training pairs $\{\langle x^i, y^i\rangle\}$ from $\mathcal{D}$
    for all $\langle x^i, y^i\rangle$ **do**
        Retrieve top-k examples similar to code x^i and filter them by ε according to (12), build training set $\mathcal{S}^i$
        Evaluate $\nabla_\theta \mathcal{L}_{\mathcal{S}^i}(f_\theta)$ according to (15)
        Compute $\theta_i' = \theta - \alpha \nabla_\theta \mathcal{L}_{\mathcal{S}^i}(f_\theta)$
    end for
    Update $\theta \leftarrow \theta - \beta \nabla_\theta \sum_i \mathcal{L}_{\langle x^i, y^i\rangle}\left(f_{\theta_i'}\right)$
end while

Text Edit Distance

$$sim(x^i, x^j) = 1 - \frac{dis(x^i, x^j)}{\max(|x^i|, |x^j|)} \quad (12)$$

$$w_j = u_j \cdot (sim(x^i, x^{ij}) + 0.5) \quad (14)$$

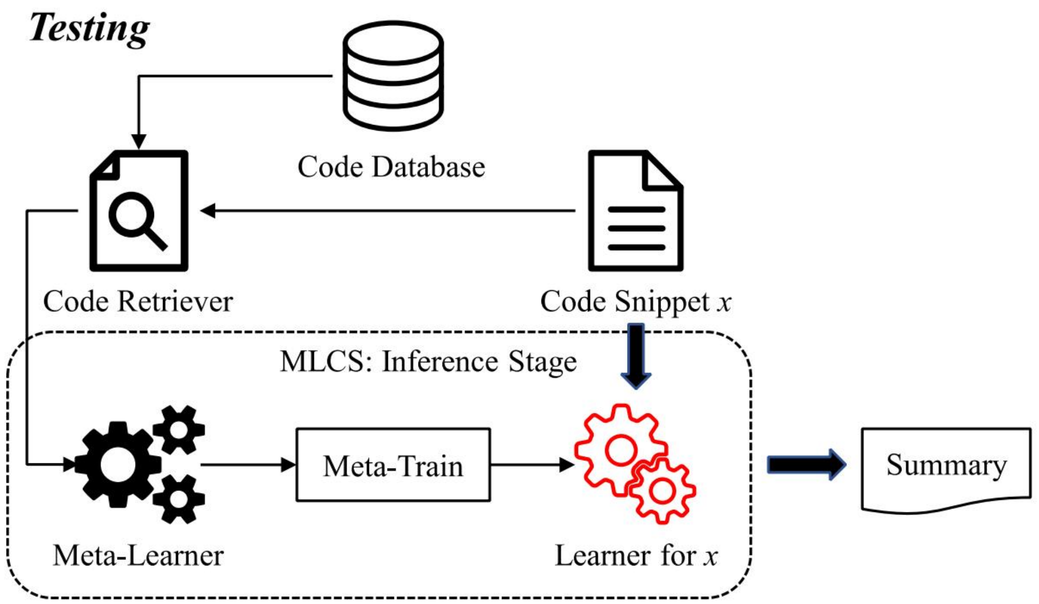
Gradient

$$\nabla_\theta \mathcal{L}_{\mathcal{S}^i}(f_\theta) = \sum_{j=1}^{k_i} w_j \nabla_\theta \mathcal{L}_{\langle x^{ij}, y^{ij}\rangle}(f_\theta) \quad (15)$$

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

Testing Stage



Algorithm 2 MLCS: Inference Stage

Require: Hyper-parameters α, n, k, ε
Require: Meta-learner f'_θ obtained via Algorithm 1
Require: Target code snippet x^i

1: $\gamma = \min\left(\frac{2}{3}\alpha, \frac{\alpha}{n-1}\right)$
2: Retrieve top-k examples similar to code x^i and filter them by ε according to (12), build training set $\mathcal{S}^i$
3: **for** *step* **in** n
4:     Evaluate $\nabla_\theta \mathcal{L}_{\mathcal{S}^i}(f'_\theta)$ using $\mathcal{S}^i$ according to (15)
5:     Compute $\theta \leftarrow \theta - \gamma \nabla_\theta \mathcal{L}_{\mathcal{S}^i}(f'_\theta)$
6: **end for**
7: **output** Target summary $\hat{y}^i = f'_\theta(x^i)$

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

Example

Example 1			
Source Code	`public void add(IMultiPoint pt){ if (inProgress) { throw new ConcurrentModificationException("Can't add point while iterator in progress"); } points.add(pt); }`		
	sim = 0.89	Signature	`public void add(TwoDNode dn)`

Example 2			
Source Code	see code below		
Retrieved	*sim* = 0.27	Signature	`def flatten_errors(cfg, res, levels=None, results=None)`
		Summary	an example function that will turn a nested dictionary of results into a flat list
Generated	Reference		filter duplicates from given errors list
	NMT		remove all duplicates from a list of errors
	Rencos		remove all duplicates from a list of errors
	Re^{2}Com		an example function that will turn a nested dictionary of exceptions
	NMT + MLCS		remove duplicates from the given errors

Source Code (Example 2):

```
def remove_duplicates(errors):
    passed = defaultdict(list)
    for error in errors:
        key = (error.linter, error.number)
        if (key in DUPLICATES):
            if (key in passed[error.lnum]):
                continue
            passed[error.lnum] = DUPLICATES[key]
        (yield error)
```

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

Dataset

Dataset	<method, summary>
JCSD(Java code summarization dataset)	87136
PCSD(Python code summarization dataset)	108726

Evaluation Metrics

- **BLEU-N**
- **Rouge-L**
- **Meteor**

$$BLEU = BP \cdot \exp\left(\sum_{n=1}^{N} w_n \, log \, p_n\right)$$

$$BP = \begin{cases} 1 & c > r \\ \exp\left(1 - \frac{r}{c}\right) & c \leq r \end{cases} \tag{17}$$

$$P_{lcs} = \frac{LCS(X,Y)}{c}$$

$$R_{lcs} = \frac{LCS(X,Y)}{r}$$

$$F_{lcs} = \frac{(1+\beta^2)P_{lcs}R_{lcs}}{R_{lcs} + \beta^2 P_{lcs}} \tag{18}$$

$$METEOR = (1 - Pen) \cdot F_{mean}$$

$$F_{mean} = \frac{PR}{\alpha P + (1-\alpha)R}$$

$$Pen = \gamma \cdot \left(\frac{ch}{m}\right)^{\beta} \tag{19}$$

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

RQ1: *How does the proposed approach perform generally compared to the baselines?*

RQ2: *To what extent could MLCS improve existing neural code summarizers?*

RQ3: *What is the effect of using different numbers of retrieved examples, and how could the model benefit from multiple similar examples via MLCS?*

RQ4: *How well could MLCS adapt to different code retrieval methods?*

RQ5: *What is the quality of the summaries produced by our approach in practice?*

RQ6: *What is the impact of cross-project data splitting and near-duplicate removal on retrieval-based code summarizers?*

RQ7: *How does MLCS perform against existing code category-based method, and how does it perform on different code categories?*

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

RQ1: *Overall Performance*

TABLE 2. Overall Performance of Proposed Approach Compared to Baselines.

Dataset	Model	BLEU-1	BLEU-2	BLEU-3	BLEU-4	BLEU-4(s)	ROUGE-L	METEOR
JCSD	DeepCom	26.03	16.53	11.57	8.91	14.95	32.71	15.95
	ast-attendgru	32.90	22.05	15.26	11.32	18.61	42.04	20.66
	NMT	32.66	21.87	15.54	12.02	18.99	41.69	20.72
	NMT (Tf)	32.02	21.93	17.01	14.33	19.80	37.32	19.88
	Rencos	34.94	24.15	17.55	13.55	20.59	42.85	21.56
	HGNN	31.80	23.54	18.95	15.59	21.11	36.95	19.07
	Re²Com	36.37	26.50	20.89	17.38	23.69	44.40	22.81
	NMT + MLCS	**37.86**	**29.00**	**23.47**	**20.16**	**26.47**	**46.94**	**23.89**
PCSD	DeepCom	24.92	15.86	11.64	8.92	15.07	30.23	14.95
	ast-attendgru	31.32	20.38	14.07	10.26	18.11	38.64	18.99
	NMT	30.58	19.62	13.95	10.61	17.98	37.29	18.51
	NMT (Tf)	30.02	20.26	15.90	13.16	18.91	34.21	18.55
	Rencos	33.70	22.80	16.77	12.85	20.44	39.97	20.12
	HGNN	31.92	22.51	17.42	13.97	20.67	38.04	19.29
	Re²Com	35.17	25.25	19.78	15.97	23.09	41.57	21.37
	NMT + MLCS	**35.89**	**26.43**	**21.16**	**17.38**	**24.70**	**43.06**	**21.98**

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

RQ2: *Improvements of Applying MLCS*

TABLE 3. The Improvements of The Baselines By Utilizing MLCS.

Dataset	Model	BLEU-1	BLEU-2	BLEU-3	BLEU-4	BLEU-4(s)	ROUGE-L	METEOR
JCSD	DeepCom + MLCS	34.36 **(+32.0%)**	26.40 **(+59.7%)**	21.94 **(+89.6%)**	18.95 **(+112.7%)**	24.37 **(+63.0%)**	40.30 **(+23.2%)**	20.97 **(+31.5%)**
	ast-attendgru + MLCS	**37.88** (+15.1%)	28.60 (+29.7%)	22.74 (+49.0%)	19.17 (+69.3%)	25.76 (+38.4%)	**47.37** (+12.7%)	23.78 (+15.1%)
	NMT + MLCS	37.86 (+15.9%)	**29.00** (+32.6%)	**23.47** (+51.0%)	**20.16** (+67.7%)	**26.47** (+39.4%)	46.94 (+12.6%)	**23.89** (+15.3%)
	NMT (Tf) + MLCS	36.40 (+13.7%)	26.51 (+20.9%)	21.49 (+26.3%)	18.54 (+29.4%)	24.28 (+22.6%)	42.11 (+12.8%)	22.41 (+12.7%)
PCSD	DeepCom + MLCS	32.60 **(+30.8%)**	24.39 **(+53.8%)**	19.93 **(+71.2%)**	16.65 **(+86.7%)**	22.90 **(+52.0%)**	37.81 **(+25.1%)**	19.66 **(+31.5%)**
	ast-attendgru + MLCS	35.75 (+14.1%)	26.07 (+27.9%)	20.31 (+44.3%)	16.33 (+59.2%)	23.97 (+32.4%)	**43.65** (+13.0%)	21.79 (+14.7%)
	NMT + MLCS	**35.89** (+17.4%)	**26.43** (+34.7%)	**21.16** (+51.7%)	**17.38** (+63.8%)	**24.70** (+37.4%)	43.06 (+15.5%)	**21.98** (+18.7%)
	NMT (Tf) + MLCS	33.81 (+12.6%)	24.28 (+19.8%)	19.58 (+23.1%)	16.50 (+25.4%)	22.83 (+20.7%)	38.70 (+13.1%)	20.63 (+11.2%)

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

RQ3: *Effect of Gradient Aggregation Under Different Numbers of Similar Examples*

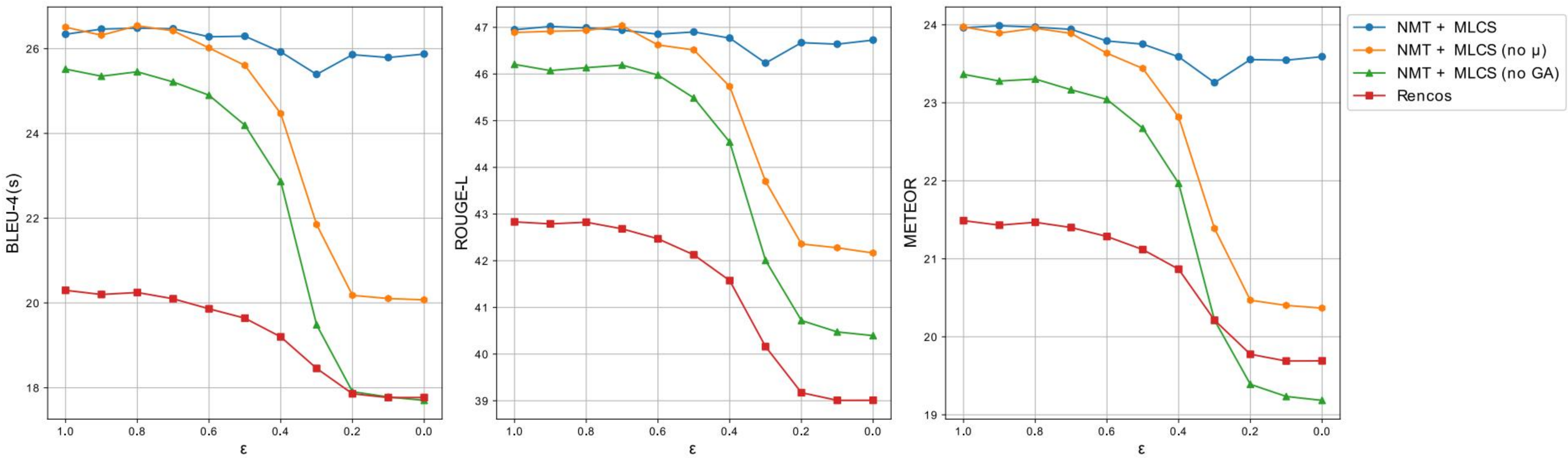


Fig. 3. Effect of changing ε on performance with k = 10.

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

RQ4: *Performance of Using Different Code Retrievers*

TABLE 4. THE PERFORMANCE OF USING DIFFERENT CODE RETRIEVERS.

Model	Retriever	BLEU-1	BLEU-2	BLEU-3	BLEU-4	BLEU-4(s)	ROUGE-L	METEOR
Retrieval Only	VSM	34.71	26.31	22.51	19.85	24.39	38.22	21.27
	LSI	33.91	25.45	21.86	19.30	23.73	37.30	20.75
	FastText	32.23	24.38	20.92	18.47	22.75	35.25	19.63
	biLSTM	35.68	27.40	23.33	**20.43**	25.10	39.24	21.87
Re^2Com	VSM	34.45	24.47	18.88	15.63	21.83	42.34	21.62
	LSI	34.51	24.48	18.86	15.61	21.87	42.43	21.61
	FastText	33.57	23.70	18.15	14.92	21.17	41.48	21.13
	biLSTM	36.37	26.50	20.89	17.38	23.69	44.40	22.81
NMT + MLCS	VSM	36.95	28.00	22.42	19.17	25.56	46.40	23.43
	LSI	36.94	27.97	22.46	19.22	25.59	46.22	23.36
	FastText	36.48	27.47	21.83	18.54	25.07	45.78	23.14
	biLSTM	**37.86**	**29.00**	**23.47**	20.16	**26.47**	**46.94**	**23.89**

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

RQ5: *Human Evaluation*

TABLE 5. THE RESULTS OF HUMAN EVALUATION.

Model	Naturalness	Informativeness
Reference	4.64	4.48
Retrieval Only	**4.53**	2.45
NMT	4.37	2.76
Rencos	4.47	2.95
Re^{2}Com	4.35	2.88
NMT + MLCS	4.48	**3.20**

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

RQ6: *Cross-project Evaluation with Near- duplicates Removal*

TABLE 6. Performance of Different Approaches with Near-duplicates Removal.

Type	Model	BLEU-1	BLEU-2	BLEU-3	BLEU-4	BLEU-4(s)	ROUGE-L	METEOR
Non-retrieval-based	DeepCom	22.51	13.38	9.35	6.89	12.88	27.72	13.56
	ast-attendgru	29.15	17.98	11.78	8.22	15.99	36.46	17.75
	NMT	28.31	17.09	11.53	8.44	15.74	34.98	17.19
	NMT (Tf)	26.87	16.83	12.60	10.12	15.78	31.07	16.68
Retrieval-based	Rencos	30.87	19.86	14.02	10.33	17.85	37.27	18.52
	HGNN	29.05	19.46	14.53	11.37	17.96	35.26	17.65
	Re^2Com	32.30	22.11	16.70	13.17	20.20	38.88	19.72
	NMT + MLCS	**32.98**	**23.16**	**17.82**	**14.22**	**21.45**	**40.05**	**20.15**

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

RQ6: *Cross-project Evaluation with Near- duplicates Removal*

TABLE 7. CROSS-PROJECT PERFORMANCE OF DIFFERENT APPROACHES WITH NEAR-DUPLICATES REMOVAL.

Type	Model	BLEU-1	BLEU-2	BLEU-3	BLEU-4	BLEU-4(s)	ROUGE-L	METEOR
Non-retrieval-based	DeepCom	10.75	3.07	1.13	0.48	4.35	14.56	6.80
	ast-attendgru	18.06	6.59	2.71	1.32	7.32	24.23	11.60
	NMT	17.79	6.24	2.51	1.16	7.21	23.96	11.44
	NMT (Tf)	12.94	3.91	1.95	1.33	5.32	16.22	8.41
Retrieval-based	Retrieval Only	10.44	3.82	2.72	2.34	5.19	12.89	6.93
	Rencos	14.75	5.03	2.28	1.29	6.03	18.99	9.24
	HGNN	13.06	4.41	2.22	1.43	5.97	18.41	8.56
	Re^{2}Com	16.80	6.23	2.92	1.80	7.26	22.93	10.92
	ast-attendgru + MLCS	**18.14**	**7.42**	**3.80**	**2.38**	**8.45**	**25.33**	**11.86**

Towards Retrieval-Based Neural Code Summarization- A Meta-Learning Approach

2023 TSE

RQ7: *Comparison with Code Category-based Method*

TABLE 8. Smoothed BLEU-4 Scores of Different Approaches in Each Comment Category.

Model	What	Why	How-to-use	How-it-is-done	Property	Others	All
DeepCom	13.53	7.32	8.57	8.23	17.66	5.60	9.62
ast-attendgru	20.24	8.24	12.60	11.48	19.68	7.02	13.78
NMT	20.38	8.40	12.23	10.86	20.48	6.60	13.56
NMT(Tf)	16.91	6.46	9.44	9.25	20.14	3.34	10.88
Rencos	16.31	9.80	11.10	10.42	17.35	7.53	12.11
HGNN	13.78	7.56	8.86	7.77	20.97	5.90	9.77
Re^2Com	19.75	10.89	12.91	12.14	22.96	**8.92**	14.25
ast-attendgru + MLCS	**21.56**	**11.97**	**13.76**	**13.13**	**25.09**	8.81	**15.37**

谢谢聆听

演讲完毕