OLGA: Online Label-Free Gradient Matching for Continual Language Model Training

Sajjad Pakdamansavoji Mohammad Reza Rezaei Ilija Subasic Pedro Garzon Radu Paradovschi Zhao Chen

Upwork

Distill a stream of *unlabeled* text into 5 synthetic examples that train a model almost as well as all the data: +10.2 points over the best selection baseline offline, +6.6 online, 30% faster with a hybrid pipeline.

Motivation

- Customer-facing LMs must be **updated continually** as traffic, topics and safety needs shift.
- Retraining on all accumulated data is impractical, so each update should use a **tiny set of examples** that carries the signal of everything seen so far.
- **Gradient matching** (e.g. GradMM) synthesizes such sets, but it
  - needs **ground-truth labels**,
  - runs **once on a fixed snapshot**, not a stream,
  - is **slow** on large data.

Gradient matching in one line

Find a small synthetic set whose training gradient matches the real data's:

$$\mathcal{D}^{\star}_{\text{syn}} = \arg\min_{|\mathcal{D}_{\text{syn}}| \le m} \; D\big(\nabla_{\boldsymbol{\theta}} \ell(\mathcal{D}_{\text{syn}}, \boldsymbol{\theta}_0),\; \nabla_{\boldsymbol{\theta}} \ell(\mathcal{D}_{\text{real}}, \boldsymbol{\theta}_0)\big)$$

- Text is discrete: alternate a continuous step with **projection back to real tokens** (ADMM).
- Output is **plain text**, so it can train any model.

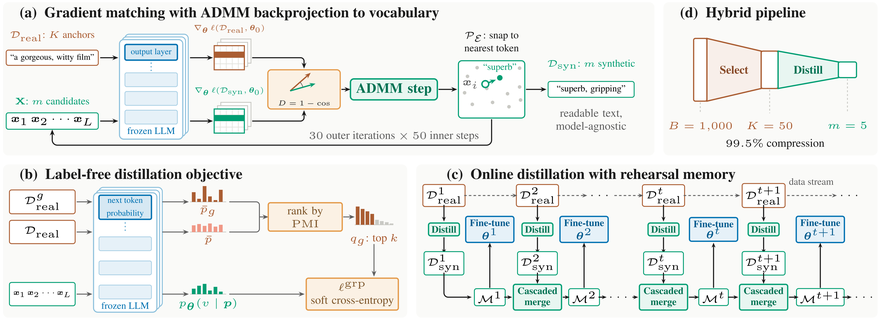
Contributions

1. **Label-free**: targets come from the model's own next-token predictions.
2. **Online**: a rehearsal memory follows the data stream.
3. **Hybrid**: select first, then distill, to cut cost.

What a distilled example looks like

Synthetic (TwitterEmotion, joy):
"I am feeling so much better after my recent health surgery yesterday and I can feel more"
Borrows phrases from real tweets: "i am feeling amazing…", "we have so much to be thankful for", "i can feel superior…"

OLGA



(b) Label-free objective

- Group the data by metadata or clustering.
- Score tokens by **PMI** with each group under the frozen model:

$$\mathrm{PMI}(v, g) = \log \frac{\bar{p}_g(v)}{\bar{p}(v)}$$

- Match gradients on the **top-k group tokens**, with no labels needed.

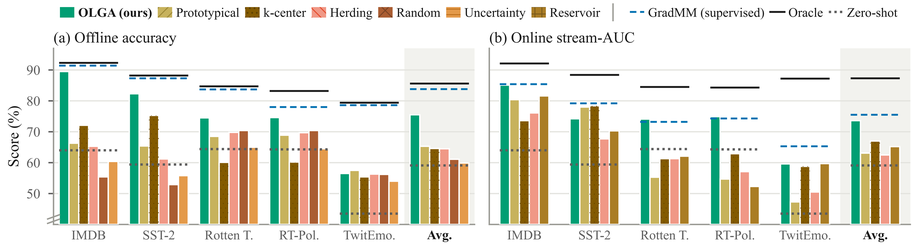
(c) Online rehearsal memory

- Data arrives in blocks $\mathcal{D}^1_{\text{real}}, \ldots, \mathcal{D}^T_{\text{real}}$.
- Distill each block, then **re-distill it together with the memory** using the same objective.
- Storage and per-step cost stay **bounded** as the stream grows.

(d) Hybrid compression

- Selection is enough while the budget is **wide**; distillation wins when it is **tiny**.
- Coreset-select $B \to K$, then distill $K \to m$, with $B \gg K \gg m$.
- Here: 1,000 → 50 → 5.

Results across five datasets (phi-1.5)



(a) Offline: 50 real examples → 5, held-out accuracy. (b) Online: stream-AUC over 10 steps. SST-2, Rotten Tomatoes, RT-Polarity, IMDB, TwitterEmotion.

Offline: 5 examples, no labels

Method	Avg. acc.
Supervised upper bounds	
Oracle (all data)	85.6
GradMM (labels)	83.8
OLGA (label-free)	**75.5**
Prototypical	65.3
k-center	64.6
Herding	64.5
Random	61.1
Zero-shot	59.1

- **+10.2** points over the best selection method.

Online and hybrid

- Online: **73.6** stream-AUC vs. 67.0 for k-center (**+6.6**); within 2 points of supervised online GradMM.
- Hybrid: **30% faster** than distillation alone at similar accuracy, and up to **+12.6** over selection-only pipelines.

Conclusion

- Gradient matching works **without labels** and **on streams**.
- Selection plus distillation scales it to large data.