Concept-1K: A Novel Benchmark for Instance Incremental Learning

Anonymous ACL submission

Abstract

Incremental learning (IL) is essential to realize the human-level intelligence in the neural network. However, existing IL scenarios and datasets are unqualified for assessing forgetting in PLMs, giving an illusion that PLMs do not suffer from catastrophic forgetting. To this end, we propose a challenging IL scenario called instance-incremental learning (IIL) and a novel dataset called Concept-1K, which supports an order of magnitude larger IL steps. Based on the experiments on Concept-1K, we reveal that billion-parameter PLMs still suffer from catastrophic forgetting, and the forgetting is affected by both model scale, pretraining, and buffer size. Furthermore, existing IL methods and a popular finetuning technique, LoRA, fail to achieve satisfactory performance. Our study provides a novel scenario for future studies to explore the catastrophic forgetting of PLMs and encourage more powerful techniques to be designed for alleviating the forgetting in PLMs. The data, code and scripts are in the supplementary material and will be publicly available [1].

1 Introduction

Incremental Learning (IL) aims to equip models with new knowledge without forgetting the old knowledge. It is of critical importance for research on the verge of human-level intelligence. Recent years have witnessed significant breakthroughs of Pretrained Language Models (OpenAI, 2023) (PLMs) in not only downstream tasks in natural language processing but also in real-world applications. Although there is a wealth of research on incremental learning with PLMs, we still lack a fundamental understanding of the incremental learning ability of PLMs. For instance, we can not clearly answer the questions "Do PLMs still suffer from catastrophic forgetting?", "Do larger PLMs forget less than the smaller ones?", "What is the role of pretraining in incremental learning?". The root cause is the lack of suitable IL benchmarks designed for PLMs.

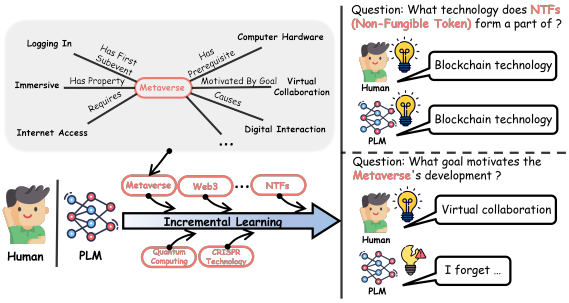


Figure 1: The illustration of Instance Incremental Learning (IIL).

There are primarily three IL scenarios for PLMs: class-incremental learning (**CIL**), task-incremental learning (**TIL**), and continual pretraining (**CPT**). However, none of them are appropriate to evaluate the IL ability of PLMs. **CIL** is designed for classification tasks such as text classification, and its goal is to learn new classes incrementally. On the one hand, state-of-the-art (SOTA) PLMs with billion-level parameters are overqualified for the above classification tasks with only dozens of categories. On the other hand, the pretraining corpus is likely to contain the knowledge required for the classification tasks (data leakage issue). As shown empirically by (Zheng et al., 2023), sequential training frozen PLMs with expanding classifiers yields comparable or even superior performance with SOTA IL methods. **TIL** aims to learn new tasks incrementally (Sun et al., 2020; Qin and Joty, 2022). Apart from the data leakage issue, the diversity of tasks and orders across research makes it difficult to readily and fairly compare IL algorithms. The last scenario **CPT** aims at continual pretraining models on the corpus from different domains. However, the evaluation relies on the performance of down-

[1] Anonymous URL: https://anonymous.4open.science/r/code-for-concept-1k-B366

stream tasks, where we can hardly identify what knowledge is learned or forgotten.

In this paper, we propose a novel incremental learning scenario called Instance Incremental Learning (**IIL**), for benchmarking the IL ability of SOTA PLMs. Unlike the IL scenario mentioned above, IIL regards each concept as an instance and is more practical and challenging for existing PLMs. Specifically, we are motivated by the human learning process and expect PLMs to learn new concepts incrementally without forgetting. For example, in Figure 1, humans can learn new concepts that are constantly emerging, such as "Metaverse" and "Quantum Computing". After learning more concepts such as "Web3.0" and "Non-Fungible Token", humans will not immediately forget the previously learned concepts such as "Metaverse". In contrast, we will later empirically show that PLMs with billions of parameters fail to recall the knowledge of old concepts.

Additionally, we present a dataset called Concept-1K for the IIL scenario. Concretely, we utilize the vast knowledge of GPT4 (OpenAI, 2023) to generate recently emerged concepts from six domains, including economy, culture, science and technology, environment, education, and health and medical, resulting in a total of 1023 concepts. Next, we prompted GPT4 to construct triplets for each concept with the relation in ConceptNet (Speer et al., 2017) where each triplet instantiates a knowledge point. Finally, we obtain 16653 training-test QA pairs corresponding to 16653 knowledge points from 1023 novel concepts.

We conduct experiments on different scales of PLMs with various pretraining steps, buffer size, and finetuning methods under the IIL scenario. The results show that PLMs suffer from catastrophic forgetting seriously and have a huge gap to human beings. In addition, we reveal that existing rehearsal-free IL methods fail to achieve superior performance. Furthermore, we find that in-context learning fails to incrementally adapt PLMs to novel concepts. In summary, the contribution is two-fold:

- We propose a novel incremental learning scenario called IIL, and a dataset called Concept-1K for better evaluating the IL ability of PLMs.

- Based the experiments on Concept-1K, we reveal that PLMs still suffer from catastrophic forgetting and existing techniques fail to achieve superior performance when no old data is stored.

Table 1: Comparison between Concept-1K and the widely-used datasets for incremental learning with PLMs. Concept-1k supports **an order of magnitude larger** incremental learning steps than existing ones.

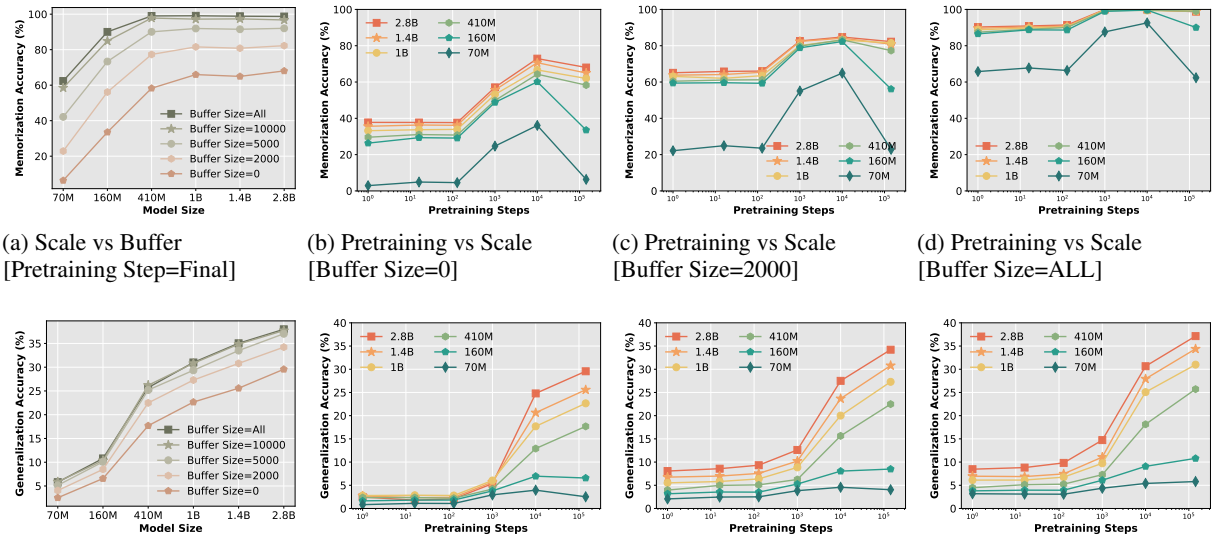
Dataset	# Classes / Concepts	Task Type
AGNews (Zhang et al., 2015)	4	Topic Classification
DBPedia (Zhang et al., 2015)	14	
YaHoo (Zhang et al., 2015)	10	
Clinic150 (Larson et al., 2019)	150	Intent Classification
Banking77 (Casanueva et al., 2020)	77	
FewRel (Han et al., 2018)	80	Relation Extraction
TACRED (Zhang et al., 2017)	40	
Few-NERD (Ding et al., 2021)	66	Named Entity Recognition
Ontonotes5 (Hovy et al., 2006)	18	
I2B2 (Murphy et al., 2010)	16	
Concept-1K	1023	Question and Answering

2 Instance Incremental Learning

Problem Formulation. Formally, IIL aims to learn a model $f_\theta : \mathbf{x} \rightarrow \mathbf{y}$ from the sequence of concepts $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_n, \cdots, \mathcal{C}_N\}$, where N is the number of concepts, and both input $\mathbf{x}$ and output $\mathbf{y}$ are natural language. The n-th concept $\mathcal{C}_n$ contains M_n training-test pairs $\mathcal{D}^{(n)} = \{\mathbf{x}_i^{(n),train}, \mathbf{x}_i^{(n),test}, \mathbf{y}_i^{(n)}\}_{i=1}^{M_n}$, where $\mathbf{x}_i^{(n),train}$ and $\mathbf{x}_i^{(n),test}$ are the training and test inputs, and $\mathbf{y}_i^{(n)}$ is the target output of them. Each training-test pair corresponds to the same knowledge point about the concept $\mathcal{C}_n$. For practical training and evaluation, we evenly divide N concepts into T ($T \leq N$) groups corresponding to T incremental steps. The model will be evaluated after learning each group of concepts.

Metric. We adopt four evaluation metrics for IIL, including Memorization Accuracy (MA), Memorization Forgetting rate (MF), Generalization Accuracy (GA), and Generalization Forgetting rate (GF). Specifically, MA and MF measure how much knowledge of training samples is memorized and forgotten. GA and GF measure how much knowledge is generalized to the test samples and is forgotten. The formal definition is provided in Appendix A.

Concept-1K To avoid the data leakage issue, we collect novel concepts from six domains, including economy, culture, science and technology, environment, education, and health and medical. To begin with, we generate 600 concepts for each domain with GPT4. We then manually filter out outdated, vague, or imaginary concepts and select the latest, most specific, and most informative concepts, resulting in a total of 1023 concepts. The concept



(a) Scale vs Buffer [Pretraining Step=Final]

(b) Pretraining vs Scale [Buffer Size=0]

(c) Pretraining vs Scale [Buffer Size=2000]

(d) Pretraining vs Scale [Buffer Size=ALL]

(e) Scale vs Buffer [Pretraining Step=Final]

(f) Pretraining vs Scale [Buffer Size=0]

(g) Pretraining vs Scale [Buffer Size=2000]

(h) Pretraining vs Scale [Buffer Size=ALL]

Figure 2: The analysis of memorization (top row) and generalization (bottom row) accuracy on Concept-1K. The backbone model is in {Pythia-70M, 160M, 410M, 1B, 1.4B, 2.8B}. The pretraining step is in {0, 16, 128, 1000, 10000, 143000 (final version)}. Each point represents the result of sequential finetuning under the IIL scenario. The detailed results with standard deveriation are provided in Table 7, 8, 9, 10, 11, 12, 13, and 14.

list is provided in Table 17. Next, we use triplets to represent "knowledge" and prompted GPT4 to construct 20 triplets for each concept with the relation in ConceptNet (Speer et al., 2017). To avoid conflict, we filter out the triplets with the same concept and relation. In addition, we filter out the triplets with relations such as "RelatedTo" and "HasContext" to ensure specificity. Finally, we utilize GPT4 to convert each triplet into a pair of training and test instances in QA format. We provide the examples and statistics of Concept-1K in Table 1, 15, and Figure 5, 6.

3 Experimental Results

We divide 1023 concepts of Concept-1K into 10 groups for IIL. We explore the impact of the model scale, pretraining, buffer size, and finetuning methods (FULL or LoRA) on the performance under the IIL scenario. We also investigate whether in-context learning is qualified to adapt PLMs to novel concepts in Concept-1K. In addition, we analyze the IIL performance at the concept level in the Appendix C.4. The detailed experimental settings, implementation details, and introduction of baselines are provided in the Appendix B.

The Role of Model Scale and Buffer Size. The results are summarized in Figure 2 and Table 2. The findings are summarized as follows: (1) Figure 2 (a) and (e) show that a larger buffer size or a larger PLM improve the accuracy of memorization and generalization. (2) The model scale determines the upper limit of the generalization performance. (3) However, the memorization accuracy of the 2.8B model is still unsatisfactory when no buffer is used. It indicates that the billion-parameter PLMs still suffer from catastrophic forgetting, and data replay is an effective technique for alleviating it. (4) Table 2 shows that when PLMs become larger, the memorization forgetting decreases while the generalization forgetting increases. It indicates that larger PLMs forget less training samples while forget more generalized knowledge since more knowledge are generalized.

Table 2: The forgetting of PLMs with different scales. The pretraining step is final and buffer size is 0.

	160M	410M	1B	1.4B	2.8B
MF (↓)	76.07±1.08	65.19±0.31	55.36±0.21	56.70±0.24	51.12±0.42
GF (↓)	5.35±0.59	15.39±0.16	18.22±0.07	21.70±0.07	23.97±0.18

The Role of Pretraining. Figure 2 (b) shows that the memorization performance increases at the early stage of pretraining (step 0 - step 10000). It indicates that pretraining enhances the memorization ability of PLMs to novel concepts. However, when performing more pretraining steps, the memorization performance will be degraded. In contrast,

Table 3: The performance of SOTA methods on Concept-1K. The detailed results are in Figure 4.

Method	MA (↑)	GA (↑)	MF (↓)	GF (↓)
SEQ	$58.28_{\pm 0.64}$	$17.68_{\pm 0.31}$	$65.19_{\pm 0.31}$	$15.39_{\pm 0.16}$
EWC (Kirkpatrick et al., 2017)	$59.83_{\pm 0.62}$	$18.09_{\pm 0.28}$	$62.46_{\pm 0.69}$	$15.07_{\pm 0.33}$
LAMOL_g (Sun et al., 2020)	$58.76_{\pm 0.53}$	$15.35_{\pm 0.46}$	$64.64_{\pm 0.59}$	$13.61_{\pm 0.47}$
LAMOL_t (Sun et al., 2020)	$58.29_{\pm 2.17}$	$15.24_{\pm 0.37}$	$66.66_{\pm 2.42}$	$14.05_{\pm 0.38}$
L2KD (Chuang et al., 2020)	$28.34_{\pm 0.29}$	$10.87_{\pm 0.01}$	$32.45_{\pm 0.63}$	$8.55_{\pm 0.42}$
PCLL (Zhao et al., 2022)	$\underline{61.94}_{\pm 1.41}$	$\mathbf{20.06}_{\pm 0.42}$	$63.04_{\pm 1.57}$	$16.27_{\pm 0.37}$
LFPT5 (Qin and Joty, 2022)	$0.63_{\pm 0.04}$	$0.84_{\pm 0.01}$	$0.04_{\pm 0.03}$	$0.03_{\pm 0.01}$
LAMOL_KD (Zheng et al., 2023)	$\mathbf{72.33}_{\pm 0.45}$	$\underline{18.20}_{\pm 0.37}$	$49.25_{\pm 0.32}$	$10.61_{\pm 0.42}$
REPLAY (buffer size=2000)	$77.31_{\pm 0.22}$	$22.48_{\pm 0.29}$	$46.97_{\pm 1.50}$	$10.57_{\pm 0.24}$
REPLAY (buffer size=Alll)	$99.01_{\pm 0.11}$	$25.70_{\pm 0.44}$	$0.70_{\pm 0.16}$	$1.44_{\pm 0.88}$

Table 4: The IIL performance of full finetuning (FULL) and LoRA on various backbones.

	MA (↑)	GA (↑)	MF (↓)	GF (↓)
Pythia-410M (FULL)	$58.28_{\pm 0.64}$	$17.68_{\pm 0.31}$	$65.19_{\pm 0.31}$	$15.39_{\pm 0.16}$
Pythia-410M (LoRA)	$15.72_{\pm 0.72}$	$6.58_{\pm 0.15}$	$30.97_{\pm 1.15}$	$2.96_{\pm 0.14}$
Pythia-2.8B (LoRA)	$36.56_{\pm 0.93}$	$10.93_{\pm 0.17}$	$83.94_{\pm 0.64}$	$6.04_{\pm 0.23}$
Vicuna 7B (LoRA)	$42.28_{\pm 0.25}$	$16.74_{\pm 0.33}$	$78.32_{\pm 0.54}$	$7.69_{\pm 0.47}$
LLaMa 7B (LoRA)	$41.76_{\pm 0.27}$	$16.20_{\pm 0.18}$	$80.55_{\pm 0.39}$	$8.08_{\pm 0.31}$
LLaMa 13B (LoRA)	$48.90_{\pm 0.68}$	$22.67_{\pm 0.30}$	$74.55_{\pm 0.46}$	$12.26_{\pm 0.18}$

Figure 2 (f) shows that the generalization performance increases monotonically for PLMs larger than 160M. The reason may be that PLMs will gradually learn to extract the knowledge behind the text during the pretraining process instead of simply remembering specific texts. In addition, especially for larger models, pretraining is more beneficial to generalization ability rather than memorization ability. Furthermore, Figure 2 (c)(d)(g)(f) indicates that a larger buffer size improves both memorization and generalization ability for all pretraining steps. The improvement of memorization ability is more significant than that of generalization ability.

Performance of SOTA Methods. We compare 8 rehearsal-free methods and data replay (REPLAY) under the IIL scenario. Table 3 indicates that although existing methods have improved sequential finetuning (SEQ), there is still a gap in the data replay with only 2000 samples (about 12 % of the total sample). The gap in memorization accuracy is much greater than generalization accuracy, indicating that PLMS still suffers from catastrophic forgetting under the challenging IIL scenario. The result urges us to design a more powerful IL algorithm to reduce the need for data replay.

LoRA v.s. Full Finetuning. LoRA (Hu et al., 2021) only finetunes a small proportion of parameters (usually less than 1%) and effectively adapt PLMs to specific domains. We compare the performance of full finetuning with LoRA in the IIL setting. Table 4 shows that LoRA limits the ability to memorize new knowledge or generalize to new instances. More specifically, the memorization and generalization accuracy of Pythia-410M is higher than LLaMa 7B (LoRA). This reminds us that when we need PLM to learn much new knowledge, we should prioritise using full fine-tuning instead of LoRA.

Table 5: The accuracy of in-context learning on the full Concept-1K (not IIL). "Rand.", "Same Conc.", and "Same Know." represent that the demonstration samples are selected randomly, or from the instances related to the same concept, or from the instance related to the same knowledge (i.e., the same training-test pair).

	1 Shot			**5 Shot**		
	Rand.	**Same Conc.**	**Same Know.**	**Rand.**	**Same Conc.**	**Same Know.**
Pythia-70M	$0.02_{\pm 0.00}$	$0.10_{\pm 0.04}$	$91.78_{\pm 0.00}$	$0.05_{\pm 0.02}$	$0.17_{\pm 0.01}$	$15.02_{\pm 0.00}$
Pythia-160M	$0.10_{\pm 0.02}$	$0.15_{\pm 0.02}$	$41.36_{\pm 0.00}$	$0.37_{\pm 0.02}$	$0.55_{\pm 0.15}$	$13.49_{\pm 0.00}$
Pythia-410M	$0.49_{\pm 0.17}$	$0.75_{\pm 0.23}$	$40.66_{\pm 0.00}$	$2.32_{\pm 0.80}$	$2.04_{\pm 0.63}$	$18.45_{\pm 0.00}$
Pythia-1B	$0.88_{\pm 0.36}$	$1.21_{\pm 0.13}$	$43.73_{\pm 0.00}$	$3.03_{\pm 0.46}$	$2.69_{\pm 0.56}$	$35.75_{\pm 0.00}$
Pythia-1.4B	$1.92_{\pm 0.28}$	$2.52_{\pm 0.16}$	$54.37_{\pm 0.00}$	$4.56_{\pm 0.58}$	$3.80_{\pm 0.01}$	$45.59_{\pm 0.00}$
Pythia-2.8B	$1.81_{\pm 0.45}$	$2.50_{\pm 0.27}$	$53.56_{\pm 0.00}$	$5.69_{\pm 0.35}$	$4.07_{\pm 0.93}$	$60.20_{\pm 0.00}$
LLaMa 7B	$4.32_{\pm 0.11}$	$4.93_{\pm 0.72}$	$83.57_{\pm 0.00}$	$8.79_{\pm 0.16}$	$6.24_{\pm 0.23}$	$66.25_{\pm 0.00}$
Vicuna 7B	$6.67_{\pm 0.74}$	$7.00_{\pm 0.88}$	$55.67_{\pm 0.00}$	$9.63_{\pm 0.78}$	$7.75_{\pm 0.15}$	$36.04_{\pm 0.00}$
GPT 3.5	6.60	8.20	51.60	8.60	13.00	74.80
GPT 4	10.20	10.40	76.60	7.40	21.80	86.20

In-Context Learning. Due to computing resources and privacy concerns, it is impractical to fully or partially fine-tune SOTA PLMs such as GPT4. We evaluate the in-context learning performance on the whole Concept-1K to investigate the ability to learn novel concepts. Table 5 indicates that the performance is unsatisfactory for all PLMs when the demonstration instances are less related to the test instance. Additionally, we can improve the performance by leveraging demonstration instances that share the same knowledge as the test instances. In other words, PLMs achieves superior performance only when the demonstration instances containing exactly the same knowledge as the test sample. Therefore, in-context learning can not meet the goal of adapting PLMs to new knowledge.

4 Conclusion

Incremental learning is a research field that has developed rapidly in recent years. However, the existing IL scenario and datasets are not suitable for SOTA PLMs and even give an illusion that PLMs do not suffer from catastrophic forgetting. This study presents a novel IL scenario called instance-incremental learning and a novel dataset called Concept-1K, supporting an order of magnitude larger steps for IL. We hope our study will further inspire follow-up works to explore the catastrophic forgetting in PLMs and to design more powerful techniques for finetuning and incremental learning.

Limitations

There are two limitations of this research: (1) The knowledge of Concept-1K is defined in the form of triplets, which can not cover the knowledge in a broad sense. (2) Apart from the experiments of in-context learning, other experiments are conducted on PLMs with less than 13B parameters. The conclusion of these experiments may not hold when finetuning SOTA PLMs such as GPT4 with more than 100B parameters.

Ethical Considerations

The ethical considerations of our research are carefully addressed to ensure compliance with relevant standards and transparency. To this end, we provide the following clarifications:

1. **Dataset Collection**: Our research employs GPT4 to construct Concept-1K and filter out offensive or harmful instances. The use of GPT4 was consistent with their intended use. The dataset Concept-1K is publicly available for academic and research purposes.

2. **Reproducibility**: We provide a detailed setting of our experiments. The source code, data, and scripts will all be publicly available. Our findings are in alignment with observed empirical outcomes.

References

Stella Biderman, Hailey Schoelkopf, Quentin Gregory Anthony, Herbie Bradley, Kyle O'Brien, Eric Hallahan, Mohammad Aflah Khan, Shivanshu Purohit, USVSN Sai Prashanth, Edward Raff, et al. 2023. Pythia: A suite for analyzing large language models across training and scaling. In *International Conference on Machine Learning*, pages 2397–2430. PMLR.

Sidney Black, Stella Biderman, Eric Hallahan, Quentin Anthony, Leo Gao, Laurence Golding, Horace He, Connor Leahy, Kyle McDonell, Jason Phang, Michael Pieler, Usvsn Sai Prashanth, Shivanshu Purohit, Laria Reynolds, Jonathan Tow, Ben Wang, and Samuel Weinbach. 2022. GPT-NeoX-20B: An open-source autoregressive language model. In *Proceedings of BigScience Episode #5 – Workshop on Challenges & Perspectives in Creating Large Language Models*, pages 95–136, virtual+Dublin. Association for Computational Linguistics.

Iñigo Casanueva, Tadas Temčinas, Daniela Gerz, Matthew Henderson, and Ivan Vulić. 2020. Efficient intent detection with dual sentence encoders. In *Proceedings of the 2nd Workshop on Natural Language Processing for Conversational AI*, pages 38–45, Online. Association for Computational Linguistics.

Yung-Sung Chuang, Shang-Yu Su, and Yun-Nung Chen. 2020. Lifelong language knowledge distillation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2914–2924, Online. Association for Computational Linguistics.

Ning Ding, Guangwei Xu, Yulin Chen, Xiaobin Wang, Xu Han, Pengjun Xie, Haitao Zheng, and Zhiyuan Liu. 2021. Few-NERD: A few-shot named entity recognition dataset. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 3198–3213, Online. Association for Computational Linguistics.

Xu Han, Hao Zhu, Pengfei Yu, Ziyun Wang, Yuan Yao, Zhiyuan Liu, and Maosong Sun. 2018. FewRel: A large-scale supervised few-shot relation classification dataset with state-of-the-art evaluation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 4803–4809, Brussels, Belgium. Association for Computational Linguistics.

Eduard Hovy, Mitchell Marcus, Martha Palmer, Lance Ramshaw, and Ralph Weischedel. 2006. OntoNotes: The 90% solution. In *Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers*, pages 57–60, New York City, USA. Association for Computational Linguistics.

Edward J Hu, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, et al. 2021. Lora: Low-rank adaptation of large language models. In *International Conference on Learning Representations*.

James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, et al. 2017. Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences*, 114(13):3521–3526.

Stefan Larson, Anish Mahendran, Joseph J. Peper, Christopher Clarke, Andrew Lee, Parker Hill, Jonathan K. Kummerfeld, Kevin Leach, Michael A. Laurenzano, Lingjia Tang, and Jason Mars. 2019. An evaluation dataset for intent classification and out-of-scope prediction. 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)*, pages 1311–1316, Hong Kong, China. Association for Computational Linguistics.

Ilya Loshchilov and Frank Hutter. 2018. Fixing weight decay regularization in adam.

Sourab Mangrulkar, Sylvain Gugger, Lysandre Debut, Younes Belkada, Sayak Paul, and Benjamin

Bossan. 2022. Peft: State-of-the-art parameter-efficient fine-tuning methods. https://github.com/huggingface/peft.

Shawn N Murphy, Griffin Weber, Michael Mendis, Vivian Gainer, Henry C Chueh, Susanne Churchill, and Isaac Kohane. 2010. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *Journal of the American Medical Informatics Association*, 17(2):124–130.

OpenAI. 2023. Gpt-4 technical report. *ArXiv*, abs/2303.08774.

Chengwei Qin and Shafiq Joty. 2022. Lfpt5: A unified framework for lifelong few-shot language learning based on prompt tuning of t5. In *International Conference on Learning Representations*.

Anastasia Razdaibiedina, Yuning Mao, Rui Hou, Madian Khabsa, Mike Lewis, and Amjad Almahairi. 2023. Progressive prompts: Continual learning for language models. In *The Eleventh International Conference on Learning Representations*.

Yijia Shao, Yiduo Guo, Dongyan Zhao, and Bing Liu. 2023. Class-incremental learning based on label generation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 1263–1276, Toronto, Canada. Association for Computational Linguistics.

Robyn Speer, Joshua Chin, and Catherine Havasi. 2017. Conceptnet 5.5: An open multilingual graph of general knowledge. In *Proceedings of the AAAI conference on artificial intelligence*, volume 31.

Fan-Keng Sun, Cheng-Hao Ho, and Hung-Yi Lee. 2020. Lamol: Language modeling for lifelong language learning. In *International Conference on Learning Representations*.

Zifeng Wang, Zizhao Zhang, Chen-Yu Lee, Han Zhang, Ruoxi Sun, Xiaoqi Ren, Guolong Su, Vincent Perot, Jennifer Dy, and Tomas Pfister. 2022. Learning to prompt for continual learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 139–149.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. 2019. Huggingface's transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*.

Xiang Zhang, Junbo Zhao, and Yann LeCun. 2015. Character-level convolutional networks for text classification. *Advances in neural information processing systems*, 28.

Yuhao Zhang, Victor Zhong, Danqi Chen, Gabor Angeli, and Christopher D. Manning. 2017. Position-aware attention and supervised data improve slot filling. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 35–45, Copenhagen, Denmark. Association for Computational Linguistics.

Yingxiu Zhao, Yinhe Zheng, Zhiliang Tian, Chang Gao, Jian Sun, and Nevin L. Zhang. 2022. Prompt conditioned VAE: Enhancing generative replay for lifelong learning in task-oriented dialogue. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 11153–11169, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Junhao Zheng, Shengjie Qiu, and Qianli Ma. 2023. Learn or recall? revisiting incremental learning with pre-trained language models. *arXiv preprint arXiv:2312.07887*.

A Evaluation Metrics

The memorization accuracy (MA) is defined as follows:

$$MA = \frac{1}{T}\sum_{t=1}^{T}\mathcal{A}_t, \tag{1}$$

where T is the number of incremental steps. $\mathcal{A}_t$ represents average accuracy on the training instances from all learned concepts. Specifically, $\mathcal{A}_t$ is computed as follows:

$$\mathcal{A}_t = \frac{1}{t}\sum_{i=1}^{t} a_{t,i}, \tag{2}$$

where $a_{t,i}$ represents the accuracy evaluated on the task i after training the model incrementally from concepts belonging to group $1, \cdots, t$. The accuracy is calculated as the exact match between the model output and the target output.

The memorization forgetting (MF) is computed as the average accuracy on all training instances of all learned concepts:

$$MF = \frac{1}{T-1}\Sigma_{i=1}^{T-1}[max_{j<T}(\{a_{j,i}\}_j) - a_{T,i}]. \tag{3}$$

Concretely, $max_{j<T}(\{a_{j,i}\}_j)$ represents the highest accuracy of task i since it has been learned. $a_{T,i}$ represents the accuracy of task i at step T. $[max_{j<T}(\{a_{j,i}\}_j) - a_{T,i}]$ computes the decrease of the accuracy of task i at incremental step T.

The generalization accuracy (GA) and generalization forgetting (GF) are computed in a similar way except that the model is evaluated on the test set instead of the training set.

B Experimental Settings

Backbones. We use the Pythia suite (Biderman et al., 2023) and the popular open-source models including LLaMa and Vicuna for experiments. Pythia

Table 6: The statistics of the PLMs used in this paper. †: Non-embedding parameters according to Biderman et al. (2023).

Model Class	Pretrained Weights	Parameters	Layers	Hidden Dim	Link
GPT-NeoX	Pythia-70m	19M	6	512	Link
	Pythia-160m	85M†	12	768	Link
	Pythia-410m	302M†	24	1024	Link
	Pythia-1b	805M†	16	2048	Link
	Pythia-1.4b	1.21B†	24	2048	Link
	Pythia-2.8b	2.52B†	32	2560	Link
LLaMa	llama-7b-hf	7B	32	4096	Link
	vicuna-7b-v1.1	7B	32	4096	Link
	llama-2-13b-hf	13B	40	5120	Link

is based on GPT-NeoX (Black et al., 2022), which contains 8 model sizes and 154 pre-training checkpoints, enabling research in interpretability and learning dynamics. The statistics of the 9 PLMs used in this paper are summarised in Table 6. We download the pre-trained weights from Huggingface (Wolf et al., 2019).

Datasets. We split the 1023 concepts in Concept-1K into 10 groups for IIL. Specifically, they sorted the concepts alphabetically and shuffled the order with random seed 1. The first group contains 105 concepts, while the others contain 102 concepts.

Baseline Methods. The introduction of the baseline methods is as follows:

- SEQ: SEQ refers to sequential fine-tuning. It can be considered as the lower bound of incremental learning.

- REPLAY: Experience replay stores representative old samples and jointly optimizes old and new samples when learning new tasks. It is a practical and popular technique in incremental learning.

- EWC (Kirkpatrick et al., 2017): EWC is a regularization-based method. The weight of each parameter is determined by the diagonal of the Fisher information matrix. The weight of the regularization loss is searched in $\{5\times10^3, 1\times10^4, 5\times10^4, 1\times10^5, 1\times10^6, 1\times10^7\}$.

- LAMOL (Sun et al., 2020): LAMOL trains PLMs with question-answering and generative objectives. It generates pseudo-samples before learning each new task for data replay. The weight of the generation loss is $\lambda = 0.25$. The proportion of pseudo samples is $\gamma = 0.20$. There are two variants of LAMOL, including LAMOL_t and LAMOL_g. The difference is whether a task-specific token is used for generation.

- L2KD (Chuang et al., 2020): L2KD adds a knowledge distillation target based on LAMOL. The teacher model is trained from scratch. We implemented the word-level variant since it performs best on text classification tasks.

- LAMOL_KD (Zheng et al., 2023): LAMOL_KD utilizes knowledge distillation based on LAMOL_t. Unlike L2KD, the teacher model in LAMOL_KD is trained on all previous tasks. The new data are used to learn the LAMOL objectives, and pseudo data are used for word-level knowledge distillation as a regularization term.

- PCLL (Zhao et al., 2022): PCLL combines the idea of variational autoencoders and word-level knowledge distillation with the objective of LAMOL.

- LFPT5 (Qin and Joty, 2022): LFPT5 only learns soft prompts for each new task. The training objective is the same as LAMOL. The number of soft prompt tokens is 10.

- LoRA (Hu et al., 2021): LoRA only trains a small proportion of parameters of PLMs. We set the rank $r = 8$ and the scaling parameter $\alpha = 16$. We use the implementation

of LoRA from the PEFT library (Mangrulkar et al., 2022).

We note that some IL methods are not compared since they are not applicable in the IIL scenario. For example, Progressive Prompt (Razdaibiedina et al., 2023) requires task ID during both training and inference stages. VAG (Shao et al., 2023) requires storing the vocabulary of class labels and do not apply to the generation tasks without class labels. Besides, prompt-based IL methods such as L2P (Wang et al., 2022) also do not apply to generation tasks.

Implementation Details. The maximum input and output lengths are set as 32 and 10, respectively. The batch size is 32. The learning rate of PLMs is 1×10^{-5}. We use AdamW (Loshchilov and Hutter, 2018) optimizer. We use A800 GPUs for our experiments. Each experiment is repeated three times, and the average and standard deviations are reported. We search for the best hyper-parameters for each baseline method.

C Additional Experimental Results and Analysis

C.1 The Role of Model Scale, Pretraining, and Buffer Size

We provide the detailed result of Figure 2 in Table 7, 8, 9, 10, 11, 12, 13, and 14.

C.2 The performance of SOTA methods

The backbone model is Pythia-410M. We provide the step-wise average accuracy on training and test set in Figure 4 (a) and (b). The results demonstrate a large gap between SOTA IL methods and data replay. We also provided a comparison between SOTA methods in terms of memorization accuracy, generalization accuracy, memorization forgetting, and generalization forgetting in Figure 4 (c), (d), (e), and (f). In addition, from Figure 4 (g), we find that the training loss of the prompt-tuning-based method LFPT5 can not decrease to a small value. It indicates that only finetuning prompts are not practical for learning new knowledge, which is also consistent with the result of LoRA.

C.3 In-Context Learning

Apart from the analysis in the main paper, Table 5 shows that the smallest PLM (Pythia-70M) achieve high accuracy under the "1 shot" and "Same Know." settings. The reason is that small PLMs simply copy the output in the demonstration instance as the final output. Under the "5 shot" and "Same Know." settings, the accuracy of Pythia-70M is only 15.02%.

We use the model "gpt-3.5-turbo" and "gpt-4-turbo" for GPT3.5 and GPT4, respectively. Due to the high cost of evaluating on the whole concept-1K, we randomly sample 500 instances for GPT3.5 and GPT4. The output and the target of GPT3.5 and GPT4 on the 500 instances are provided in the supplementary material. Table 5 shows that GPT4 achieves 86.20% under the "5 shot" and "Same Know." settings. It indicates that the training and the test instance of Concept-1K share the same knowledge.

C.4 Concept-Level Accuracy on Concept-1K

We analyze the memorization accuracy of the concepts in the first task after IIL. The backbone model is Pythia-2.8B, and it is sequential finetuning on Concept-1K. From Figure 3, we have the following observation:

- The concepts such as "Cultural Competence", "Augmented Reality Displays", "Geopolitical Risks" that achieve high accuracy are broad, interdisciplinary topics that are likely rich in diverse data sources and content. These subjects span various fields and are likely to have clear definitions, significant public interest, and extensive coverage in the pretraining datasets, facilitating high memorization accuracy.

- The lowest accuracy is observed in highly technical, cutting-edge, or niche fields such as "Spatial Audio", "Biomaterials", "Sustainable Tech", and "Space Tourism". These topics are likely challenging due to rapidly evolving information, specialized knowledge requirements, and potentially limited representation in available datasets.

In summary, the tendency from high to low accuracy reflects the transition from broad, interdisciplinary topics with abundant and diverse data to more specialized, technical, or emerging fields with potentially limited data exposure. This suggests that the PLMs' memory accuracy is influenced by the richness and diversity of information available in its training data and the clarity and stability of the concepts themselves. Additionally, despite their significance, emerging technologies and areas of

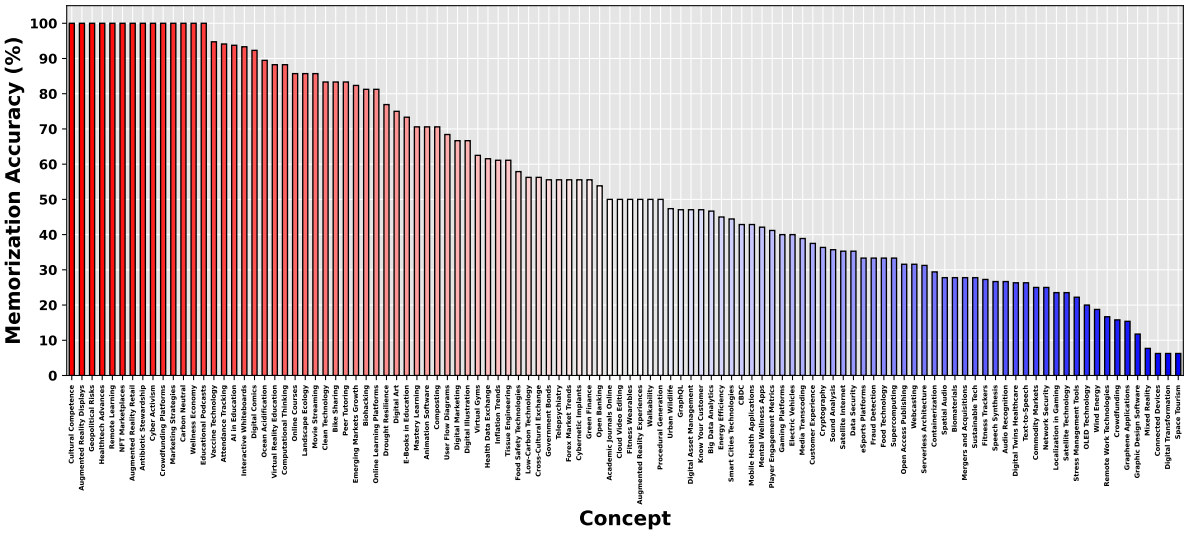


Figure 3: The accuracy of different concepts in the first task after sequential finetuning Pythia-2.8B on the 10 tasks of Concept-1K.

innovation might be harder for PLMs to learn accurately due to the rapid evolution of knowledge and possibly less structured data.

Table 7: The **memorization accuracy** when different **model scales** and **buffer size** are selected. The pretraining step is 143000 (final version). Figure 2 (a) summarises this figure's content.

	70M	160M	410M	1B	1.4B	2.8B
Buffer Size=0	$6.34_{\pm 0.86}$	$33.53_{\pm 1.39}$	$58.28_{\pm 0.61}$	$65.97_{\pm 0.31}$	$64.95_{\pm 0.55}$	$68.03_{\pm 0.46}$
Buffer Size=2000	$22.87_{\pm 1.28}$	$56.12_{\pm 0.81}$	$77.31_{\pm 0.23}$	$81.57_{\pm 0.32}$	$80.81_{\pm 0.62}$	$82.26_{\pm 0.66}$
Buffer Size=5000	$42.13_{\pm 1.88}$	$73.32_{\pm 1.00}$	$90.05_{\pm 0.64}$	$91.93_{\pm 0.49}$	$91.49_{\pm 0.32}$	$92.03_{\pm 0.35}$
Buffer Size=10000	$58.50_{\pm 0.50}$	$84.91_{\pm 1.64}$	$97.81_{\pm 0.12}$	$97.21_{\pm 0.73}$	$97.35_{\pm 0.84}$	$96.74_{\pm 0.43}$
Buffer Size=All	$62.37_{\pm 1.82}$	$90.04_{\pm 0.64}$	$99.01_{\pm 0.14}$	$99.10_{\pm 0.57}$	$98.86_{\pm 0.55}$	$98.72_{\pm 0.34}$

Table 8: The **generalization accuracy** when different **model scales** and **buffer size** are selected. The pretraining step is 143000 (final version). Figure 2 (e) summarises this figure's content.

	70M	160M	410M	1B	1.4B	2.8B
Buffer Size=0	$2.52_{\pm 0.12}$	$6.57_{\pm 0.28}$	$17.69_{\pm 0.35}$	$22.65_{\pm 0.39}$	$25.57_{\pm 0.36}$	$29.56_{\pm 0.35}$
Buffer Size=2000	$4.04_{\pm 0.07}$	$8.49_{\pm 0.23}$	$22.49_{\pm 0.32}$	$27.29_{\pm 0.55}$	$30.80_{\pm 0.28}$	$34.21_{\pm 0.40}$
Buffer Size=5000	$5.14_{\pm 0.09}$	$10.08_{\pm 0.40}$	$25.19_{\pm 0.46}$	$29.34_{\pm 0.38}$	$33.50_{\pm 0.46}$	$37.04_{\pm 0.36}$
Buffer Size=10000	$5.88_{\pm 0.10}$	$10.36_{\pm 0.52}$	$26.20_{\pm 0.35}$	$30.81_{\pm 0.24}$	$34.85_{\pm 0.57}$	$37.72_{\pm 0.52}$
Buffer Size=All	$5.80_{\pm 0.08}$	$10.80_{\pm 0.46}$	$25.70_{\pm 0.28}$	$30.99_{\pm 0.42}$	$35.03_{\pm 0.27}$	$37.97_{\pm 0.37}$

Table 9: The **memorization accuracy** when different **pretraining steps** and **model scales** are selected. The buffer size is 0. Figure 2 (b) summarizes the content of this figure.

	Step=0	Step=16	Step=128	Step=1000	Step=10000	Step=143000
Pythia-70M	$2.92_{\pm 0.06}$	$4.89_{\pm 0.16}$	$4.58_{\pm 0.10}$	$24.69_{\pm 0.04}$	$36.08_{\pm 0.63}$	$6.34_{\pm 0.86}$
Pythia-160M	$26.33_{\pm 0.31}$	$29.34_{\pm 0.25}$	$29.11_{\pm 0.11}$	$48.78_{\pm 0.09}$	$60.12_{\pm 0.34}$	$33.53_{\pm 1.39}$
Pythia-410M	$29.58_{\pm 1.10}$	$31.06_{\pm 0.24}$	$30.78_{\pm 0.22}$	$49.84_{\pm 0.06}$	$64.31_{\pm 0.38}$	$58.28_{\pm 0.61}$
Pythia-1B	$33.18_{\pm 0.59}$	$33.71_{\pm 0.17}$	$33.91_{\pm 0.49}$	$53.14_{\pm 0.59}$	$66.66_{\pm 0.59}$	$65.97_{\pm 0.31}$
Pythia-1.4B	$35.61_{\pm 0.59}$	$36.31_{\pm 0.51}$	$36.15_{\pm 0.32}$	$55.61_{\pm 0.59}$	$70.66_{\pm 0.51}$	$64.95_{\pm 0.55}$
Pythia-2.8B	$37.76_{\pm 0.29}$	$37.73_{\pm 0.38}$	$37.65_{\pm 0.39}$	$57.09_{\pm 0.34}$	$72.92_{\pm 0.59}$	$68.03_{\pm 0.46}$

Table 10: The **generalization accuracy** when different **pretraining steps** and **model scales** are selected. The buffer size is 0. Figure 2 (f) summarises this figure's content.

	Step=0	Step=16	Step=128	Step=1000	Step=10000	Step=143000
Pythia-70M	$0.83_{\pm 0.02}$	$1.10_{\pm 0.01}$	$1.08_{\pm 0.04}$	$2.94_{\pm 0.04}$	$3.95_{\pm 0.03}$	$2.52_{\pm 0.12}$
Pythia-160M	$1.64_{\pm 0.04}$	$1.81_{\pm 0.02}$	$1.82_{\pm 0.06}$	$3.79_{\pm 0.05}$	$6.96_{\pm 0.04}$	$6.57_{\pm 0.28}$
Pythia-410M	$2.34_{\pm 0.07}$	$2.35_{\pm 0.04}$	$2.21_{\pm 0.04}$	$4.18_{\pm 0.10}$	$12.91_{\pm 0.14}$	$17.69_{\pm 0.35}$
Pythia-1B	$2.78_{\pm 0.14}$	$2.85_{\pm 0.24}$	$2.80_{\pm 0.24}$	$6.03_{\pm 0.24}$	$17.70_{\pm 0.24}$	$22.65_{\pm 0.39}$
Pythia-1.4B	$2.71_{\pm 0.37}$	$2.32_{\pm 0.24}$	$2.40_{\pm 0.53}$	$5.68_{\pm 0.24}$	$20.64_{\pm 0.24}$	$25.57_{\pm 0.36}$
Pythia-2.8B	$2.48_{\pm 0.23}$	$1.79_{\pm 0.29}$	$2.09_{\pm 0.32}$	$5.35_{\pm 0.49}$	$24.79_{\pm 0.24}$	$29.56_{\pm 0.35}$

Table 11: The **memorization accuracy** when different **pretraining steps** and **model scales** are selected. The buffer size is 2000. Figure 2 (c) summarises this figure's content.

	Step=0	**Step=16**	**Step=128**	**Step=1000**	**Step=10000**	**Step=143000**
Pythia-70M	22.15±0.13	24.91±0.20	23.54±0.15	55.06±0.16	64.89±0.23	22.87±1.28
Pythia-160M	59.42±0.13	59.62±0.77	59.30±0.18	78.84±0.44	82.30±0.18	56.12±0.81
Pythia-410M	60.43±0.59	61.09±0.66	61.24±0.27	79.77±0.15	83.24±0.12	77.31±0.23
Pythia-1B	63.11±0.15	62.05±0.40	63.51±0.31	80.11±0.50	83.11±0.38	81.57±0.32
Pythia-1.4B	63.77±0.15	64.14±0.23	65.40±0.32	82.38±0.34	84.19±0.22	80.81±0.62
Pythia-2.8B	65.10±0.45	65.84±0.57	66.04±0.53	82.63±0.18	84.80±0.44	82.26±0.66

Table 12: The **generalization accuracy** when different **pretraining steps** and **model scales** are selected. The buffer size is 2000. Figure 2 (g) summarises this figure's content.

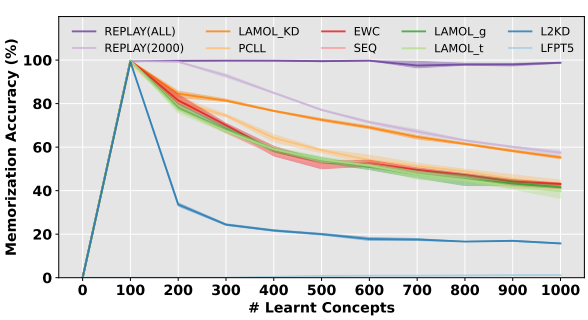
	Step=0	**Step=16**	**Step=128**	**Step=1000**	**Step=10000**	**Step=143000**
Pythia-70M	2.07±0.08	2.46±0.06	2.51±0.13	3.84±0.06	4.57±0.07	4.04±0.07
Pythia-160M	3.17±0.08	3.55±0.14	3.52±0.06	5.26±0.14	8.05±0.08	8.49±0.23
Pythia-410M	3.96±0.06	4.97±0.09	5.09±0.04	6.25±0.13	15.64±0.06	22.49±0.32
Pythia-1B	5.51±0.17	5.81±0.43	6.37±0.56	8.85±0.55	20.01±0.51	27.29±0.55
Pythia-1.4B	6.76±0.21	6.98±0.16	7.55±0.55	10.21±0.42	23.67±0.18	30.80±0.28
Pythia-2.8B	8.07±0.28	8.57±0.54	9.32±0.40	12.61±0.25	27.50±0.23	34.21±0.40

Table 13: The **memorization accuracy** when different **pretraining steps** and **model scales** are selected. The buffer size is unlimited (all old samples are stored). Figure 2 (d) summarizes the content of this figure.

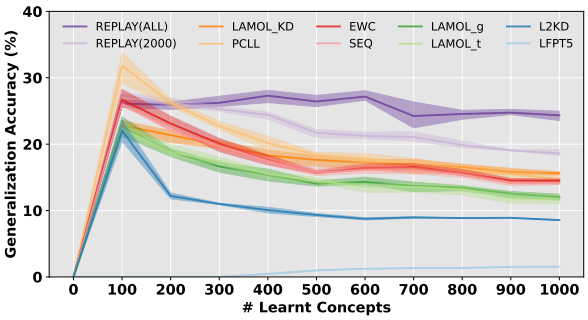
	Step=0	**Step=16**	**Step=128**	**Step=1000**	**Step=10000**	**Step=143000**
Pythia-70M	65.77±0.56	67.75±0.19	66.36±0.09	87.61±0.14	92.59±0.11	62.37±1.82
Pythia-160M	86.54±0.15	88.75±0.28	88.64±0.07	98.95±0.03	99.62±0.05	90.04±0.64
Pythia-410M	87.46±0.19	89.07±0.27	90.03±0.16	99.30±0.22	99.59±0.02	99.01±0.14
Pythia-1B	88.77±0.02	89.75±0.19	90.36±0.09	99.75±0.19	99.36±0.09	99.10±0.57
Pythia-1.4B	89.65±0.02	90.34±0.03	90.76±0.54	99.85±0.21	99.76±0.02	98.86±0.55
Pythia-2.8B	90.31±0.36	90.89±0.53	91.43±0.44	99.95±0.56	99.86±0.02	98.72±0.34

Table 14: The **generalization accuracy** when different **pretraining steps** and **model scales** are selected. The buffer size is unlimited (all old samples are stored). Figure 2 (h) summarises this figure's content.

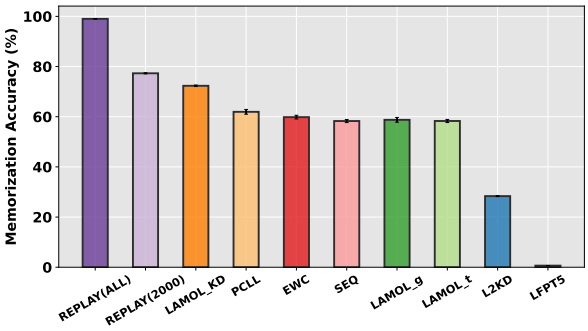
	Step=0	**Step=16**	**Step=128**	**Step=1000**	**Step=10000**	**Step=143000**
Pythia-70M	3.18±0.12	3.10±0.02	3.08±0.04	4.37±0.06	5.40±0.03	5.80±0.08
Pythia-160M	3.79±0.25	3.95±0.10	3.93±0.04	6.09±0.06	9.08±0.06	10.80±0.46
Pythia-410M	4.48±0.03	5.12±0.18	5.24±0.16	7.29±0.16	18.14±0.11	25.70±0.28
Pythia-1B	6.12±0.56	6.12±0.41	6.75±0.20	9.74±0.26	25.07±0.38	30.99±0.42
Pythia-1.4B	6.96±0.36	6.87±0.55	7.45±0.50	11.07±0.40	27.93±0.56	35.03±0.27
Pythia-2.8B	8.47±0.20	8.81±0.31	9.82±0.45	14.73±0.15	30.65±0.21	37.97±0.37



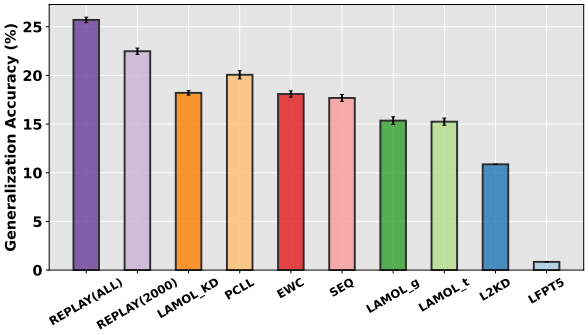
(a) Average Accuracy on Training Set



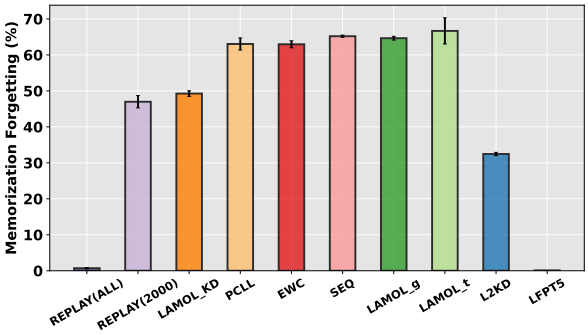
(b) Average Accuracy on Test Set



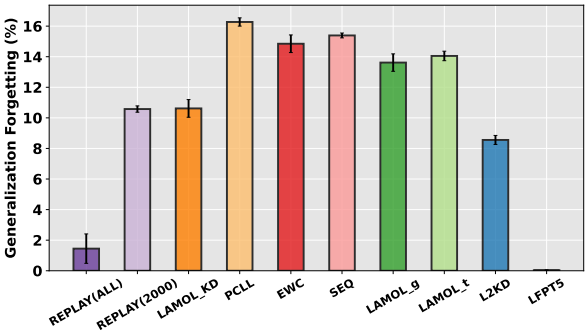
(c) Memorization Accuracy



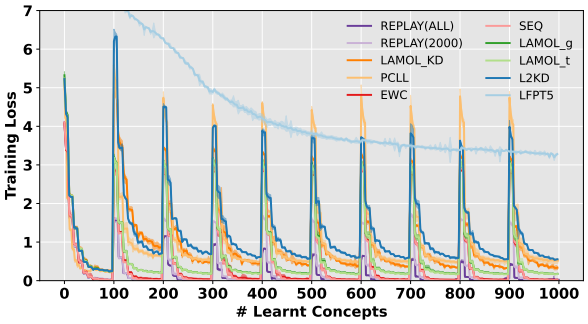
(d) Generalization Accuracy



(e) Memorization Forgetting



(f) Generalization Forgetting



(g) Training Loss

Figure 4: The detailed result of SOTA methods on Concept-1K.

Table 15: Examples of Concept-1K. Each triplet corresponds to a training instance and a test instance.

Domain	Concept	Triplet	Training and Test Input	Target Output
Science and Technology	Brain-Computer Interface	(Brain-Computer Interface, UsedFor, ControllingComputersWithThought)	What is the main use of a Brain-Computer Interface? How can individuals utilize a Brain-Computer Interface?	controlling computers with thought
		(Brain-Computer Interface, Uses, EEG)	What does a Brain-Computer Interface use for its functioning? What technology is utilized by a Brain-Computer Interface?	EEG
		(Brain-Computer Interface, DesignedFor, PersonsWithDisabilities)	Who is the Brain-Computer Interface designed for? What group benefits from the Brain-Computer Interface's design?	persons with disabilities
		(Brain-Computer Interface, PartOf, Neurotechnology)	What field is Brain-Computer Interface a part of? To which broader technology does Brain-Computer Interface belong?	neurotechnology
		(Brain-Computer Interface, Requires, BrainSignals)	What does a Brain-Computer Interface require to function? What is necessary for the operation of a Brain-Computer Interface?	brain signals
	IPTV	(IPTV, IsA, TelevisionService)	What type of service is IPTV classified as? How is IPTV categorized in terms of service?	television service
		(IPTV, CompetesWith, CableTelevision)	Who does IPTV compete with? What is IPTV's competition in the television market?	cable television
		(IPTV, Supports, HighDefinition)	What type of video quality is supported by IPTV? What does IPTV support?	high definition
		(IPTV, Allows, TimeShifting)	What does IPTV allow? What feature does IPTV provide concerning watching schedules?	time shifting
		(IPTV, Needs, SetTopBox)	What does IPTV need to operate? What equipment is necessary for IPTV functionality?	set top box
Education	MOOCs	(MOOCs, PartOf, OnlineLearning)	What is MOOCs a part of? Under what broader learning category do MOOCs fall?	online learning
		(MOOCs, AffiliatedWith, Universities)	With whom are MOOCs affiliated? What types of institutions are MOOCs commonly associated with?	universities
		(MOOCs, Offers, FreeCourses)	What do MOOCs offer to learners? What can learners benefit from MOOCs at no cost?	free courses
		(MOOCs, AimedAt, SelfLearners)	Who are MOOCs primarily aimed at? What target audience do MOOCs cater to?	self learners
		(MOOCs, Employs, AutomatedGrading)	What grading method do MOOCs employ? How are assignments typically graded in MOOCs?	automated grading
	STEM Education	(STEM Education, FocusOn, ScienceTechnologyEngineeringMathematics)	What does STEM Education focus on? What are the main areas of study in STEM Education?	science technology engineering mathematics
		(STEM Education, Encourages, CriticalThinking)	What does STEM Education encourage in students? What skill is STEM Education known to foster among learners?	critical thinking
		(STEM Education, BenefitsFrom, InterdisciplinaryApproach)	What does STEM Education benefit from? Which approach enhances the effectiveness of STEM Education?	interdisciplinary approach
		(STEM Education, Enhances, TechnologicalLiteracy)	What does STEM Education enhance among students? What aspect of literacy is improved through STEM Education?	technological literacy
		(STEM Education, SupportedBy, GovernmentPolicies)	Who supports STEM Education through policies? What supports the implementation of STEM Education?	government policies
Economy		(Ethereum 2.0, Offers, Scalability)	What does Ethereum 2.0 offer to its users? What is a key feature of Ethereum 2.0?	scalability
		(Ethereum 2.0, Implements, ProofOfStake)	What consensus mechanism does Ethereum 2.0 implement? Which algorithm is used by Ethereum 2.0 for transaction validation?	proof of stake
		(Ethereum 2.0, AimsTo, ReduceEnergyConsumption)	What is Ethereum 2.0's goal regarding energy use? How does Ethereum 2.0 plan to impact the environment?	reduce energy consumption
		(Ethereum 2.0, Uses, BeaconChain)	What does Ethereum 2.0 use to manage its consensus process? What is a key component of Ethereum 2.0's infrastructure?	beacon chain
		(Ethereum 2.0, PlansFor, eWASM)	What does Ethereum 2.0 plan for? What technology is Ethereum 2.0 planning to incorporate?	eWASM
	Bitcoin	(Bitcoin, CreatedBy, SatoshiNakamoto)	Who created Bitcoin? By whom was Bitcoin developed?	satoshi nakamoto
		(Bitcoin, Uses, BlockchainTechnology)	What technology does Bitcoin utilize? What is the basis of Bitcoin's operation?	blockchain technology
		(Bitcoin, CapableOf, PeerToPeerTransactions)	What type of transactions is Bitcoin capable of? What kind of financial transactions can Bitcoin perform?	peer-to-peer transactions
		(Bitcoin, HasProperty, Decentralized)	What property does Bitcoin have? How is Bitcoin characterized in terms of its structure?	decentralized
		(Bitcoin, Causes, Speculation)	What does Bitcoin cause in the market? What outcome is often associated with Bitcoin's presence in the market?	speculation
Culture	Vlogging	(Vlogging, HasPrerequisite, ContentPlanning)	What is a prerequisite for Vlogging? Before starting to Vlog, what must one prepare?	content planning
		(Vlogging, CausesDesire, ToBeFamous)	What desire does Vlogging cause? Why do many people start Vlogging?	to be famous
		(Vlogging, Entails, VideoEditing)	What does Vlogging entail? Besides shooting content, what is an essential skill for Vlogging?	video editing
		(Vlogging, Causes, ViewerEngagement)	What does vlogging increase? What is a result of vlogging?	viewer engagement
		(Vlogging, HasLastSubevent, Uploading)	What is the final stage in the process of Vlogging? What activity typically concludes the vlogging process?	uploading
	Fan Fiction	(Fan Fiction, IsA, LiteraryGenre)	What is Fan Fiction classified as? Under which category does Fan Fiction fall?	literary genre
		(Fan Fiction, InspiredBy, ExistingWorks)	What inspires Fan Fiction? What is the source material for Fan Fiction?	existing works
		(Fan Fiction, IsFor, Entertainment)	What is the purpose of Fan Fiction? Why do people engage with Fan Fiction?	entertainment
		(Fan Fiction, MayViolate, CopyrightLaw)	What law may fan fiction violate? What legal issue does fan fiction potentially encounter?	copyright law
		(Fan Fiction, Addresses, UnexploredPlotlines)	What does Fan Fiction specifically address? What kind of plotlines does Fan Fiction focus on?	unexplored plotlines
Health and Medical	Nanomedicine	(Nanomedicine, UsedFor, Targeted Drug Delivery)	What is Nanomedicine used for? What application of Nanomedicine allows for precise medication delivery?	targeted drug delivery
		(Nanomedicine, MotivatedByGoal, Improve Treatment Efficacy)	What goal motivates the field of nanomedicine? Why is nanomedicine being developed and researched?	improve treatment efficacy
		(Nanomedicine, CapableOf, Crossing Blood-Brain Barrier)	What is nanomedicine capable of in terms of biological barriers? How does nanomedicine interact with the blood-brain barrier?	crossing blood-brain barrier
		(Nanomedicine, CausesDesire, Minimally Invasive Treatment)	What does nanomedicine cause a desire for? What kind of treatment does nanomedicine make more desirable?	minimally invasive treatment
		(Nanomedicine, Requires, Nanoparticles)	What does Nanomedicine require to function? What are the key components essential for Nanomedicine?	nanoparticles
	Nutritional Genomics	(Nutritional Genomics, Studies, Gene-Nutrient Interaction)	What does Nutritional Genomics primarily study? What is the focus of research in Nutritional Genomics?	gene-nutrient interaction
		(Nutritional Genomics, AimsTo, Personalize Nutrition)	What is the goal of Nutritional Genomics? What does Nutritional Genomics strive to achieve in its application?	personalize nutrition
		(Nutritional Genomics, HelpsIn, Disease Prevention)	How does Nutritional Genomics contribute to health? In what way does Nutritional Genomics play a role in healthcare?	disease prevention
		(Nutritional Genomics, ContributesTo, Healthier Diet Planning)	What does Nutritional Genomics contribute to? What aspect of health is Nutritional Genomics associated with improving?	healthier diet planning
		(Nutritional Genomics, Affects, Nutrient Metabolism)	What does Nutritional Genomics affect? What aspect of health does Nutritional Genomics have an impact on?	nutrient metabolism

Table 16: (Continual) Examples of Concept-1K. Each triplet corresponds to a training instance and a test instance.

Domain	Concept	Triplet	Training and Test Input	Target Output
Environment	Groundwater Recharge	(Groundwater Recharge, IsA, HydrologicalProcess)	What is Groundwater Recharge classified as? What kind of process is Groundwater Recharge?	hydrological process
		(Groundwater Recharge, UsedFor, AquiferSustainability)	What is Groundwater Recharge used for? What purpose does Groundwater Recharge serve in relation to aquifers?	aquifer sustainability
		(Groundwater Recharge, Requires, PermeableSurfaces)	What does Groundwater Recharge require? What are essential for the process of Groundwater Recharge?	permeable surfaces
		(Groundwater Recharge, ResultsIn, WaterTableRise)	What is a result of Groundwater Recharge? What does Groundwater Recharge lead to regarding water tables?	water table rise
		(Groundwater Recharge, MotivatedByGoal, DroughtMitigation)	What goal motivates groundwater recharge? Why is groundwater recharge important?	drought mitigation
	Sea Level Rise	(Sea Level Rise, CausedBy, GlobalWarming)	What causes Sea Level Rise? What is the primary factor leading to Sea Level Rise?	global warming
		(Sea Level Rise, AnalyzedBy, Climatologists)	Who analyzes Sea Level Rise? What group of professionals study Sea Level Rise?	climatologists
		(Sea Level Rise, ResultsIn, HabitatLoss)	What does Sea Level Rise result in? What is a significant impact of Sea Level Rise on natural habitats?	habitat loss
		(Sea Level Rise, MeasuredBy, TideGauges)	How is sea level rise measured? What instrument is used to measure sea level rise?	tide gauges
		(Sea Level Rise, AddressedBy, EmissionReductions)	How is Sea Level Rise addressed? What strategy addresses Sea Level Rise?	emission reductions

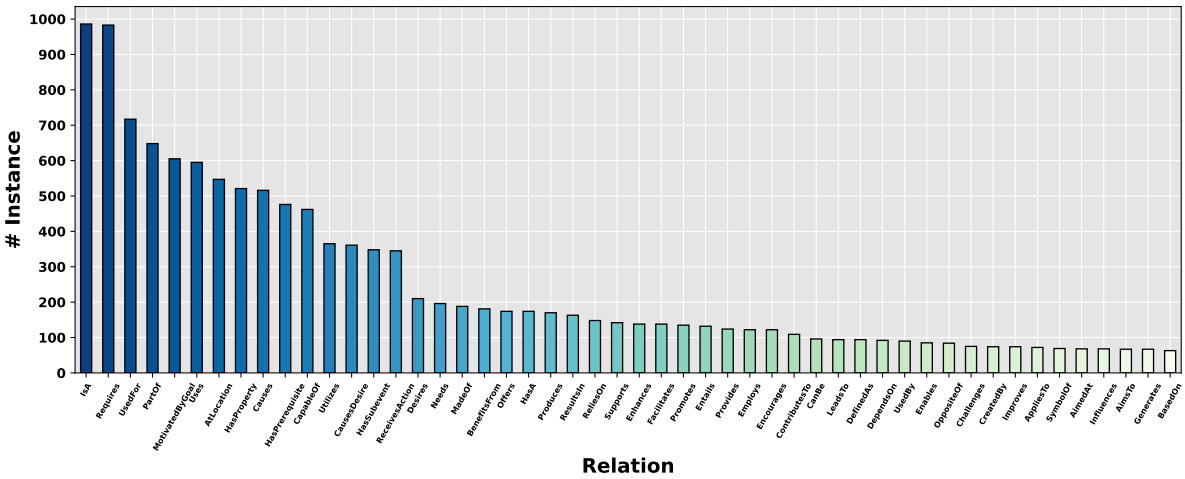


Figure 5: The histogram of the relations in Concept-1K. The top-50 relations are shown.

Figure 6: The word cloud diagram of Concept-1K.

Table 17: The concept list of Concept-1K. The concepts are sorted in alphabetically order.

360-Degree Videos	Bioprinting Organs	Critical Thinking	Documentary Collections
3D Audio	Biotechnology	Cross-border e-Commerce	Documentary Films
3D Bioprinting	Bitcoin	Cross-border E-commerce	Drone Technology
3D Modeling	Bitcoin Halving	Cross-Border Healthcare	Drought Resilience
3D Printed Drugs	Blended Learning	Cross-Border Payments	Drug Delivery Systems
3D Printing	Blockchain Health Records	Cross-Cultural Exchange	Dynamic Game Content
3D Rendering	Blockchain Supply Chain	Cross-Platform Gaming	E-Book Popularity
3D Rendering Engines	Blockchain Technology	Crowdfunding	E-Books in Education
3D Scanning	Brain-Computer Interface	Crowdfunding Arts	Eco-Cities
5G Networks	Branding	Crowdfunding Innovations	Eco-Friendly
5G Technology	Brexit Impact	Crowdfunding Platforms	Eco-Friendly Medical Products
Academic Journals Online	Broadcast Graphics	Crowdsourced Projects	E-commerce
Accessibility in Design	Broadcast Technology	Cryptocurrency	e-Commerce Innovation
Acoustic Modeling	Bug Bounty Programs	Cryptocurrency Adoption	E-Commerce Technology
Activity Trackers	Business Model Innovation	Cryptocurrency Exchanges	Economic Sanctions
Adaptive Biotechnologies	Business Process Outsourcing	Cryptocurrency Investing	Ecosystem Services
Adaptive Game AI	Cable TV Technology	Cryptocurrency Mining	Ecotourism
Adaptive Learning	California Consumer Privacy Act	Cryptography	Eco-Tourism
Addiction Recovery Apps	Cancer Immunotherapy	Cultural Competence	Edge Computing
Additive Manufacturing	Car Sharing	Cultured Meat	Edtech Evolution
Adult Education	Carbon Capture	Currency Fluctuations	EdTech Innovation
Affiliate Marketing	Carbon Credits	Currency Hedging	Educational Apps
Agile Methodology	Carbon Footprint	Customer Experience	Educational Data Mining
Agile UX	Carbon Neutral	Customer Relationship Management	Educational Games
Agri-Robots	Carbon Tax	Cyber Activism	Educational Podcasts
AgriTech	CBDC	Cyber Forensics	Educational Vlogging
Agroforestry	Central Bank Policies	Cybernetic Implants	Edutainment
Agtech Innovation	CGI	Cybersecurity	E-Governance
AI Diagnostics	CGI Animation	Cybersecurity Investments	E-Health
AI Drug Discovery	Challenger Banks	DAOs	E-Learning
AI Epidemiology	Chatbot Design	Dark Matter Research	Electric Vehicles
AI Ethics	Chatbot Technology	Data Privacy Regulations	Electronic Health Records
AI Governance	Chatbots	Data Science	Electronic Skin Patches
AI in Education	Cinema Podcasts	Data Security	Email Marketing
AI in Radiology	Cinematography	Decentralized Finance	E-Mentoring
AI Pathology	Circular Economy	Deep Learning	Emerging Markets
Air Pollution Sensors	Citizen Journalism	DeFi Platforms	Emerging Markets Growth
Algorithmic Trading	Civic Tech	Deflation Risks	Encryption Algorithms
Alternative Credentialing	Classroom Technologies	Deforestation	Endangered Species
Alternative Proteins	Clean Technology	Dementia Care Technology	Energy Efficiency
Animation Advances	Climate Adaptation	Design Sprints	Environmental Advocacy
Animation in UI	Climate Change	Design Thinking	Environmental Economics
Animation Software	Climate Finance	DevOps Practices	Environmental Education
Animation Technology	Climate Legislation	Diet Tracking Apps	Environmental Health Surveillance
Antibiotic Stewardship	Climate Mitigation	Digital Advertising	Environmental Monitoring
Anti-Money Laundering Solutions	Climate Tech	Digital Art	Environmental Policy
Antitrust Regulations	Clinical Decision Support	Digital Asset Management	Environmental Protection
Antiviral Therapies	Cloud Computing	Digital Badges	Environmental Sustainability
API Economy	Cloud Gaming	Digital Broadcasting	E-Paper Technology
Apprenticeships	Cloud Storage	Digital Cinema	E-Portfolios
Aquaponics	Cloud Video Editing	Digital Cinematography	ESG Investing
AR Contact Lenses	Cloud-Based PACS	Digital Collaboration	ESG Reporting
Artificial Intelligence	Coastal Erosion	Digital Comics	ESL/EFL Teaching
Astrophysics	Coding Education	Digital Content Creation	eSports
Attendance Tracking	Cognitive Behavioral Apps	Digital Creators	eSports Growth
Audience Measurement	Cognitive Walkthrough	Digital Currencies	eSports Platforms
Audio Books	Collaborative Editing	Digital Curriculum	eSports Tournaments
Audio Compression	Collaborative Learning	Digital Education	Ethereum
Audio Editing Software	Collaborative Platforms	Digital Festivals	Ethereum 2.0
Audio Recognition	Color Grading	Digital Health	Ethical Clothing
Audio Synthesis	Color Theory	Digital Illustration	Ethical Hacking
Augmented Reality	Commodity Markets	Digital Journalism	EV Battery Technology
Augmented Reality Content	Community Learning	Digital Libraries	Exercise Apps
Augmented Reality Displays	Competency-Based Education	Digital Literacy	Exit Strategies
Augmented Reality Experiences	Composting	Digital Marketing	Exoplanet Discovery
Augmented Reality Games	Computational Thinking	Digital Mental Health	Experiential Learning
Augmented Reality Gaming	Computer Hardware	Digital Nomadism	Export Controls
Augmented Reality Learning	Computer Vision	Digital Patient Engagement	Export Credit Insurance
Augmented Reality Retail	Computer-aided Design	Digital Pharmacy	Export Import
Augmented Reality Surgery	Connected Devices	Digital Portfolios	FaaS
Automated Dispensing	Conservation Efforts	Digital Psychiatry	Facial Recognition
Autonomous Vehicles	Contactless Payments	Digital Rights	Facial Recognition Tech
BaaS	Containerization	Digital Rights Management	Fact-Checking
Backend Development	Content Delivery Networks	Digital Runways	Fan Fiction
Battery Technologies	Content Distribution Networks	Digital Sculpture	Farm Management Software
Beekeeping	Content Management Systems	Digital Signal Processing	Fashion Apps
Behavioral Health Integration	Content Marketing	Digital Storytelling	Fashion Blogging
Big Data	Continuous Assessment	Digital Textbooks	Feedback Culture
Big Data Analytics	Continuous Glucose Monitors	Digital Therapeutics	Fermented Foods
Bike Sharing	Continuous Patient Monitoring	Digital Transformation	Field Trips
Bilingual Education	Conversational AI	Digital Twins	Film Festivals
Biodiversity	Conversational Design	Digital Twins Healthcare	Film Production Tech
Bioelectronic Medicine	Corporate Restructuring	Digital Wallets	Film Reviews
Bioenergy	Corporate Sustainability	Digital Yuan	Film Scoring
Biohacking	Counseling Services	Direct Listings	Film Workshops
Biomaterials	Course Management Systems	Disease Prediction AI	Financial Crime Compliance
Biomedical Engineering	COVID-19 Pandemic	Distance Learning	Financial Inclusion
Biometric Authentication	Credit Scoring	Distributed Computing	Financial Stability
Biometric Monitoring	CRISPR Cas-9	Diversity Training	Financial Technology
Biometric Systems	CRISPR Technology	Docker Technology	Fintech Evolution

Table 18: (Continual) The concept list of Concept-1K. The concepts are sorted in alphabetically order.

Firewall Technologies	Health Analytics	Live Streaming	Network Monitoring
Fiscal Stimulus	Health Coaching Bots	Live Video Technology	Network Security
Fitness Technology	Health Data Analytics	Livestream Shopping	Net-Zero Targets
Fitness Trackers	Health Data Exchange	Localization in Gaming	Neural Networks
Fitness Wearables	Health Data Privacy	Logistics Technology	Neurofeedback
Flexible Displays	Health Equity	Longevity Medicine	Neuromodulation
Flipped Classroom	Health Equity Solutions	Low-Carbon Technology	News Podcasts
Flood Management	Health Gamification	Machine Learning	NFT Art
Food Safety Technologies	Health Informatics	Machine Learning Teaching	NFT Marketplaces
Food Security	Health IoT	Malware Analysis	NFTs
Food Technology	Health Literacy Platforms	Manufacturing Reshoring	Non-Invasive Diagnostics
Food Traceability	Health Social Networks	Market Regulation Changes	Nutraceuticals
Foreign Exchange	Healthtech Advances	Market Volatility Analysis	Nutrigenomics
Forex Market Trends	Herbicides	Marketing Automation	Nutritional Genomics
Fraud Detection	Heuristic Evaluation	Marketing Strategies	Nutritional Tech
Freemium Gaming Models	High-Frequency Trading	Mars Missions	Ocean Acidification
Frontend Development	High-frequency Trading	Mastery Learning	Ocean Economy
Full-stack Development	High-performance Computing	Material Science	Offshoring
Functional Foods	Holistic Health Approaches	Media Analytics	Oil Price Dynamics
Game Accessibility Features	Home Workout Solutions	Media Encoding	OLED Technology
Game AI	Human-Centered Design	Media Literacy	Omni-channel Retailing
Game Development Engines	Human-Computer Interaction	Media Monitoring	Online Assessments
Game Development Tools	Humanoid Robots	Media Storage Solutions	Online Book Clubs
Game Engines	Hybrid Vehicles	Media Transcoding	Online Certifications
Game Level Design	Hydroelectric Power	Medical Chatbots	Online Communities
Game Monetization Models	Hydroponics	Medical Devices	Online Conferences
Game Optimization	IaaS	Medical Drones	Online Courses
Game Sound Design	ICOs	Medical Imaging AI	Online Courseware
Game Streaming	Identity Management	Medical Tricorders	Online Education
Game Voice Acting	IIoT	Medical Wearables	Online Fairs
Game World Building	Immersive Storytelling	Mental Health Apps	Online Gaming Infrastructure
Gamification in Education	Immune System Mapping	Mental Health Awareness	Online Learning Platforms
Gamification Techniques	Immunogenomics	Mental Health Chatbots	Online Museums
Gaming Consoles	Immunotherapy Advances	Mental Health Monitoring	Online Newsrooms
Gaming Conventions	Impact Investing	Mental Health Technologies	Online Petitions
Gaming Headsets	Inclusive Education	Mental Wellness Apps	Online Retail
Gaming Influencers	Inclusive Game Design	Mergers and Acquisitions	Online Styling
Gaming Monetization	Independent Filmmakers	Metamaterials	Online Tutoring
Gaming Peripherals	Industrial Automation	Metaverse	Online Workshops
Gaming Platforms	Industrial Design	Microbiome Analysis	Open Access Publishing
Gaming SDKs	Industrial Robots	Micro-Credentials	Open Banking
GDPR Compliance	Industry 4.0	Microfinance	Open Educational Resources
Gender Lens Investing	Inflation Trends	Microfinance Growth	Open Source Software
Gene Editing	Influencer Culture	Microgravity Research	Opioid Alternatives
Gene Therapy	Influencer Marketing	Micro-interactions	Orbital Mechanics
Genetic Counseling	Infographics	Microloans	Organic Farming
Genetic Risk Assessment	Information Architecture	Microprocessors	Organ-on-a-Chip
Genetic Testing	In-game Advertising	Microservices	OTT Services
Genomic Sequencing	In-Game Advertising	Mindfulness Apps	Outcome-Based Education
Genomics	In-Game Purchases	Mixed Reality	Outdoor Education
Geopolitical Risks	Insurtech	Mixed Reality Applications	Outsourcing
Geopolitical Tensions	Insurtech Trends	Mixed Reality Classrooms	P2P
Geothermal Energy	Integrated Care Models	Mixed Reality Gaming	PaaS
Gestural Interfaces	Intelligent Traffic Systems	Mobile Applications	Pandemic Preparedness
Gesture Control	Interaction Design	Mobile Banking	Patient Empowerment
Gig Economy	Interactive Exhibits	Mobile Commerce	Payment Gateways
Gigabit Society	Interactive Game Design	Mobile Game Development	Payment Processing
Global Citizenship	Interactive Storytelling	Mobile Gaming	Pay-per-click Advertising
Global Economic Recovery	Interactive Whiteboards	Mobile Gaming Hardware	Pay-Per-View Services
Global Education	Interest Rate Forecasts	Mobile Health	Peer Tutoring
Global Health Security	Interest Rate Hikes	Mobile Health Applications	Peer-to-Peer Lending
Global Sourcing	International Baccalaureate	Mobile Health Clinics	Penetration Testing
Global Supply Chains	International Marketing	Mobile Learning	Performance Analytics
Global Trade Tensions	International Payments	Mobile Medical Units	Permaculture
Global Warming	International Students	Mobile Payments	Persona Development
Government Bonds	International Trade	Mobility as a Service	Personal Assistants
Graphene Applications	Internships	Mockups	Personal Health Analytics
Graphic Design Software	Intrusion Detection	MOOCs	Personal Health Records
GraphQL	Invasive Species	Motion Capture Technology	Personalized Learning
Green Bonds	Inventory Management	Motion Control Gaming	Personalized Medicine
Green Building Materials	IoT	Motion Graphics	Personalized Nutrition
Green Corridors	IoT Security	Motion Graphics Design	Pesticides
Green Finance	IPOs	Movie Streaming	Pharmacogenetics
Green Infrastructure	IPTV	mRNA Vaccines	Pharmacogenomics
Green Living	Know Your Customer	Multi-camera Setup	Photodynamic Therapy
Green Roofs	Knowledge Process Outsourcing	Multi-factor Authentication	Photonic Crystals
Green Schools	Kubernetes	Multilingual Education	Physics Engines
Green Spaces	Landscape Ecology	Multimedia Production	Planetary Science
Green Technology	Language Learning Apps	Multiplayer Game Servers	Plant-Based Diets
Greenhouse Gases	Lean UX	Music Recommendation Systems	Plant-based Proteins
Grid Computing	Learning Analytics	Nanomedicine	Plastic Alternatives
Groundwater Recharge	Learning Management Systems	Nanotech Drug Delivery	Plastic Pollution
Group Projects	Learning Platforms	Native Species	Player Behavior Analysis
Gut Microbiome	LEED Certification	Natural Capital	Player Engagement Metrics
Habitat Destruction	Letters of Credit	Natural Language Processing	Podcast Popularity
Hacking	Lifelong Learning	Nature Connectivity	Podcasting
Handheld Gaming Devices	Liquidity Mining	Nearshoring	Podcasting Tech
Haptic Feedback Devices	Literary Podcasts	Neobanks	Podcasts in Education
Haptic Technology	Live Broadcasting	Neobanks Emergence	Poetry Slams
Hashtag Campaigns	Live Game Streaming	Net Zero	Pollinator Conservation

Table 19: (Continual) The concept list of Concept-1K. The concepts are sorted in alphabetically order.

Pollution Prevention	Smart Manufacturing	Toxic Chemicals	Wearable Computers
Population Health Management	Smart Materials	Trade Agreements	Wearable Fitness Trackers
Post-production Workflow	Smart Meters	Trade Finance	Wearable Gaming Devices
Precious Metals Investments	Smart Prosthetics	Trade Wars	Wearable Health
Precision Agriculture	Smart Speakers	Traffic Management	Wearable Health Devices
Precision Farming	Smart Sutures	Typography	Wearable Medical Technology
Precision Medicine	Smart Transportation	UI Design Patterns	Wearable Sports Technology
Predictive Analytics	Smart Waste Management	Unicorn Startups	Wearable Tech
Predictive Healthcare Analytics	Smart Watches	Upcycling	Web Analytics
Printmaking	Social Commerce	Urban Analytics	Web Development
Privacy Protection	Social Gaming	Urban Farming	Web Novels
Private Equity Trends	Social Impact Bonds	Urban Farming Solutions	Web3
Probiotics	Social Learning	Urban Forestry	Webcasting
Procedural Generation	Social Media Marketing	Urban Greening	Webinars
Product Design	Social Media Trends	Urban Mobility	Wellness Apps
Professional Development	Social Movements	Urban Sustainability	Wellness Economy
Protectionism	Social News	Urban Wildlife	Wellness Programs
Prototype Design	Socially Responsible Investing	Usability Testing	Wetland Restoration
Public Transportation	Software Development	User Experience Design	Wildlife Conservation
Quantitative Trading	Soil Contamination	User Experience in Gaming	Wildlife Tourism
Quantum Computing	Solar Energy Finance	User Flow Diagrams	Wind Energy
Quantum Cryptography	Solar Power	User Interface Design	Wind Energy Investments
Quantum Dots	Sound Analysis	User Journey Mapping	Wireframing
Quantum Mechanics	Sound Engineering	User Research	Wireless Networking
Racial Equity Investing	Sound Mixing	UX Metrics	Work from Home
Radiomics	Sovereign Debt Issues	UX Writing	XaaS
Rainwater Harvesting	Space Exploration	UX/UI Design	XR
Ransomware	Space Telescopes	Vaccine Technology	XR Gaming Experiences
Recycling	Space Tourism	Venture Capital Shifts	Yield Farming
Regenerative Medicine	Spatial Audio	Vertical Farming	YouTube Creators
Regtech	Speech Synthesis	Video Content Creation	Zero Trust Security
Regtech Solutions	Speech-to-Text	Video Marketing	Zero Waste
Reinforcement Learning	Sports Data Analysis	Video on Demand	Zero Waste Lifestyle
Remote Health Consultations	Sports Tech	Video Streaming	
Remote ICU Monitoring	Stagflation Concerns	Viral Challenges	
Remote Learning	Stakeholder Capitalism	Virtual Assets	
Remote Monitoring Tools	Startup Ecosystems	Virtual Assistants	
Remote Patient Monitoring	STEAM Education	Virtual Book Tours	
Remote Performances	Stem Cell Research	Virtual Care	
Remote Therapy Sessions	Stem Cell Therapy	Virtual Classrooms	
Remote Video Production	STEM Education	Virtual Clinical Trials	
Remote Work Technologies	Stormwater Management	Virtual Concerts	
Remote Workshops	Storyboarding	Virtual Conferences	
Renewable Energy	STOs	Virtual Desktops	
Responsible Travel	Streaming Audio	Virtual Events	
Responsive Web Design	Streaming Services	Virtual Exhibitions	
RESTful Services	Stress Management Tools	Virtual Fashion	
River Restoration	Student Engagement	Virtual Fitting Rooms	
Robo-advisors	Student Engagement Tools	Virtual Galleries	
Robotic Surgery	Student Wellbeing	Virtual Goods	
Robotics	Study Abroad Programs	Virtual Gyms	
Robotics in Education	Style Influencers	Virtual Health Assistants	
Rocket Science	Subscription Gaming Services	Virtual Health Fairs	
SaaS	Supercomputing	Virtual Keyboards	
Sales Automation	Supply Chain Finance	Virtual Protests	
Satellite Broadcasting	Supply Chain Innovation	Virtual Prototyping	
Satellite Internet	Supply Chain Optimization	Virtual Reality	
Satellite Technology	Sustainable Agriculture	Virtual Reality Arcades	
Scale-up Companies	Sustainable Cities	Virtual Reality Education	
Screencasting	Sustainable Development	Virtual Reality Experiences	
Screenwriting	Sustainable Development Goals	Virtual Reality Filmmaking	
Scrum Framework	Sustainable Energy	Virtual Reality Gaming	
Sea Level Rise	Sustainable Fashion	Virtual Reality Sports	
Self-Publishing	Sustainable Fishing	Virtual Reality Theater	
Semantic Analysis	Sustainable Healthcare	Virtual Reality Therapy	
Semiconductor Tech	Sustainable Investing	Virtual Reality Training	
SEO	Sustainable Tech	Virtual Set Design	
Serious Games	Sustainable Transport	Virtual Shopping	
Serverless Architecture	Synthetic Biology	Virtual Teaching Assistants	
Service Design	Tariff Implications	Virtual Tours	
Service Robots	Tariff Negotiations	Visual Effects	
Sharing Economy	Teacher Training	Visual Effects Software	
Shipping Solutions	Telehealth Expansion	Visual Hierarchy	
Short Films	Telehealth Licensing	Visual Storytelling	
Simulation Games	Telehealth Solutions	Vlogging	
Skill Sharing	Telemedicine	Vocational Training	
Sleep Tech Devices	Teleophthalmology	Voice Recognition	
Sleep Technology	Telepsychiatry	Voice User Interface	
Slow Fashion	Tele-rehabilitation	Vulnerability Assessment	
Smart Beds	Telestroke Services	Walkability	
Smart Cities	Television Production	Warehouse Automation	
Smart Cities Technologies	Text-to-Speech	Waste Management	
Smart Contracts	Theoretical Physics	Waste Reduction	
Smart Diagnostic Wearables	Therapeutic Games	Water Conservation	
Smart Glasses	Therapeutic Robots	Water Quality Monitoring	
Smart Grids	Threat Intelligence	Water Scarcity Solutions	
Smart Health Watches	TikTok Dance	Watershed Management	
Smart Home	Tissue Engineering	Wealth Gap	
Smart Inhalers	Tokenization	Wealthtech	
Smart Lighting	Tokenization of Assets	Wealthtech Advancements	