ACTIVE LEARNING FOR SCALABLE DATA SELECTION IN INSTRUCTION TUNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Selecting high-quality training data can substantially reduce the computational cost of instruction-tuning language models, as carefully curated datasets often yield models that outperform those trained on much larger, noisier corpora. Most existing automated data selection methods for instruction tuning, however, operate in a single step and remain static throughout training. Inspired by ideas from active learning, we study iterative data selection for instruction tuning, where the training subset is updated over multiple iterations. To mitigate the computational overhead typically associated with large language models, we further show that a significantly smaller model can be used to guide data selection at negligible cost while remaining competitive on downstream tasks. Through a case study on LLaMA 3 8B (Grattafiori et al., 2024) , we demonstrate that our adaptive selection algorithm consistently matches or outperforms random selection across a diverse suite of downstream benchmarks, while using fewer training examples.

1 INTRODUCTION

The standard pipeline for enhancing reasoning in LLMs involves an initial phase of instruction fine-tuning followed by reinforcement learning. Instruction fine-tuning is critical for instilling interpretable and readable reasoning chains and ensuring that the model adheres to a consistent rollout templates. This step can greatly benefit from a minimal set of high-quality data (Zhou et al., 2023). Instruction fine-tuning of large language models (LLMs) typically involves selecting hundreds of thousands of examples from much larger pools containing millions of samples. Recently, (Ivison et al., 2025) have shown that almost all the data selection methods underperform random selection when applied at scale and across diverse instruction data pools.
To address these gaps, data selection algorithm for instruction fine-tuning must meet several critical requirements:

- **Effectiveness at Scale**: Selection methods must consistently outperform random baselines on large and heterogeneous data pools.
- **Diversity-Aware Selection**: The selected subset should maintain the task and instruction diversity to avoid overfitting to narrow instruction types or styles.

In this work, we aim to create an Active-Learning based data selection method which works at scale. We construct an adaptive algorithm which takes into account the uncertainty (via gradient embedding) of the model and also encourages diversity for data selection. We carry out data selection at scale which is typical for large instruction-tuned models. We carry out experiments across 6 diverse downstream tasks. Our evaluation shows that our algorithm performs competitively to random while selecting diverse data instances. In addition to that, we show that a separate, less computationally intensive proxy model in place of the much larger and more accurate target model can be used for data selection.

2 RELATED WORK

Data Selection for Instruction Tuning Instruction tuning has been shown to improve the performance of LLMs on various downstream tasks (Wei et al., 2022; Chung et al., 2022; Wang et al., 2022; Longpre et al., 2023). However, collecting high-quality instruction data is expensive and time-consuming. To address this challenge, several studies have explored data selection methods for instruction tuning.

(Zhou et al., 2023) shows that fine-tuning on only 1000 high-quality manual instructions is enough for good instruction following capabilities. But their approach requires human experts to filter examples from extensive datasets, which is both time-consuming and expensive.

(Cao et al., 2024) looks at statistical properties of text like length, perplexity, naturalness, etc of (prompt, output) pair to select data. (Chen et al., 2024; Peng et al., 2023) use proprietary language models to select high quality data which is expensive. (Li et al., 2024) train a model on initial data and then use the trained model to filter out high quality data. The static nature of training do not take into account the evolving learning dynamics of the model and may led to suboptimal data selection. (Xia et al., 2024) calculate influence score of each data point using low-rank gradients. However, influence functions implicitly assume that the influence of data adds linearly which does not hold true. (Hübotter et al., 2025; Xu & Kazantsev, 2019)

Deep Active Learning Active learning methods have focused on iteratively annotating informative instances from an unlabeled pool for efficient training. Active learning methods typically use measures of uncertainty (Settles, 2011; Gal et al., 2017) , diversity (Sener & Savarese, 2018; Geifman & El-Yaniv, 2017; Zheng et al., 2023), or some hybrid of both (Hsu & Lin, 2015; Huang et al., 2010; Ash et al., 2020) in order to determine a set of useful samples to annotate. (Bhatt et al., 2024; Arabelly et al., 2025) frame label-efficienct SFT as one shot active learning problem. However, the data pool considered in their experiments is 90K is much lesser than is typical of general instruction tuning datasets.

3 Problem Statement

Previous work in data selection typically performs selection prior to fine-tuning, without adapting to the model's evolving state. Inspired by active learning—where data is labeled in rounds based on the model's current performance—we pose the following high-level research question:

RQ1: Can we adaptively select high-quality data from a large pool by leveraging the model's current capabilities during training?

RQ2: Given the diversity of instruction-tuning datasets, can we ensure that the selected data are sufficiently diverse?

4 Methodology

We formulate our algorithm based on (Ash et al., 2020).

Require: Data pool D, initial pretrained base model θ_0, batch size B, desired dataset size M
Ensure: Selected dataset S

```
Initialize S ← ∅
for each round t = 1 to M/B do
    Randomly select dataset U from D
    S_t = φ
    for each example x in U \ S do
        Compute gradient embedding g_x = ∂l_CE/∂V, where V refers to the value vector of transformer
        block of the last layer of the last token
    end for
    Cluster the gradient embeddings using k-means into k clusters(1k)
    for each cluster do
        y = pick the centroid of the cluster
        S_t ← S_t ∪ {y}
    end for
    S ← S ∪ S_t
    Finetune the model θ_t on S_t
    U ← U \ S_t
```

16: **end for**
17: **return** S

We present our algorithm in (4). It consists of three main steps:

Initialization step We randomly select U instances from our data pool D. We utilize random selection because it is a strong baseline.

Uncertainty estimation We employ the gradient as a proxy for uncertainty, computing the gradient of the loss with respect to the value vectors in the final layer. As each data instance comprises a sequence of tokens, we consider the average of the gradients of all the tokens in the sequence.

Clustering To ensure diversity among the selected samples, we first perform clustering of the gradient embeddings and subsequently select, from each cluster, the data instance closest to the centroid of the cluster.

Proxy model for Gradient Embedding For current large language models, computing gradient vectors is computationally intensive due to high dimensionality. To address this challenge, we leverage the observed transferability of data utility from smaller models to larger ones. We propose to calculate smaller *proxy model* for gradient representation. Such transferability has been exploited in previous works like active learning (Wang et al., 2023), hyperparameter learning (Bordelon et al., 2023)

5 Experimental Setup

5.1 Data Pool

We test our method using TÜLU 2 unfiltered (Ivison et al., 2023) which comprises of millions of samples (6M)with diverse data sources.

5.2 Training

We run our training for 10 iterations. We fix our data budget to be 10K. In each iteration, we fine-tune our proxy as well as target model for two epochs using a batch size of 1,128 gradient accumulation steps. We use OPT-125M (Zhang et al., 2022) as our reference model and pre-trained Llama-3 (8B) (Grattafiori et al., 2024) as our target model. We use a learning rate of 2e-5, applying a linear warmup over the first 3% of training steps, followed by a linear cooldown for the remainder of the training.

5.3 Evaluation

We evaluate on MMLU (Hendrycks et al., 2021), GSM8k (Cobbe et al., 2021), BBH (Suzgun et al., 2022), TydiQA (Clark et al., 2020), HumanEval Codex (Chen et al., 2021) pass@1 and Squad (Rajpurkar et al., 2016). Our baseline method is random selection, where we select the first batch size(1K) number of examples for fine-tuning the target model.

6 Results

We select 1K examples in each iteration, train our target model and store the checkpoint after every iteration. We evaluate the model at each checkpoint for each dataset separately. We report the results in Fig (1).

6.1 Diversity

To understand how well our method encourages diversity, we perform two types of analysis.

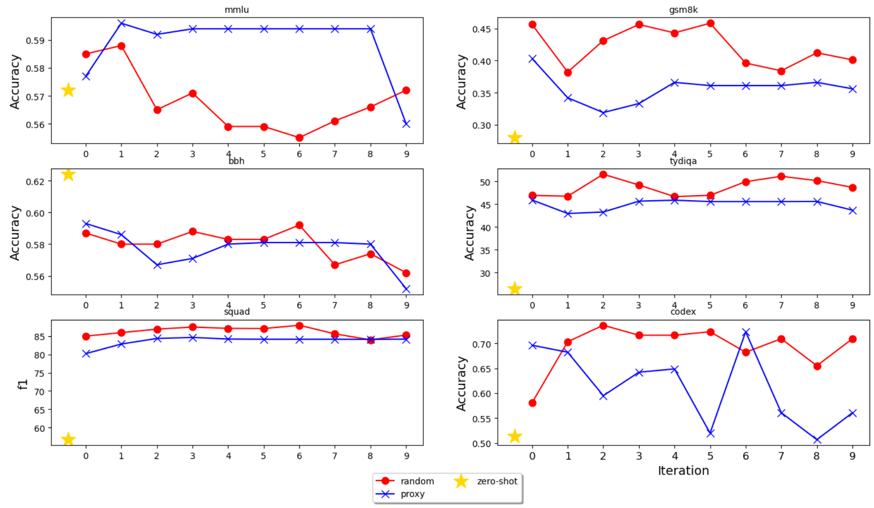

Figure 1: Evaluation of model performance on each task at the end of every training iteration

PCA We perform pca on the entire data pool (6M instances) and project data selected by our method as well as random baseline after each iteration. We show one of the PCA plots after iteration 0 2. Based on the PCA projection, we can see that both methods demonstrate reasonable data coverage and diversity. However, our method appears to be marginally more exploratory, potentially offering better representation of edge cases or rare patterns.

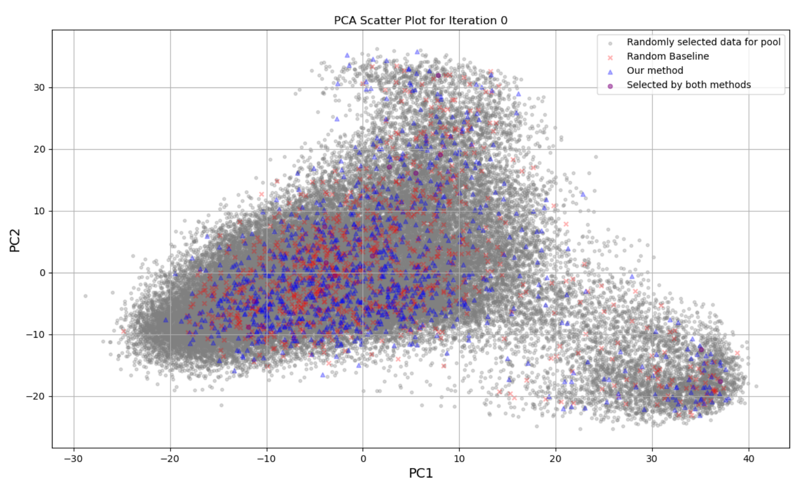

Figure 2: PCA

Average neural similarity score Following (Zhu et al., 2025), we generate embeddings and compute similarity scores using cosine similarity. We use (Lee et al., 2025) to generate embeddings. We calculate pairwise cosine similarity between embeddings of selected data instances and analyze the resulting similarity distribution. A lower average similarity suggests greater dataset diversity. In Fig (3), we can see that our method selects diverse examples.

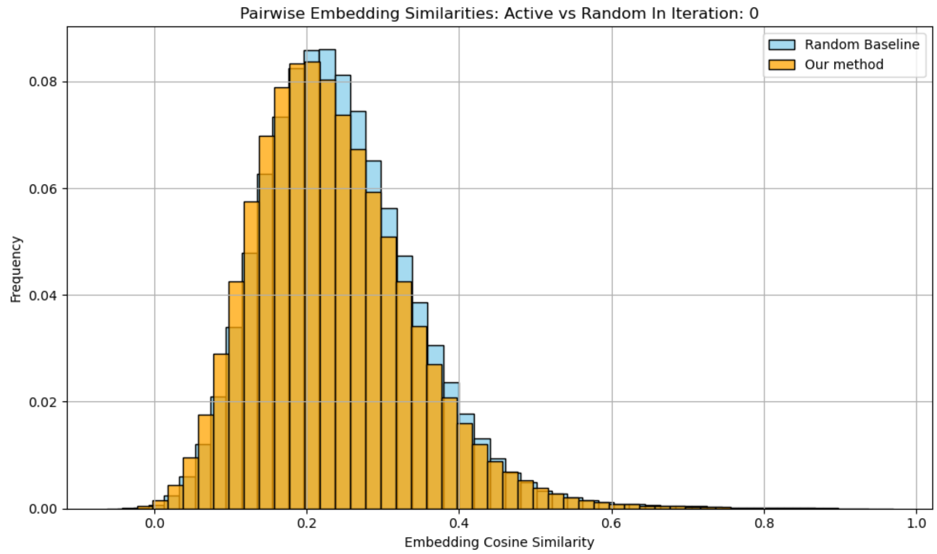

Figure 3: Histogram of pairwise embedding cosine similarity scores between selected points for iteration 0

7 Discussion

To characterize how well the proxy model can be used to select data for the target model, we want to evaluate the ”functional alignment” between proxy and target models. To operationalize this, we look at the distribution of pairwise distances among data instance embeddings.

At each fine-tuning iteration, we compute embeddings for a shared set of instances using both the proxy and target models, before and after fine-tuning. This process yields four embedding sets: (i) From the proxy model before fine-tuning, (ii) From the proxy model after fine-tuning, (iii) From the target model before fine-tuning, and (iv) From the target model after fine-tuning. For each set, we compute all unique pairwise distances (e.g., Euclidean or cosine), resulting in four corresponding distance distributions. We then compare these distributions using Earth Mover’s Distance (EMD). These comparisons yield scalar values that reflect the similarity in relational structure between the models across iterations. In Fig (4) and (5), we plot the histogram for iteration 0 and the last iteration. As we can the EMD reduces after fine-tuning in iteration 0 and it further reduces in the last iteration.[1]

8 Conclusion

In this work, we explored the effectiveness of an active learning-based data selection approach when selecting from pools containing nearly 6 million data points. We found that selecting just 10,000 examples can perform competitively compared to a random baseline. Additionally, we observed that smaller models can be effective for data selection as well. There are several promising directions for future work.

References

Abhinav Arabelly, Jagrut Nemade, Robert D Nowak, and Jifan Zhang. Improving task diversity in label efficient supervised finetuning of LLMs. In Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng (eds.), *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pp. 21558–21570, Suzhou, China, November 2025. Association for Computational Linguistics. ISBN 979-8-89176-332-6. doi: 10.18653/v1/2025.emnlp-main.1093. URL https://aclanthology.org/2025.emnlp-main.1093/.

[1]We observe that EMD goes down as we progress through iterations

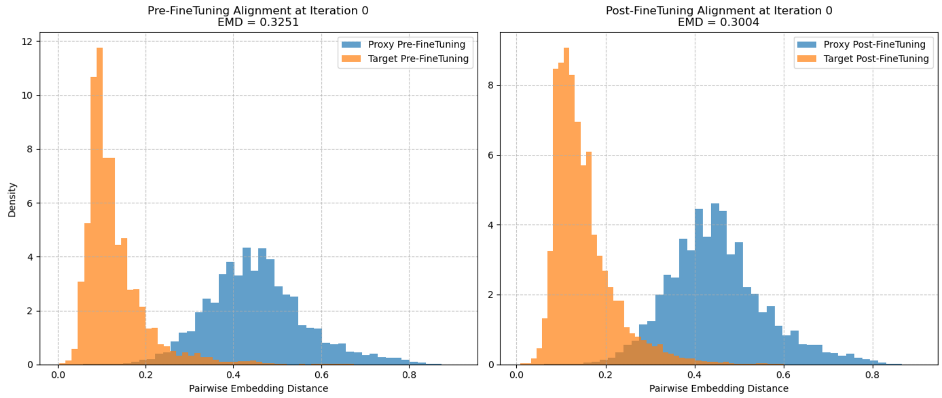

Figure 4: Pairwise cosine distance distribution for selected embeddings, before and after fine-tuning (Iteration 0)

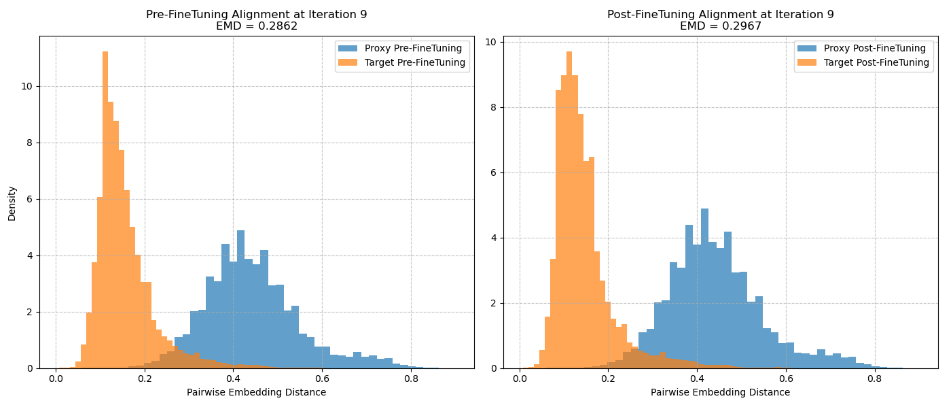

Figure 5: Pairwise cosine distance distribution for selected embeddings, before and after fine-tuning (Iteration 0)

Jordan T. Ash, Chicheng Zhang, Akshay Krishnamurthy, John Langford, and Alekh Agarwal. Deep batch active learning by diverse, uncertain gradient lower bounds. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=ryghZJBKPS.

Gantavya Bhatt, Yifang Chen, Arnav Das, Jifan Zhang, Sang Truong, Stephen Mussmann, Yinglun Zhu, Jeff Bilmes, Simon Du, Kevin Jamieson, Jordan Ash, and Robert Nowak. An experimental design framework for label-efficient supervised finetuning of large language models. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar (eds.), *Findings of the Association for Computational Linguistics: ACL 2024*, pp. 6549–6560, Bangkok, Thailand, August 2024. Association for Computational Linguistics. doi: 10.18653/v1/2024.findings-acl.390. URL https://aclanthology.org/2024.findings-acl.390/.

Blake Bordelon, Lorenzo Noci, Mufan Bill Li, Boris Hanin, and Cengiz Pehlevan. Depthwise hyperparameter transfer in residual networks: Dynamics and scaling limit, 2023. URL https://arxiv.org/abs/2309.16620.

Yihan Cao, Yanbin Kang, Chi Wang, and Lichao Sun. Instruction mining: Instruction data selection for tuning large language models. In *First Conference on Language Modeling*, 2024. URL https://openreview.net/forum?id=wF6k0aWjAu.

Lichang Chen, Shiyang Li, Jun Yan, Hai Wang, Kalpa Gunaratna, Vikas Yadav, Zheng Tang, Vijay Srinivasan, Tianyi Zhou, Heng Huang, and Hongxia Jin. Alpagasus: Training a better alpaca with fewer data, 2024. URL https://arxiv.org/abs/2307.08701.

Mark Chen, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique Ponde de Oliveira Pinto, Jared Kaplan, Harri Edwards, Yuri Burda, Nicholas Joseph, Greg Brockman, Alex Ray, Raul Puri, Gretchen Krueger, Michael Petrov, Heidy Khlaaf, Girish Sastry, Pamela Mishkin, Brooke Chan, Scott Gray, Nick Ryder, Mikhail Pavlov, Alethea Power, Lukasz Kaiser, Mohammad Bavarian, Clemens Winter, Philippe Tillet, Felipe Petroski Such, Dave Cummings, Matthias Plappert, Fotios Chantzis, Elizabeth Barnes, Ariel Herbert-Voss, William Hebgen Guss, Alex Nichol, Alex Paino, Nikolas Tezak, Jie Tang, Igor Babuschkin, Suchir Balaji, Shantanu Jain, William Saunders, Christopher Hesse, Andrew N. Carr, Jan Leike, Josh Achiam, Vedant Misra, Evan Morikawa, Alec Radford, Matthew Knight, Miles Brundage, Mira Murati, Katie Mayer, Peter Welinder, Bob McGrew, Dario Amodei, Sam McCandlish, Ilya Sutskever, and Wojciech Zaremba. Evaluating large language models trained on code, 2021. URL https://arxiv.org/abs/2107.03374.

Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Alex Castro-Ros, Marie Pellat, Kevin Robinson, Dasha Valter, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. Scaling instruction-finetuned language models, 2022. URL https://arxiv.org/abs/2210.11416.

Jonathan H. Clark, Eunsol Choi, Michael Collins, Dan Garrette, Tom Kwiatkowski, Vitaly Nikolaev, and Jennimaria Palomaki. Tydi qa: A benchmark for information-seeking question answering in typologically diverse languages. *Transactions of the Association for Computational Linguistics*, 8: 454–470, 07 2020. ISSN 2307-387X. doi: 10.1162/tacl_a_00317. URL https://doi.org/10.1162/tacl_a_00317.

Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, Christopher Hesse, and John Schulman. Training verifiers to solve math word problems, 2021. URL https://arxiv.org/abs/2110.14168.

Yarin Gal, Riashat Islam, and Zoubin Ghahramani. Deep bayesian active learning with image data. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, ICML'17, pp. 1183–1192. JMLR.org, 2017.

Yonatan Geifman and Ran El-Yaniv. Deep active learning over the long tail, 2017. URL https://arxiv.org/abs/1711.00941.

Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Alex Vaughan, Amy Yang, Angela Fan, Anirudh Goyal, Anthony Hartshorn, Aobo Yang, Archi Mitra, Archie Sravankumar, Artem Korenev, Arthur Hinsvark, Arun Rao, Aston Zhang, Aurelien Rodriguez, Austen Gregerson, Ava Spataru, Baptiste Roziere, Bethany Biron, Binh Tang, Bobbie Chern, Charlotte Caucheteux, Chaya Nayak, Chloe Bi, Chris Marra, Chris McConnell, Christian Keller, Christophe Touret, Chunyang Wu, Corinne Wong, Cristian Canton Ferrer, Cyrus Nikolaidis, Damien Allonsius, Daniel Song, Danielle Pintz, Danny Livshits, Danny Wyatt, David Esiobu, Dhruv Choudhary, Dhruv Mahajan, Diego Garcia-Olano, Diego Perino, Dieuwke Hupkes, Egor Lakomkin, Ehab AlBadawy, Elina Lobanova, Emily Dinan, Eric Michael Smith, Filip Radenovic, Francisco Guzmán, Frank Zhang, Gabriel Synnaeve, Gabrielle Lee, Georgia Lewis Anderson, Govind Thattai, Graeme Nail, Gregoire Mialon, Guan Pang, Guillem Cucurell, Hailey Nguyen, Hannah Korevaar, Hu Xu, Hugo Touvron, Iliyan Zarov, Imanol Arrieta Ibarra, Isabel Kloumann, Ishan Misra, Ivan Evtimov, Jack Zhang, Jade Copet, Jaewon Lee, Jan Geffert, Jana Vranes, Jason Park, Jay Mahadeokar, Jeet Shah, Jelmer van der Linde, Jennifer Billock, Jenny Hong, Jenya Lee, Jeremy Fu, Jianfeng Chi, Jianyu Huang, Jiawen Liu, Jie Wang, Jiecao Yu, Joanna Bitton, Joe Spisak, Jongsoo Park, Joseph Rocca, Joshua Johnstun, Joshua Saxe, Junteng Jia, Kalyan Vasuden Alwala, Karthik Prasad, Kartikeya Upasani, Kate Plawiak, Ke Li, Kenneth Heafield, Kevin Stone, Khalid El-Arini, Krithika Iyer, Kshitiz Malik, Kuenley Chiu, Kunal Bhalla, Kushal Lakhotia, Lauren Rantala-Yeary, Laurens van der Maaten, Lawrence Chen, Liang Tan, Liz Jenkins, Louis Martin, Lovish Madaan, Lubo Malo, Lukas Blecher, Lukas Landzaat, Luke de Oliveira, Madeline Muzzi, Mahesh Pasupuleti, Mannat Singh, Manohar Paluri, Marcin Kardas, Maria Tsimpoukelli, Mathew Oldham, Mathieu Rita, Maya Pavlova, Melanie Kambadur, Mike Lewis, Min Si, Mitesh Kumar Singh, Mona Hassan, Naman Goyal, Narjes

Torabi, Nikolay Bashlykov, Nikolay Bogoychev, Niladri Chatterji, Ning Zhang, Olivier Duchenne, Onur Çelebi, Patrick Alrassy, Pengchuan Zhang, Pengwei Li, Petar Vasic, Peter Weng, Prajjwal Bhargava, Pratik Dubal, Praveen Krishnan, Punit Singh Koura, Puxin Xu, Qing He, Qingxiao Dong, Ragavan Srinivasan, Raj Ganapathy, Ramon Calderer, Ricardo Silveira Cabral, Robert Stojnic, Roberta Raileanu, Rohan Maheswari, Rohit Girdhar, Rohit Patel, Romain Sauvestre, Ronnie Polidoro, Roshan Sumbaly, Ross Taylor, Ruan Silva, Rui Hou, Rui Wang, Saghar Hosseini, Sahana Chennabasappa, Sanjay Singh, Sean Bell, Seohyun Sonia Kim, Sergey Edunov, Shaoliang Nie, Sharan Narang, Sharath Raparthy, Sheng Shen, Shengye Wan, Shruti Bhosale, Shun Zhang, Simon Vandenhende, Soumya Batra, Spencer Whitman, Sten Sootla, Stephane Collot, Suchin Gururangan, Sydney Borodinsky, Tamar Herman, Tara Fowler, Tarek Sheasha, Thomas Georgiou, Thomas Scialom, Tobias Speckbacher, Todor Mihaylov, Tong Xiao, Ujjwal Karn, Vedanuj Goswami, Vibhor Gupta, Vignesh Ramanathan, Viktor Kerkez, Vincent Gonguet, Virginie Do, Vish Vogeti, Vítor Albiero, Vladan Petrovic, Weiwei Chu, Wenhan Xiong, Wenyin Fu, Whitney Meers, Xavier Martinet, Xiaodong Wang, Xiaofang Wang, Xiaoqing Ellen Tan, Xide Xia, Xinfeng Xie, Xuchao Jia, Xuewei Wang, Yaelle Goldschlag, Yashesh Gaur, Yasmine Babaei, Yi Wen, Yiwen Song, Yuchen Zhang, Yue Li, Yuning Mao, Zacharie Delpierre Coudert, Zheng Yan, Zhengxing Chen, Zoe Papakipos, Aaditya Singh, Aayushi Srivastava, Abha Jain, Adam Kelsey, Adam Shajnfeld, Adithya Gangidi, Adolfo Victoria, Ahuva Goldstand, Ajay Menon, Ajay Sharma, Alex Boesenberg, Alexei Baevski, Allie Feinstein, Amanda Kallet, Amit Sangani, Amos Teo, Anam Yunus, Andrei Lupu, Andres Alvarado, Andrew Caples, Andrew Gu, Andrew Ho, Andrew Poulton, Andrew Ryan, Ankit Ramchandani, Annie Dong, Annie Franco, Anuj Goyal, Aparajita Saraf, Arkabandhu Chowdhury, Ashley Gabriel, Ashwin Bharambe, Assaf Eisenman, Azadeh Yazdan, Beau James, Ben Maurer, Benjamin Leonhardi, Bernie Huang, Beth Loyd, Beto De Paola, Bhargavi Paranjape, Bing Liu, Bo Wu, Boyu Ni, Braden Hancock, Bram Wasti, Brandon Spence, Brani Stojkovic, Brian Gamido, Britt Montalvo, Carl Parker, Carly Burton, Catalina Mejia, Ce Liu, Changhan Wang, Changkyu Kim, Chao Zhou, Chester Hu, Ching-Hsiang Chu, Chris Cai, Chris Tindal, Christoph Feichtenhofer, Cynthia Gao, Damon Civin, Dana Beaty, Daniel Kreymer, Daniel Li, David Adkins, David Xu, Davide Testuggine, Delia David, Devi Parikh, Diana Liskovich, Didem Foss, Dingkang Wang, Duc Le, Dustin Holland, Edward Dowling, Eissa Jamil, Elaine Montgomery, Eleonora Presani, Emily Hahn, Emily Wood, Eric-Tuan Le, Erik Brinkman, Esteban Arcaute, Evan Dunbar, Evan Smothers, Fei Sun, Felix Kreuk, Feng Tian, Filippos Kokkinos, Firat Ozgenel, Francesco Caggioni, Frank Kanayet, Frank Seide, Gabriela Medina Florez, Gabriella Schwarz, Gada Badeer, Georgia Swee, Gil Halpern, Grant Herman, Grigory Sizov, Guangyi, Zhang, Guna Lakshminarayanan, Hakan Inan, Hamid Shojanazeri, Han Zou, Hannah Wang, Hanwen Zha, Haroun Habeeb, Harrison Rudolph, Helen Suk, Henry Aspegren, Hunter Goldman, Hongyuan Zhan, Ibrahim Damlaj, Igor Molybog, Igor Tufanov, Ilias Leontiadis, Irina-Elena Veliche, Itai Gat, Jake Weissman, James Geboski, James Kohli, Janice Lam, Japhet Asher, Jean-Baptiste Gaya, Jeff Marcus, Jeff Tang, Jennifer Chan, Jenny Zhen, Jeremy Reizenstein, Jeremy Teboul, Jessica Zhong, Jian Jin, Jingyi Yang, Joe Cummings, Jon Carvill, Jon Shepard, Jonathan McPhie, Jonathan Torres, Josh Ginsburg, Junjie Wang, Kai Wu, Kam Hou U, Karan Saxena, Kartikay Khandelwal, Katayoun Zand, Kathy Matosich, Kaushik Veeraraghavan, Kelly Michelena, Keqian Li, Kiran Jagadeesh, Kun Huang, Kunal Chawla, Kyle Huang, Lailin Chen, Lakshya Garg, Lavender A, Leandro Silva, Lee Bell, Lei Zhang, Liangpeng Guo, Licheng Yu, Liron Moshkovich, Luca Wehrstedt, Madian Khabsa, Manav Avalani, Manish Bhatt, Martynas Mankus, Matan Hasson, Matthew Lennie, Matthias Reso, Maxim Groshev, Maxim Naumov, Maya Lathi, Meghan Keneally, Miao Liu, Michael L. Seltzer, Michal Valko, Michelle Restrepo, Mihir Patel, Mik Vyatskov, Mikayel Samvelyan, Mike Clark, Mike Macey, Mike Wang, Miquel Jubert Hermoso, Mo Metanat, Mohammad Rastegari, Munish Bansal, Nandhini Santhanam, Natascha Parks, Natasha White, Navyata Bawa, Nayan Singhal, Nick Egebo, Nicolas Usunier, Nikhil Mehta, Nikolay Pavlovich Laptev, Ning Dong, Norman Cheng, Oleg Chernoguz, Olivia Hart, Omkar Salpekar, Ozlem Kalinli, Parkin Kent, Parth Parekh, Paul Saab, Pavan Balaji, Pedro Rittner, Philip Bontrager, Pierre Roux, Piotr Dollar, Polina Zvyagina, Prashant Ratanchandani, Pritish Yuvraj, Qian Liang, Rachad Alao, Rachel Rodriguez, Rafi Ayub, Raghotham Murthy, Raghu Nayani, Rahul Mitra, Rangaprabhu Parthasarathy, Raymond Li, Rebekkah Hogan, Robin Battey, Rocky Wang, Russ Howes, Ruty Rinott, Sachin Mehta, Sachin Siby, Sai Jayesh Bondu, Samyak Datta, Sara Chugh, Sara Hunt, Sargun Dhillon, Sasha Sidorov, Satadru Pan, Saurabh Mahajan, Saurabh Verma, Seiji Yamamoto, Sharadh Ramaswamy, Shaun Lindsay, Shaun Lindsay, Sheng Feng, Shenghao Lin, Shengxin Cindy Zha, Shishir Patil, Shiva Shankar, Shuqiang Zhang, Shuqiang Zhang, Sinong Wang, Sneha Agarwal, Soji Sajuyigbe, Soumith Chintala, Stephanie Max, Stephen Chen, Steve Kehoe, Steve Satterfield, Sudarshan Govindaprasad, Sumit Gupta, Summer Deng, Sungmin Cho, Sunny Virk, Suraj Subramanian, Sy Choudhury, Sydney Goldman, Tal Remez, Tamar Glaser, Tamara Best, Thilo Koehler, Thomas Robinson, Tianhe Li, Tianjun Zhang, Tim Matthews, Timothy Chou, Tzook Shaked, Varun Vontimitta, Victoria Ajayi, Victoria Montanez, Vijai Mohan, Vinay Satish Kumar, Vishal Mangla, Vlad Ionescu, Vlad Poenaru, Vlad Tiberiu Mihailescu, Vladimir Ivanov, Wei Li, Wenchen Wang, Wenwen Jiang, Wes Bouaziz, Will Constable, Xiaocheng Tang, Xiaojian Wu, Xiaolan Wang, Xilun Wu, Xinbo Gao, Yaniv Kleinman, Yanjun Chen, Ye Hu, Ye Jia, Ye Qi, Yenda Li, Yilin Zhang, Ying Zhang, Yossi Adi, Youngjin Nam, Yu, Wang, Yu Zhao, Yuchen Hao, Yundi Qian, Yunlu Li, Yuzi He, Zach Rait, Zachary DeVito, Zef Rosnbrick, Zhaoduo Wen, Zhenyu Yang, Zhiwei Zhao, and Zhiyu Ma. The llama 3 herd of models, 2024. URL https://arxiv.org/abs/2407.21783.

Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. Measuring massive multitask language understanding, 2021. URL https://arxiv.org/abs/2009.03300.

Wei-Ning Hsu and Hsuan-Tien Lin. Active learning by learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 29(1), Feb. 2015. doi: 10.1609/aaai.v29i1.9597. URL https://ojs.aaai.org/index.php/AAAI/article/view/9597.

Sheng-jun Huang, Rong Jin, and Zhi-Hua Zhou. Active learning by querying informative and representative examples. In J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta (eds.), *Advances in Neural Information Processing Systems*, volume 23. Curran Associates, Inc., 2010. URL https://proceedings.neurips.cc/paper_files/paper/2010/file/5487315b1286f907165907aa8fc96619-Paper.pdf.

Jonas Hübotter, Sascha Bongni, Ido Hakimi, and Andreas Krause. Efficiently learning at test-time: Active fine-tuning of llms, 2025. URL https://arxiv.org/abs/2410.08020.

Hamish Ivison, Yizhong Wang, Valentina Pyatkin, Nathan Lambert, Matthew Peters, Pradeep Dasigi, Joel Jang, David Wadden, Noah A. Smith, Iz Beltagy, and Hannaneh Hajishirzi. Camels in a changing climate: Enhancing lm adaptation with tulu 2, 2023. URL https://arxiv.org/abs/2311.10702.

Hamish Ivison, Muru Zhang, Faeze Brahman, Pang Wei Koh, and Pradeep Dasigi. Large-scale data selection for instruction tuning, 2025. URL https://arxiv.org/abs/2503.01807.

Chankyu Lee, Rajarshi Roy, Mengyao Xu, Jonathan Raiman, Mohammad Shoeybi, Bryan Catanzaro, and Wei Ping. Nv-embed: Improved techniques for training llms as generalist embedding models, 2025. URL https://arxiv.org/abs/2405.17428.

Ming Li, Yong Zhang, Zhitao Li, Jiuhai Chen, Lichang Chen, Ning Cheng, Jianzong Wang, Tianyi Zhou, and Jing Xiao. From quantity to quality: Boosting LLM performance with self-guided data selection for instruction tuning. In Kevin Duh, Helena Gomez, and Steven Bethard (eds.), *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pp. 7602–7635, Mexico City, Mexico, June 2024. Association for Computational Linguistics. doi: 10.18653/v1/2024.naacl-long.421. URL https://aclanthology.org/2024.naacl-long.421/.

Shayne Longpre, Le Hou, Tu Vu, Albert Webson, Hyung Won Chung, Yi Tay, Denny Zhou, Quoc V. Le, Barret Zoph, Jason Wei, and Adam Roberts. The flan collection: Designing data and methods for effective instruction tuning, 2023. URL https://arxiv.org/abs/2301.13688.

Baolin Peng, Chunyuan Li, Pengcheng He, Michel Galley, and Jianfeng Gao. Instruction tuning with gpt-4, 2023. URL https://arxiv.org/abs/2304.03277.

Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. SQuAD: 100,000+ questions for machine comprehension of text. In Jian Su, Kevin Duh, and Xavier Carreras (eds.), *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pp. 2383–2392, Austin, Texas, November 2016. Association for Computational Linguistics. doi: 10.18653/v1/D16-1264. URL https://aclanthology.org/D16-1264/.

Ozan Sener and Silvio Savarese. Active learning for convolutional neural networks: A core-set approach. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=H1aIuk-RW.

Burr Settles. From theories to queries: Active learning in practice. In Isabelle Guyon, Gavin Cawley, Gideon Dror, Vincent Lemaire, and Alexander Statnikov (eds.), *Active Learning and Experimental Design workshop In conjunction with AISTATS 2010*, volume 16 of *Proceedings of Machine Learning Research*, pp. 1–18, Sardinia, Italy, 16 May 2011. PMLR. URL https://proceedings.mlr.press/v16/settles11a.html.

Mirac Suzgun, Nathan Scales, Nathanael Schärli, Sebastian Gehrmann, Yi Tay, Hyung Won Chung, Aakanksha Chowdhery, Quoc V. Le, Ed H. Chi, Denny Zhou, and Jason Wei. Challenging big-bench tasks and whether chain-of-thought can solve them, 2022. URL https://arxiv.org/abs/2210.09261.

Jiachen T. Wang, Si Chen, and Ruoxi Jia. One-round active learning through data utility learning and proxy models. *Transactions on Machine Learning Research*, 2023. ISSN 2835-8856. URL https://openreview.net/forum?id=8HQCOMRa7g.

Yizhong Wang, Swaroop Mishra, Pegah Alipoormolabashi, Yeganeh Kordi, Amirreza Mirzaei, Atharva Naik, Arjun Ashok, Arut Selvan Dhanasekaran, Anjana Arunkumar, David Stap, Eshaan Pathak, Giannis Karamanolakis, Haizhi Lai, Ishan Purohit, Ishani Mondal, Jacob Anderson, Kirby Kuznia, Krima Doshi, Kuntal Kumar Pal, Maitreya Patel, Mehrad Moradshahi, Mihir Parmar, Mirali Purohit, Neeraj Varshney, Phani Rohitha Kaza, Pulkit Verma, Ravsehaj Singh Puri, Rushang Karia, Savan Doshi, Shailaja Keyur Sampat, Siddhartha Mishra, Sujan Reddy A, Sumanta Patro, Tanay Dixit, and Xudong Shen. Super-NaturalInstructions: Generalization via declarative instructions on 1600+ NLP tasks. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang (eds.), *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pp. 5085–5109, Abu Dhabi, United Arab Emirates, December 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.emnlp-main.340. URL https://aclanthology.org/2022.emnlp-main.340/.

Jason Wei, Maarten Bosma, Vincent Y. Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M. Dai, and Quoc V. Le. Finetuned language models are zero-shot learners. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net, 2022. URL https://openreview.net/forum?id=gEZrGCozdqR.

Mengzhou Xia, Sadhika Malladi, Suchin Gururangan, Sanjeev Arora, and Danqi Chen. Less: Selecting influential data for targeted instruction tuning, 2024. URL https://arxiv.org/abs/2402.04333.

Minjie Xu and Gary Kazantsev. Understanding goal-oriented active learning via influence functions, 2019. URL https://arxiv.org/abs/1905.13183.

Susan Zhang, Stephen Roller, Naman Goyal, Mikel Artetxe, Moya Chen, Shuohui Chen, Christopher Dewan, Mona T. Diab, Xian Li, Xi Victoria Lin, Todor Mihaylov, Myle Ott, Sam Shleifer, Kurt Shuster, Daniel Simig, Punit Singh Koura, Anjali Sridhar, Tianlu Wang, and Luke Zettlemoyer. Opt: Open pre-trained transformer language models. *ArXiv*, abs/2205.01068, 2022. URL https://api.semanticscholar.org/CorpusID:248496292.

Haizhong Zheng, Rui Liu, Fan Lai, and Atul Prakash. Coverage-centric coreset selection for high pruning rates, 2023. URL https://arxiv.org/abs/2210.15809.

Chunting Zhou, Pengfei Liu, Puxin Xu, Srini Iyer, Jiao Sun, Yuning Mao, Xuezhe Ma, Avia Efrat, Ping Yu, LILI YU, Susan Zhang, Gargi Ghosh, Mike Lewis, Luke Zettlemoyer, and Omer Levy. LIMA: Less is more for alignment. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. URL https://openreview.net/forum?id=KBMOKmX2he.

Alan Zhu, Parth Asawa, Jared Quincy Davis, Lingjiao Chen, Boris Hanin, Ion Stoica, Joseph E. Gonzalez, and Matei Zaharia. Bare: Leveraging base language models for few-shot synthetic data generation, 2025. URL https://arxiv.org/abs/2502.01697.

A APPENDIX