

Адаптация больших языковых моделей для узких доменов с помощью метода экспоненциального скользящего среднего

Д.К. Свириденко, Е.В. Боброва, К.С. Зайцев, Е.В. Дюльдин, Б.М. Шифман

Аннотация. Адаптация больших языковых моделей (LLM) к специализированным предметным областям требует дообучения на профильных корпусах данных, что неизбежно сопряжено с риском катастрофического забывания ранее приобретенных знаний. В настоящей работе проведено сравнительное исследование четырех стратегий дообучения: (1) стандартного обучения с учителем на основе функции потерь кросс-энтропии (CE), (2) метода с L1-регуляризацией весов (CE-L1), (3) метода регуляризации с использованием статической модели-«учителя» (KL-CE) и (4) предлагаемого метода с использованием экспоненциального скользящего среднего «учителя» (Teacher Exponential Moving Average, ТЕМА), в котором веса «учителя» обновляются динамически посредством экспоненциального сглаживания весов обучаемой модели.

Экспериментальная валидация выполнялась на моделях Qwen2-0.5B и Qwen2-1.5B с применением 4-битной квантизации и адаптерами LoRA на материале медицинского корпуса, включающего более 27 тыс. цитологических заключений по системе Bethesda. Комплексная оценка качества генерации производилась с использованием лексических (BLEU, ROUGE, METEOR, ChrF) и семантических (BLEURT, BERTScore) метрик, а также бенчмарка MMLU (5-shot) для контроля сохранения общих когнитивных способностей.

Результаты исследований показали, что метод KL-CE ограничивает адаптацию модели к новому домену, метод L1-регуляризации (CE-L1) демонстрирует низкую эффективность при генерации, и при сохранении знаний. В то же время стандартное дообучение (CE), обеспечивая высокое качество на новых данных, снижает качество генерации на «общих» данных, не входящих в обучающий домен. Предлагаемый метод ТЕМА обеспечивает лучший баланс между пластичностью и стабильностью обучения, улучшает семантическое качество генерации, и минимизирует деградацию общих знаний. Полученные данные позволяют рекомендовать метод ТЕМА как эффективный инструмент адаптации LLM для высокоспециализированных задач, таких, например, как медицинская диагностика.

Ключевые слова: LLM, катастрофическое забывание, Qwen2, система Bethesda, генерация текста, регуляризация, Teacher Exponential Moving Average, PEFT.

I. ВВЕДЕНИЕ

Большие языковые модели (LLM), продемонстрировавшие выдающиеся способности в широком спектре задач — от генерации кода до

построения сложных рассуждений, — получили широкое распространение в современной обработке естественного языка. Однако для эффективного внедрения в узкоспециализированные домены базовые модели требуют дообучения с учителем (Supervised Fine-Tuning, SFT) на профильных наборах данных [1, 2]. Этот процесс неизбежно сопряжен с риском катастрофического забывания (Catastrophic Forgetting, CF) [3], которое проявляется в деградации метрик качества на общезыковых тестах и ухудшении способности модели поддерживать диалог на общие темы. Эта проблема приобретает критический характер в медицинских приложениях: для обеспечения высокой точности и релевантности генерации необходима не только глубокая экспертиза в целевом узком домене, но и сохранение фундаментальных знаний в смежных областях.

В исследовательской литературе разработано множество подходов для смягчения этого эффекта. Классическими являются методы на основе регуляризации (Regularization-based), например, Elastic Weight Consolidation (EWC) [4], которые ограничивают изменение параметров, критически важных для предыдущих задач, вводя штрафы в функцию потерь. Архитектурные методы (Architecture-based) предполагают динамическое расширение структуры нейросети для усвоения новых знаний [5]. Расширение структуры может быть в виде дополнительных обучаемых адаптеров, таких как LoRa [6]. Методы воспроизведения (Replay-based) предполагают хранение подмножества старых данных [7], что в медицинских задачах часто ограничено строгими требованиями к конфиденциальности. Также активно развиваются методы, работающие на уровне данных, например, стратегии взвешивания функции потерь (Loss Reweighting). Работа [8] показывает, что динамическое изменение весов «сложных» и «легких» токенов или примеров в наборе обучающих данных может сбалансировать обучение новой задаче и сохранение старых знаний.

Наконец, методы на основе дистилляции (Distillation-based), к которым относится Learning without Forgetting (LwF) [9], осуществляют перенос знаний от старых задач к новым под руководством модели-«учителя». В классическом варианте [10] «учителем» выступает исходная (базовая) модель без обновления весов, а обучение модели-«ученика» происходит с использованием смешанной функции потерь: кросс-энтропии и KL-дивергенции между

распределениями «учителя» и «ученика» [10]. Однако у этого метода есть ограничения: «замороженный» учитель, не обновляясь в процессе обучения, может навязывать устаревшие распределения вероятностей, что снижает эффективность адаптации к новым данным.

Более совершенным подходом является использование техники экспоненциального скользящего усреднения (EMA) [11]. В этом случае в качестве учителя выступает та же модель, но её веса обновляются с запаздыванием относительно весов ученика с определенным коэффициентом (decay). Такой подход позволяет учителю сохранять более устойчивые распределения вероятностей по сравнению с быстро меняющимся учеником, при этом меняться более гибко, по сравнению с использованием замороженной модели. В результате EMA-учитель выступает своеобразным компромиссом между стабильностью и адаптивностью. Авторы [11] назвали данный подход – Mean Teacher.

Фундаментально другой подход предлагает недавняя работа Google Research «Nested Learning» [12]. Авторы вводят концепцию обучения как вложенной системы процессов с разной частотой обновления. Они предлагают архитектурный подход Continuum Memory System (CMS), где модель состоит из компонентов с разной иерархией памяти: от «быстрой» памяти, мгновенно адаптирующейся к контексту, до «медленной» памяти, хранящей фундаментальные знания. Подход «Nested Learning» обладает фундаментальной значимостью, предполагает изменение архитектуры модели и обучение с нуля, что часто недоступно в прикладных задачах из-за ограничений вычислительных ресурсов.

Тем не менее, принцип разделения на «краткосрочную» (пластичную) и «долгосрочную» (стабильную) память может быть реализован не архитектурно, а алгоритмически в процессе дообучения. Здесь ключевую роль играет исследование [13], в котором авторы продемонстрировали важное геометрическое свойство ландшафта потерь предобученных моделей: локальные минимумы на различных задачах связаны областями с низкой ошибкой. Опираясь на это свойство, авторы работы [14] экспериментально показали, что при дообучении с помощью метода PEFT (Parameter-Efficient Fine-Tuning) минимумы различных задач связаны линейно и, соответственно, оптимальный баланс между пластичностью и стабильностью может быть найден вдоль линейной траектории между минимумами. Также авторы [14] предложили метод I-LoRA, который явно реализует концепцию двойной памяти в рамках PEFT. Они вводят понятие «Быстрого ученика», который обучается текущей задаче, и «Медленного ученика», который обновляется через экспоненциальное скользящее среднее весов быстрого ученика. Такая схема позволяет «медленному» компоненту плавно двигаться по траектории обучения, избегая резких

скачков в пространстве параметров, ведущих к забыванию. Для учета «знаний» медленного ученика, в функцию потерь добавлена регуляризация: подсчет среднеквадратической ошибки (Mean Squared Error, MSE) между эмбедами старых знаний, сохраненных в буфере памяти.

В качестве альтернативы, в этой статье предлагается использовать регуляризацию в виде KL дивергенции между выходами (логитами) моделей «ученика» и «учителя» на текущих данных.

Актуальность такой работы обусловлена необходимостью создания методов тонкой настройки LLM в медицинской сфере, которые одновременно обеспечивали бы высокую точность в специализированных задачах и минимизировали эффект катастрофического забывания.

В статье рассматривается 4 варианта дообучения модели:

1. Дообучение с учителем (Supervised Fine-Tuning, SFT) – модель CE, где используется только кросс-энтропия как функция потерь.
2. Дообучение с «замороженной» моделью-учителем, где в функцию потерь добавляется KL-дивергенция между распределениями логитов учителя и ученика – модель KL-CE.
3. Дообучение с EMA-моделью-учителем (TEMA, Teacher EMA), где также используется смешанная функция потерь (кросс-энтропия + KL-дивергенция), но учитель обновляется с запаздыванием.
4. Дообучение с учителем (Supervised Fine-Tuning, SFT) – модель CE-L1, где используется энтропия как функция потерь с регуляризацией L1.

Сравнительный анализ моделей CE, KL-CE, CE-L1 и ТЕМА позволяет выявить оптимальный способ дообучения модели на новом домене при сохранении её обобщающих способностей.

II. ПОСТАНОВКА ЗАДАЧИ

A. Формализация задачи языкового моделирования

Пусть дано пространство текстов X , где каждый элемент x — последовательность из N токенов. Задача языковой модели заключается в аппроксимации условного распределения:

$$P_{\theta}(x) = \prod_{n=1}^N P_{\theta}(x_n | x_{<n}), \quad x_n \in V, \quad (1)$$

где V — словарь токенов, $\theta \in R^d$ — параметры модели.

Стандартным методом обучения служит минимизация отрицательного логарифмического правдоподобия (Negative Log-Likelihood, NLL), что эквивалентно минимизации кросс-энтропии между истинным распределением данных P_{data} и распределением модели P_{θ}

$$L_{CE}(\theta) = \mathbb{E}_{x \sim P_{data}} \left[- \sum_{n=1}^N \log P_{\theta}(x_n | x_{<n}) \right]. \quad (2)$$

Однако в задачах непрерывного обучения (continual learning) или дообучения (fine-tuning) распределение новых данных может существенно отличаться от

распределения данных, на которых модель обучалась изначально. Оптимизация только L_{CE} на новом домене приводит к смещению параметров θ в область, оптимальную для новой задачи, но не обеспечивающую требуемого качества на исходных задачах. Это явление, известное как катастрофическое забывание [3], проявляется в снижении качества модели на «старых» задачах или в деградации общезыковых способностей.

Для регуляризации процесса обучения и сохранения знаний можно ввести штраф на основе KL -дивергенции относительно некоторой референсной модели θ_{ref} :

$$L_{KL}(\theta \parallel \theta_{ref}) = \mathbb{E}_{x \sim P_{data}} \left[D_{KL} \left(P_{\theta_{ref}}(\cdot | x) \parallel P_{\theta}(\cdot | x) \right) \right] \quad (3)$$

где KL -дивергенция для двух дискретных распределений p и q определяется как:

$$D_{KL}(p \parallel q) = \sum_{y \in V} p(y) \log \frac{p(y)}{q(y)} \quad (4)$$

Полная функция потерь в таком случае:

$$L_{CE+KL}(\theta_t) = L_{CE}(\theta_t) + \beta L_{KL}(\theta \parallel \theta_{ref}), \quad (5)$$

где $\beta \geq 0$ — коэффициент регуляризации.

Существует два принципиально различных способа выбора θ_{ref} .

- **Фиксированная модель-учитель:** $\theta_{ref} = \theta_{init}$, где θ_{init} — параметры предобученной модели, которые замораживаются и не меняются в процессе дообучения.

- **Динамическая модель-учитель (ТЕМА):** параметры θ_{ref} (обозначим их через $\theta^{teacher}$) обновляются в процессе обучения как экспоненциальное скользящее среднее (ЕМА) параметров обучаемой модели ("студента").

В предлагаемом методе ТЕМА (Teacher Exponential Moving Average) параметры учителя на шаге t вычисляются следующим образом:

$$\theta_t^{teacher} = \tau \theta_{t-1}^{teacher} + (1 - \tau) \theta_t, \tau \in (0, 1), \quad (6)$$

где τ — коэффициент сглаживания.

Таким образом, учитель усредняет историю обновлений студента и служит адаптивным референсом.

$$L_{ТЕМА}(\theta_t) = L_{CE}(\theta_t) + \beta L_{KL}(\theta \parallel \theta_{ref}). \quad (7)$$

III. МЕТОДОЛОГИЯ ИССЛЕДОВАНИЯ

Для проведения экспериментов были выбраны две компактные модели семейства Qwen2, предоставленные в библиотеке Unsloth [17] в формате bnb-4bit:

- Qwen2-1.5B-bnb-4bit,
- Qwen2-0.5B-bnb-4bit.

В работе использовались модели, предоставленные библиотекой Unsloth в оптимизированном 4-битном

формате. Выбор этого фреймворка обусловлен его высокой эффективностью в ускорении обучения LLM за счет оптимизации вычислительных процессов. Ключевым фактором стала полная поддержка методов параметрически-эффективной настройки (PEFT), таких как LoRA и QLoRA, что является критически важным для задач непрерывного обучения. Данный подход позволил достичь необходимый баланс между производительностью и потреблением ресурсов, обеспечивая высокое качество генерации даже в условиях аппаратных ограничений. Эксперименты проводились на инфраструктуре сайта Kaggle с использованием ускорителя GPU T4.

А. Обучающие данные

В качестве исходных данных использовался корпус медицинских текстов, основанный на системе классификации Bethesda (The Bethesda System for Reporting Thyroid Cytopathology, TBSRTC) [15, 16]. Эта система применяется для описания результатов тонкоигольной аспирационной биопсии (ТАБ) узлов щитовидной железы и включает шесть диагностических категорий, каждая из которых характеризуется различным риском злокачественности (от 4 до 97%). Указанные категории определяют тактику ведения пациента и имеют важное клиническое значение.

Датасет включает более 27 000 реальных медицинских записей (пары текстов «описание» и «заключение»), собранных в лаборатории цитологии и цитогенетики отдела патоморфологии НМИЦ эндокринологии имени академика И.И. Дедова Минздрава России в период 2013–2023 г.г.

Ключевая особенность корпуса заключается в низкой структурированности текстов:

- поля «описание» и «заключение» частично пересекаются,
- отсутствуют единые шаблоны оформления,
- содержатся медицинские термины, сокращения и нестандартные символы.

Это делает задачу очистки данных нетривиальной.

Для экспериментов использовалась подвыборка из 1000 строк. Пример данных приведён в таблице 1.

Табл. 1 Пример обучающих данных

Описание	Заключение
П 20 (№7053/22 1). Пунктирован коллоидный зоб с морфологическими признаками регрессивных изменений и небольшим содержанием в мазке дистрофичных тироцитов на фоне коллоида и измененных эритроцитов.	Доброкачественные образования обеих долей в соответствии с критериями системы классификации Bethesda - II диагностическая категория.

В. Методы дообучения

Для адаптации моделей были использованы LoRA- (Low-Rank Adaptation) адаптеры, которые позволяют обучать только низкоранговые матрицы параметров, накладываемые на слои трансформера. Это обеспечивает значительное снижение числа

обучаемых параметров и уменьшение потребления ресурсов.

Обучение моделей проводилось с помощью касса SFTTrainer из библиотеки TRL (Transformers Reinforcement Learning). Были дообучены 4 модели:

- 1) модель с функцией потерь кроссэнтропия (CE). Реализация – SFTTrainer;
- 2) модель с функцией потерь кроссэнтропия + KL-дивергенция с фиксированной моделью учителем (KL-CE). Реализация – модифицированный SFTTrainer (код взят из открытого гитхаб-репозитория [18]);
- 3) модель с функцией потерь кроссэнтропия + KL-дивергенция с моделью-учителем, обновляемой по формуле $\text{ema} - \text{exponential moving average}$ (TEMA). Реализация – модифицированный SFTTrainer (кастомная реализация, код находится на гитхаб репозитории [19]);
- 4) модель с функцией потерь кроссэнтропия и регуляризацией L1 (CE-L1). Реализация – модифицированный SFTTrainer (кастомная реализация, код расположен на гитхаб [19]).

1. Метод с фиксированным «учителем» (KL-CE)

В модификации использовался фиксированный «учитель» — модель без LoRA-адаптеров. Функция потерь имеет вид (5) с параметром $\beta=0.05$.

Алгоритм метода:

Входные данные: тренировочный датасет \mathcal{D} , параметры базовой модели θ_{base} , скорость обучения η , коэффициент β

Инициализация параметров модели-студента

$$\theta \leftarrow \theta_{base} + \theta_{LoRA}$$

Фиксация параметров модели-учителя

$$\theta_{ref} \leftarrow \theta_{base}$$

Цикл обучения:

Для каждого батча (x, y) в \mathcal{D} :

Прямой проход модели-студента

$$z_{stu} \leftarrow f(x; \theta),$$

$$\mathcal{L}_{SFT} \leftarrow \text{CrossEntropy}(z_{stu}, y)$$

Прямой проход модели-учителя (без вычисления градиентов)

$$z_{ref} \leftarrow f(x; \theta_{ref})$$

Вычисление функции потерь и обновление параметров

$$\mathcal{L}_{KL} \leftarrow D_{KL}(\text{Softmax}(z_{ref}) \parallel \text{LogSoftmax}(z_{stu}))$$

$$\mathcal{L} \leftarrow \mathcal{L}_{SFT} + \beta \cdot \mathcal{L}_{KL}$$

$$\theta \leftarrow \theta - \eta \nabla_{\theta} \mathcal{L}$$

Конец цикла

2. Метод ТЕМА (Teacher Exponential Moving Average)

Предлагаемый метод представляет собой модификацию метода дистилляции знаний в рамках концепции Mean Teacher, специально адаптированную для параметрически-эффективной настройки (PEFT) больших языковых моделей. Хранение весов учителя осуществляется не путем дублирования полного экземпляра нейронной сети, а через создание изолированного буфера (словаря тензоров), содержащего копии исключительно

обучаемых параметров, то есть матриц LoRA. Базовые веса языковой модели остаются общими и неизменяемыми для обоих агентов обучения. Обновление параметров учителя производится непосредственно в выделенной области памяти (in-place) после каждого шага оптимизации путем линейной интерполяции текущих значений с весами модели-ученика, что исключает дополнительные затраты на аллокацию памяти в процессе итераций.

По сравнению со стандартным методом обучения с учителем (реализация SFTTrainer) предлагаемый алгоритм демонстрирует минимальный накладной расход памяти. В то время как стандартный процесс SFT требует видеопамати для хранения базовой модели, активных адаптеров, градиентов и состояний оптимизатора, метод ТЕМА добавляет потребность в хранении лишь одной дополнительной копии весов адаптеров. Учитывая, что количество параметров LoRA, как правило, составляет доли 1% от общего объема параметров большой языковой модели, дополнительный объем занимаемой памяти является незначительным в процентном соотношении. Это позволяет применять метод ТЕМА на том же вычислительном оборудовании, которое используется для стандартного дообучения.

Алгоритм метода:

Входные данные: тренировочный датасет \mathcal{D} , параметры замороженной базовой модели θ_{base} , параметры адаптеров LoRA ϕ , скорость обучения η , коэффициенты функции потерь β , коэффициент сглаживания μ .

Инициализация весов студента:

$$\theta_{stu} = \{\theta_{base}, \phi\}$$

Инициализация весов учителя (EMA):

$$\phi_{ema} \leftarrow \phi$$

Цикл обучения:

Для каждого батча (x, y) в \mathcal{D} :

Прямой проход модели-студента

$$z_{stu} \leftarrow f(x; \theta_{base}, \phi)$$

$$\mathcal{L}_{SFT} \leftarrow \text{CrossEntropy}(z_{stu}, y)$$

Прямой проход модели-учителя (без вычисления градиентов)

Используются веса базовой модели и EMA-адаптеры

$$z_{tea} \leftarrow f(x; \theta_{base}, \phi_{ema})$$

Вычисление функции потерь и обновление параметров

$$\mathcal{L}_{KL} \leftarrow D_{KL}(\text{Softmax}(z_{tea}) \parallel$$

$\text{LogSoftmax}(z_{stu}))$

$$\mathcal{L} \leftarrow \mathcal{L}_{SFT} + \beta \cdot \mathcal{L}_{KL}$$

Обновление только параметров LoRA

$$\phi \leftarrow \phi - \eta \nabla_{\phi} \mathcal{L}$$

Обновление учителя (EMA)

Сглаживание только параметров LoRA

$$\phi_{ema} \leftarrow \mu \cdot \phi_{ema} + (1 - \mu) \cdot \phi$$

Конец цикла

Правило обновления параметров модели – учителя и функция потерь соответствуют формулам (6) и (7) соответственно. Параметры: $\beta = 0.05$ и $\tau = 0.999$

3. Метод с L1-регуляризацией (CE-L1)

В этой модификации к стандартной функции потерь (Cross-Entropy) добавляется слагаемое для регуляризации — L1-норма обучаемых параметров, т.е. весов LoRA.

Входные данные: тренировочный датасет D , параметры базовой модели θ_{base} , скорость обучения η , коэффициент регуляризации β .

Инициализация параметров модели

$$\theta \leftarrow \theta_{base} + \theta_{LoRA}$$

Цикл обучения:

Для каждого батча (x, y) в D :

Прямой проход модели

$$z_{stu} \leftarrow f(x; \theta)$$

$$\mathcal{L}_{SFT} \leftarrow \text{CrossEntropy}(z_{stu}, y)$$

Вычисление регуляризации (только для параметров θ_{LoRA})

$$\mathcal{L}_{L1} \leftarrow \sum |\theta_{LoRA}|$$

Вычисление функции потерь и обновление параметров

$$\mathcal{L} \leftarrow \mathcal{L}_{SFT} + \beta \cdot \mathcal{L}_{L1}$$

$$\theta \leftarrow \theta - \eta \nabla_{\theta} \mathcal{L}$$

Конец цикла

С. Параметры дообучения

Тонкая настройка моделей производилась с использованием метода эффективной адаптации LoRA (ранг $r = 8$, коэффициент масштабирования $\alpha = 16$). Процесс обучения выполнялся в течение 4 эпох под управлением оптимизатора AdamW с коэффициентом регуляризации весов weight decay, равным 0,01. Начальная скорость обучения составляла $2 \cdot 10^{-4}$ с применением стратегии линейного спада. Максимальная длина входной последовательности была ограничена 1024 токенами. Выбор гиперпараметров регуляризации не является однозначным и зависит от требований к балансу между стабильностью и пластичностью решения конкретной задачи. Чем выше параметр регуляризации, тем больше баланс смещен в сторону стабильности, т.е. сгенерированные тексты будут ближе по содержанию к изначальной модели. Анализ влияния гиперпараметров на процесс тонкой настройки моделей представлен ниже в разделе V.

Д. Методы оценки качества генерации

Оценка качества автоматически сгенерированных текстов представляет собой нетривиальную задачу, особенно в медицинской сфере, где даже при использовании различных формулировок критически важно сохранить точность, полноту и достоверность врачебных заключений. На сегодняшний день не существует единой универсальной метрики, позволяющей в полной мере отразить качество медицинских текстов. В связи с этим для анализа применяются несколько взаимодополняющих метрик, таких как BLEU [20], ROUGE [21], METEOR [22], ChrF [23], BERTScore [24], BLEURT [25]. Каждая из них по-разному отражает степень сходства между автоматически сгенерированным и эталонным заключениями.

BLEU (Bilingual Evaluation Understudy) — метрика, оценивающая качество машинного перевода через сравнение n -грамм в сгенерированном и эталонном текстах [20]

Для расчёта метрики BLEU сначала определяют модифицированные показатели точности (precision) для n -грамм в тестовом корпусе, после чего вычисляют их геометрическое среднее и умножают на экспоненциальный штраф за сокращение длины перевода. Перед оценкой текста все символы приводятся к нижнему регистру. Формально, полагая p_n модифицированной точностью для n -грамм длины от 1 до N с весами w_n (такими, что $\sum_{n=1}^N w_n = 1$), вычисляют P :

$$P = \exp \left(\sum_{n=1}^N w_n \log p_n \right). \quad (8)$$

Далее, обозначив длину проверяемого текста перевода через c и «эффективную» длину эталонного корпуса через r , рассчитывают коэффициент штрафа BP :

$$BP = \begin{cases} 1, & \text{если } c > r \\ e^{(1-\frac{r}{c})}, & \text{если } c \leq r \end{cases}. \quad (9)$$

Тогда

$$BLEU = BP \times P.$$

Метрика BLEU принимает значения от 0 до 1. Чем ближе значение метрики к 1, тем более похож сгенерированный текст на эталонный.

ROUGE-1 [21], ROUGE-2 [21], ROUGE-L [21] (Recall-Oriented Understudy for Gisting Evaluation) — набор метрик, оценивающих покрытие n -грамм эталонного текста в сгенерированном. Эти методы подсчитывают число совпадающих элементов — n -грамм, последовательностей слов и пар слов — между оцениваемым автоматически сгенерированным резюме и эталонными вариантами, созданными ранее людьми.

ROUGE-N вычисляется по формуле:

$$\text{ROUGE-N} = \frac{\sum_{S \in \{\text{ReferenceSummaries}\}} \sum_{gram_n \in S} \text{Count}_{\text{match}}(gram_n)}{\sum_{S \in \{\text{ReferenceSummaries}\}} \sum_{gram_n \in S} \text{Count}(gram_n)}, \quad (11)$$

где n обозначает длину n -граммы $gram_n$, а $\text{Count}_{\text{match}}(gram_n)$ — максимальное число n -грамм, совпадающих в оцениваемом реферате и множестве эталонных рефератов. ROUGE-N является мерой, основанной на полноте (recall), поскольку в знаменателе формулы стоит общее количество n -грамм в эталонных рефератах. Похожая метрика BLEU, применяемая для автоматической оценки качества машинного перевода, ориентирована на точность (precision) [21].

METEOR (Metric for Evaluation of Translation with Explicit ORdering) — метрика автоматической оценки качества перевода, основанная на явном сопоставлении слов между сгенерированным и эталонным текстами, учитывающая не только точные совпадения, но и морфологические вариации и синонимии [22].

В отличие от BLEU, METEOR рассчитывает гармоническое среднее между точностью и полнотой, при этом полноте (recall) придается значительно больший вес. Для расчёта сначала выполняется выравнивание униграмм в три этапа: точное совпадение, совпадение основ слов и совпадение синонимов. Формально, обозначив количество сопоставленных униграмм через m , общую длину сгенерированного перевода через c , а длину эталонного перевода через r , рассчитывают точность $P = \frac{m}{c}$ и полноту $R = \frac{m}{r}$. Взвешенное гармоническое среднее F_{mean} вычисляется по формуле:

$$F_{mean} = \frac{10PR}{R + 9P} \quad (12)$$

Для оценки порядка слов и связности текста вводится штраф (Penalty). Сопоставленные слова группируются в наименьшее возможное количество непрерывных последовательностей (чанков), идущих в одинаковом порядке в обоих текстах. Обозначив количество таких чанков через ch , рассчитывают коэффициент штрафа Pen :

$$Pen = 0.5 \times \left(\frac{ch}{m}\right)^3 \quad (13)$$

Тогда $METEOR = F_{mean} \times (1 - Pen)$. (14)

Метрика METEOR принимает значения от 0 до 1. Более высокое значение указывает на лучшее качество перевода.

Метрика CHRF++ [23] (Character n-gram F-score with words) оценивает качество машинного перевода на основе F-меры для символьных n -грамм с добавлением униграмм и биграмм слов [23]. Для расчёта сначала вычисляют точность (Precision) и полноту (Recall) символьных n -грамм и n -грамм слов, усредняя их арифметически. Формально, пусть $ChrP$ и $ChrR$ — усредненные точность и полнота для символьных n -грамм, а $WrdP$ и $WrdR$ — усредненные точность и полнота для словных n -грамм. Итоговые показатели Precision (P_{comb}) и Recall (R_{comb}) для CHRF++ рассчитываются как среднее арифметическое символьных и словных компонентов:

$$P_{comb} = \frac{ChrP + WrdP}{2} \quad (15)$$

$$R_{comb} = \frac{ChrR + WrdR}{2} \quad (16)$$

Затем вычисляется итоговое значение CHRF++ как F-мера с параметром β :

$$CHRF++ = (1 + \beta^2) \frac{P_{comb} \cdot R_{comb}}{\beta^2 \cdot P_{comb} + R_{comb}} \quad (17)$$

Метрика CHRF++ принимает значения от 0 до 1 (или иногда от 0 до 100 в процентах). Она сочетает преимущества символьного (устойчивость к морфологии) и словного (лексическое соответствие) подходов.

Метрика BERTScore используется для автоматической оценки качества генерации текста.

Она основана на использовании контекстуальных эмбедингов (например, BERT, RoBERTa) для вычисления сходства между токенами сгенерированного и эталонного предложений [24].

В отличие от n -граммных метрик, BERTScore сопоставляет каждый токен сгенерированного предложения с наиболее похожим токеном из эталона, используя косинусное сходство их векторных представлений.

Формально, для эталонного $x = \langle x_1, \dots, x_k \rangle$ и сгенерированного $\hat{x} = \langle \hat{x}_1, \dots, \hat{x}_l \rangle$ предложений сначала вычисляются эмбединги токенов. Затем рассчитываются точность (P_{BERT}), полнота (R_{BERT}) и F-мера (F_{BERT}):

$$R_{BERT} = \frac{1}{|x|} \sum_{x_i \in x} \max_{\hat{x}_j \in \hat{x}} x_i^T \hat{x}_j \quad (18)$$

$$P_{BERT} = \frac{1}{|\hat{x}|} \sum_{\hat{x}_j \in \hat{x}} \max_{x_i \in x} x_i^T \hat{x}_j \quad (19)$$

$$F_{BERT} = 2 \frac{P_{BERT} \cdot R_{BERT}}{P_{BERT} + R_{BERT}} \quad (20)$$

где x_i и \hat{x}_j — предварительно нормализованные векторы эмбедингов токенов.

Метрика BLEURT (Bilingual Evaluation Understudy with Representations from Transformers) — автоматически оценивает качество текста, и обучается на базе модели BERT для предсказания человеческих оценок качества перевода [25].

В основе BLEURT лежит предварительно обученная модель BERT, которая дообучается в два этапа. Сначала модель проходит предварительное обучение на больших объемах синтетических данных (пары предложений с различными видами искажений), чтобы научиться распознавать широкий спектр лексических и семантических различий. Затем модель дообучается на реальных данных с человеческими оценками качества.

Формально, BLEURT принимает на вход пару предложений: эталонное предложение x и сгенерированное предложение \hat{x} . Эти предложения объединяются в одну последовательность и подаются на вход модели BERT y :

$$y = f(BERT(x, \hat{x})), \quad (21)$$

где f — линейный слой регрессии.

BLEURT способен учитывать семантические нюансы, перефразирования и различные типы ошибок, благодаря использованию контекстуальных представлений BERT и обучению на разнообразных данных [26].

Таким образом, каждая из рассмотренных метрик отражает различные аспекты качества текста: от формального совпадения на уровне слов и символов (BLEU, ROUGE, ChrF) до оценки смысловой эквивалентности (BERTScore, BLEURT). Использование комбинации этих метрик позволит получить всестороннюю и объективную картину качества генерации медицинских заключений, хотя,

конечно, ни одна из этих метрик пока не может сравниться с экспертизой специалиста.

IV. АНАЛИЗ РЕЗУЛЬТАТОВ

А. Анализ метрик качества генерации заключений

В настоящей работе были обучены 4 варианта моделей с различными функциями потерь на корпусе заключений по результатам тонкоигольной биопсии. Использовался стандартный промпт вида `### Input: {описание} ### Output: {заключение}`. Пример промпта представлен в Приложении А. Валидация проводилась на тестовой выборке из 1000 строк.

Результаты оценивания моделей на тестовой выборке приведены в таблицах 1 и 2

Табл.1. Для модели Qwen2-1.5B

Метрика	CE	KL-CE	TEMA	CE-L1
BLEU	0.858	0.849	0.856	0.854
ROUGE-1	0.917	0.913	0.913	0.912
ROUGE-2	0.865	0.863	0.864	0.861
ROUGE-L	0.917	0.913	0.913	0.912
METEOR	0.926	0.925	0.933	0.915
ChrF	93.137	93.666	94.183	92.365
BLEURT	0.904	0.921	0.925	0.897
BERTScoreF1	0.954	0.963	0.965	0.953

Табл.2. Для модели Qwen2-0.5B

Метрика	CE	KL-CE	TEMA	CE-L1
BLEU	0.849	0.827	0.834	0.841
ROUGE-1	0.907	0.886	0.900	0.902
ROUGE-2	0.859	0.834	0.850	0.854
ROUGE-L	0.907	0.886	0.900	0.902
METEOR	0.915	0.900	0.915	0.914
ChrF	92.366	91.198	92.558	92.330
BLEURT	0.908	0.897	0.910	0.910
BERTScoreF1	0.945	0.934	0.953	0.946

Таблицы 1 и 2 демонстрируют, что результаты варьируются в зависимости от выбора функции потерь.

Для модели Qwen2-1.5B:

- чистая CE-модель показывает наилучшие результаты по *символьным метрикам* (ROUGE, ChrF), что отражает её склонность воспроизводить формулировки максимально близко к исходному тексту;
- ТЕМА показывает лучший результат на *семантических метриках* (BERTScore, BLEURT, METEOR), что указывает на более адекватное

захватывание смысла текста, даже при меньшем совпадении на уровне символов;

- KL-CE показывает результаты хуже по большинству метрик, по сравнению с CE и ТЕМА — это объясняется тем, что фиксированный «учитель» имеет ограниченные способности, и жёсткое следование его распределению мешает обучению.
- CE-L1 показывает наихудшие результаты среди всех моделей

Для модели Qwen2-0.5B картина схожая: CE модель лучше на символьных метриках, а ТЕМА - на семантических. При этом модель KL-CE показала наихудшие результаты.

Проведенная оценка показывает, что модель с динамическим учителем (ТЕМА) превосходит модель со статическим учителем (KL-CE) по всем рассматриваемым метрикам. При этом результаты ТЕМА являются близкими к модели CE, а по метрикам, оценивающим семантическое сходство, — лучше. Полученные данные свидетельствуют о высокой адаптивной способности предлагаемой модели «динамический учитель – ученик» и ее эффективности при обучении на новых данных.

В. Анализ устойчивости и «забывания» на MMLU

Проверка сохранения общих знаний после дообучения на медицинском корпусе была проведена на бенчмарке MMLU в формате 5-shot, т.е. использовался промпт с инструкцией на русском языке, включающий пять примеров вопросов с ответами и тестовый вопрос в формате множественного выбора. Пример промпта представлен в Приложении А.

Валидационный датасет был сформирован на основе данных официального репозитория [27]. Тестирование проводилось на выборках объемом 100 записей для каждой отдельной области знаний. Для сводной секции «ALL» выборка была расширена до 1000 записей с целью повышения полноты оценки и обеспечения всестороннего покрытия предметных доменов. Результаты экспериментов приведены в таблицах 3 и 4.

Табл.3. Результаты MMLU модели Qwen2-0.5B

Область знаний	Модель	Точность	Средн. время генерации, с/вопрос
clinical_knowledge	KL_CE	0.35	0.61
	CE	0.27	1.06
	CE_L1	0.31	0.69
	TEMA	0.31	0.23
college_biology	KL_CE	0.33	0.44
	CE	0.32	0.76
	CE_L1	0.3	0.57
	TEMA	0.32	0.30
medical_	KL_CE	0.39	0.86

genetics	CE	0.35	1.01
	CE_L1	0.31	1.27
	TEMA	0.35	0.43
virology	KL_CE	0.38	0.58
	CE	0.31	1.09
	CE_L1	0.34	0.62
	TEMA	0.4	0.24
philosophy	KL_CE	0.27	0.76
	CE	0.25	0.95
	CE_L1	0.3	1.00
	TEMA	0.3	0.30
formal_logic	KL_CE	0.28	0.94
	CE	0.02	3.00
	CE_L1	0.19	0.46
	TEMA	0.2	0.46
global_facts	KL_CE	0.31	0.22
	CE	0.32	0.24
	CE_L1	0.3	0.28
	TEMA	0.36	0.20
public_relations	KL_CE	0.41	0.45
	CE	0.38	0.79
	CE_L1	0.38	0.98
	TEMA	0.4	0.31
All	KL_CE	0.33	0.79
	CE	0.25	1.5
	CE_L1	0.23	1.1
	TEMA	0.34	0.32

genetics	CE	0.56	0.56
	CE_L1	0.55	0.80
	TEMA	0.6	0.31
virology	KL_CE	0.43	0.44
	CE	0.41	0.53
	CE_L1	0.36	1.21
	TEMA	0.43	0.44
philosophy	KL_CE	0.57	0.26
	CE	0.55	0.25
	CE_L1	0.54	0.47
	TEMA	0.61	0.25
formal_logic	KL_CE	0.42	0.82
	CE	0.31	4.53
	CE_L1	0.18	10.54
	TEMA	0.41	2.76
global_facts	KL_CE	0.25	0.93
	CE	0.24	1.74
	CE_L1	0.17	4.60
	TEMA	0.22	0.96
public_relations	KL_CE	0.6	0.33
	CE	0.59	0.70
	CE_L1	0.58	0.94
	TEMA	0.61	0.64
All	KL_CE	0.51	0.94
	CE	0.44	3.37
	CE_L1	0.43	2.1
	TEMA	0.49	1.40

Табл.4. Результаты MMLU модели Qwen2-1.5B

Область знаний	Модель	Точность	Средн. время генерации, с/вопрос
clinical_knowledge	KL_CE	0.53	0.74
	CE	0.48	0.82
	CE_L1	0.47	1.1
	TEMA	0.54	0.76
college_biology	KL_CE	0.52	0.34
	CE	0.49	0.89
	CE_L1	0.48	1.05
	TEMA	0.54	0.42
medical_	KL_CE	0.61	0.29

Анализ результатов показывает:

- CE и CE-L1-модели демонстрируют наихудшую точность и существенно большее время генерации. Дополнительно наблюдается склонность выдавать избыточные и нерелевантные ответы, что подтверждает риск «катастрофического забывания» общих знаний
- KL-CE в ряде случаев повышает точность, но в модели 0.5B значительно уступает в скорости генерации модели ТЕМА. Это указывает на то, что статический «учитель» хоть и снижает забывание, но при этом ограничивает адаптацию к генерации согласно промпту.
- ТЕМА достигает наиболее сбалансированных результатов: точность выше или сопоставима с KL-CE, а время генерации значительно меньше. Это подтверждает гипотезу, что динамический

ЕМА-учитель помогает стабилизировать процесс обучения, при этом сохраняя базовые знания.

Также был сделан акцент на скорость генерации ответа, т.к. было замечено, что кроме ответа на вопрос, модель иногда генерирует текст, не связанный с вопросом, что напрямую указывает на качество модели.

Например, на вопрос:
 “Which of the following best describes the structure that collects urine in the body?”

Модель Qwen2-1.5B с CE функцией потерь сгенерировала ответ:

“А узловое образование цитовидной железы с реактивными изменениями и элементами воспалительной инфильтрации подозрительное в отношении злокачественного новообразования в соответствии с критериями системы классификации Bethesda - V диагностическая категория.”

Таким образом, время генерации ответа моделью характеризует также его качество, т.к. объемный ответ во-первых, противоречит промпту, и во-вторых, скорее всего связан с генерацией не связанного с вопросом текста.

С. Качественный анализ примеров

Для наглядной демонстрации влияния функции потерь на сохранение общих языковых способностей модели был проведен качественный анализ генерации текста. Ниже приведен характерный пример, иллюстрирующий феномен катастрофического забывания и эффективность методов регуляризации. Проверка метода CE-L1 не была включена в список проверок из-за ее низких показателей в предыдущих экспериментах. Базовая модель – Qwen2-1.5B.

Запрос к моделям:

“Мозг - центр управления всем, что мы делаем, чувствуем и думаем, нужен для”

Ответы моделей на этот запрос представлены в таблице 5.

Сравнительный анализ показывает, что дообучение без регуляризации (CE) приводит к сильному смещению распределения вероятностей при генерации токенов в сторону доменной лексики, характерной для обучающих данных, что ухудшает качество генерации в смежном медицинском домене.

Табл. 5 Тексты ответов, сгенерированных разными моделями

Модель	Сгенерированный текст
BASE	“того, чтобы мы могли двигаться, чувствовать, думать и чувствовать. Он также является центром, который управляет нашим поведением, эмоциями и поведением. Он также является центром, который управляет нашим поведением, эмоциями и поведением. Он также является

	центром, который управляет нашим поведением, эмоциями и поведением. Он также является центром, который управляет нашим поведением, эмоциями и поведением. Он также является центром, который управ”
CE	“того, чтобы мы могли жить и работать в мире. В мозге содержится очень много клеток и неравномерной толщины. Клетки образуют многослойные структуры и неравномерно распределенные клеточные группы. В мазке цитологическая картина характерная для коллоидного в разной степени пролиферирующего зоба с аденоматозомпунктированного в зоне кистозной дегенерации. В мазке на фон”
KL-CE	“того, чтобы мы могли бы чувствовать себя комфортно в своей жизни. Правильное питание и регулярные физические упражнения могут помочь улучшить состояние мозга и уменьшить риск возникновения узлового коллоидного зоба.”
ТЕМА	“того, чтобы мы могли двигаться по жизни. Многие из наших мыслей и чувств не связаны с тем, что происходит в нашем мозге. В мозге происходят процессы, которые мы не видим и не чувствуем. В мозге происходит многое, что мы не видим и не чувствуем. В мозге происходит многое, что мы не видим и не чувствуем. В мозге происходит многое, что мы не видим и не чувствуем. В мозге происходит многое, что мы не видим”

В то же время, методы регуляризации, и особенно метод ТЕМА, эффективно сдерживают смещение параметров. Предлагаемый метод позволяет модели сохранять баланс между адаптацией и удержанием общих знаний, предотвращая генерацию нерелевантных доменных терминов в общем контексте. Дополнительные примеры генерации, подтверждающие эти наблюдения, приведены в Приложении А.

V. АНАЛИЗ ВЛИЯНИЯ ГИПЕРПАРАМЕТРОВ РЕГУЛЯРИЗАЦИИ НА ПРОЦЕСС ТОНКОЙ НАСТРОЙКИ

В этом разделе исследуется влияние гиперпараметров предложенного метода ТЕМА на динамику обучения и свойства полученных моделей. В частности, анализируются коэффициент регуляризации β , который отвечает за вклад функции потерь KL-дивергенции и коэффициент затухания τ для экспоненциального скользящего среднего (ЕМА) весов учителя. В качестве метрик оценки использовались L2-норма изменения весов адаптеров

LoRA относительно их начального состояния (отражающая степень пластичности и величину обновления параметров) и Centered Kernel Alignment (СКА) [28], характеризующая степень сохранения внутренних представлений обучаемой модели по сравнению с базовой (стабильность).

А. Влияние коэффициента регуляризации β

На рисунке 1 представлены зависимости L2-нормы изменений весов и СКА от шага обучения при фиксированном $\tau = 0.999$ и варьировании β .

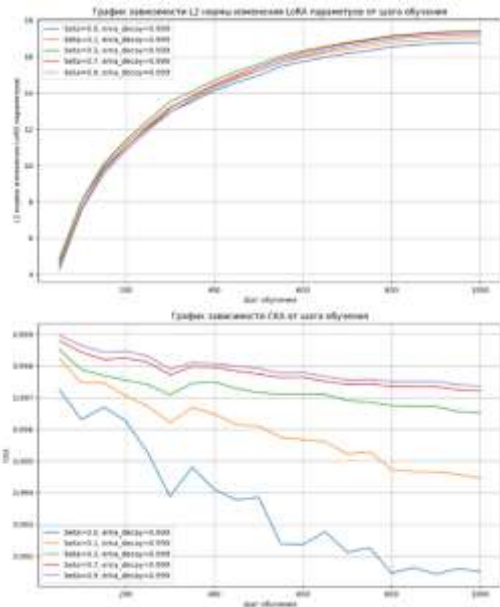


Рисунок 1 – Графики изменения метрик L2 нормы LoRA весов и СКА от шага обучения при изменении β

1. **L2-норма изменений весов:** как видно из верхнего графика рисунка 1, все кривые демонстрируют монотонный рост нормы весов, что свидетельствует о непрерывном процессе адаптации к новым данным. Примечательно, что увеличение коэффициента β (от 0.0 до 0.9) приводит к *незначительному увеличению* L2-нормы весов. Такое поведение можно объяснить более сложным ландшафтом функции потерь, т.к. использование динамического модели-учителя требует нахождения минимума не только на новой задаче (CrossEntropyLoss), но и на выполненной моделью-учителем задаче (KL-дивергенция), что может требовать более сложной траектории оптимизации и, как следствие, большей длины пути в пространстве параметров.

2. **СКА (Centered Kernel Alignment):** нижний график рисунка 1 демонстрирует ключевой эффект регуляризации. Модель без регуляризации ($\beta = 0.0$, синяя линия) показывает наиболее резкое падение СКА, что указывает на значительное изменение внутренних представлений и высокий риск катастрофического забывания. С увеличением β наблюдается систематическое увеличение значений СКА. При $\beta = 0.9$ (фиолетовая линия) СКА остается на уровне выше 0.997 на протяжении всего обучения, что свидетельствует о высокой

степени сохранения исходных знаний. Таким образом мы видим, что параметр β напрямую управляет балансом между адаптацией к новому домену и сохранением исходных репрезентаций.

В. Влияние коэффициента затухания ЕМА τ

На рисунке 2 показаны зависимости метрик при фиксированном $\beta = 0.05$ и варьировании τ .

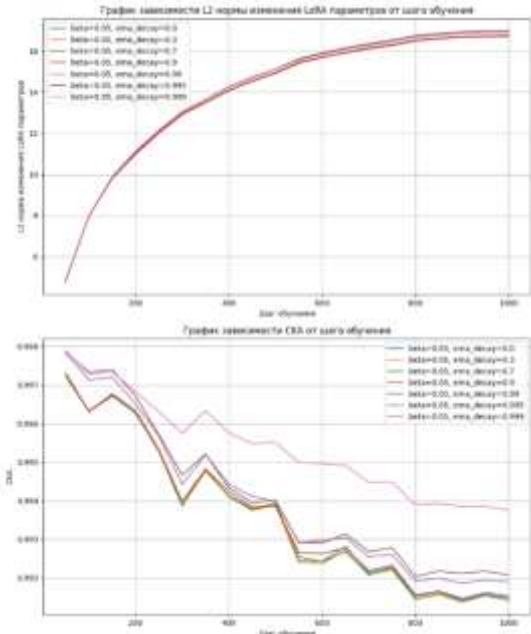


Рисунок 2 - Графики изменения метрик L2 нормы LoRA весов и СКА от шага обучения при изменении τ

1. **L2-норма изменений весов:** из верхнего графика рисунка 2 следует, что параметр τ оказывает минимальное влияние на величину изменения весов. Все кривые практически совпадают, что говорит о том, что динамика "ухода" весов от начальной точки определяется в основном скоростью обучения и функцией потерь Cross-Entropy и параметром регуляризации β , а не инерцией учителя.

2. **СКА:** Нижний график рисунка 2 выявляет критическую роль τ в стабилизации обучения. При малых значениях τ (быстрое обновление учителя, линии 0.0–0.9) СКА падает аналогично слабо регуляризованным моделям. Однако при приближении τ к 1 (0.99, 0.995, 0.999) наблюдается качественный скачок в стабильности представлений. Линия $\tau = 0.999$ (розовая) демонстрирует наивысшие значения СКА, значительно превосходя остальные конфигурации.

С. Подбор параметров

Проведенный анализ позволяет заключить, что метод ТЕМА эффективно управляет компромиссом между пластичностью и стабильностью. Коэффициент β отвечает за "силу" удержания студента около учителя, напрямую влияя не только на сохранение представлений (СКА), но и на изменение весов при адаптации к новой задаче. Коэффициент τ определяет инерцию самого учителя:

высокие значения (≈ 0.999) необходимы для формирования стабильного и надежного ориентира, предотвращающего дрейф представлений.

Оптимальной конфигурацией для задач, требующих минимизации забывания, является сочетание высокого $\tau(0.999)$ и умеренного $\beta(0.05-0.3)$, что обеспечивает высокие значения СКА при сохранении способности к обучению.

ЗАКЛЮЧЕНИЕ

В рамках настоящего исследования проведен сравнительный анализ стратегий дообучения компактных больших языковых моделей (на примере Qwen2-0.5B и Qwen2-1.5B с применением 4-битной квантизации и адаптеров LoRA) на материале специализированного медицинского корпуса. Объектом исследования выступили 4 подхода: стандартное обучение с учителем на основе кросс-энтропии (CE), метод с L1 регуляризацией (CE-L1), метод с регуляризацией относительно фиксированной базовой модели (KL-CE) и предлагаемый метод ТЕМА (Teacher Exponential Moving Average), использующий динамическое обновление параметров модели-«учителя».

Анализ экспериментальных данных позволил сформулировать следующие выводы:

1. Метод KL-CE продемонстрировал сниженную эффективность адаптации к новому домену. Фиксация весов «учителя» накладывает жесткие ограничения на процесс обучения, так как распределение вероятностей базовой модели обладает низкой релевантностью для специфической медицинской терминологии и задач, что препятствует полноценному усвоению новой информации.

2. Метод CE-L1 показал низкое качество как при генерации новых данных, на которых происходило дообучение, так и на данных MMLU.

3. Модели, обученные с использованием только функции потерь CE, показали высокие результаты по метрикам лексического сходства (BLEU, ROUGE, ChrF). Однако это сопровождалось значительной деградацией обобщающих способностей (эффект катастрофического забывания), что подтверждается падением результатов на бенчмарке MMLU. Это указывает на тенденцию модели к переобучению на поверхностных статистических паттернах целевого домена в ущерб фундаментальным знаниям.

4. Метод ТЕМА продемонстрировал лучший баланс между пластичностью (способностью к обучению) и стабильностью (сохранением знаний). Модель, обученная этим методом, превзошла альтернативные подходы по семантическим метрикам качества генерации (BLEURT, BERTScore) при минимальном снижении показателей на общезыковых тестах.

5. Анализ влияния гиперпараметров на процесс дообучения моделей показал стабильность обучающего метода при различных значениях гиперпараметров, а также оптимальную

конфигурацию параметров, обеспечивающую баланс между пластичностью и стабильностью: $\tau = 0.999$, $\beta = (0.05-0.3)$.

Таким образом, результаты подтверждают, что использование экспоненциально сглаженного «учителя» позволяет эффективно интегрировать доменную экспертизу в модель, минимизируя риск разрушения её общих когнитивных способностей. Метод ТЕМА может быть успешно применен для адаптации LLM в высокоспециализированных областях, включая медицину, в которых критически важны как точность профильной генерации, так и сохранение широкого общего контекста.

БЛАГОДАРНОСТИ

Авторы выражают благодарность Высшей инженеринговой школе НИЯУ МИФИ за помощь в возможности опубликовать результаты выполненной работы и руководству ФГБУ «НМИЦ эндокринологии им. академика И.И. Дедова» Минздрава России за предоставленные текстовые данные.

ИСТОЧНИКИ ФИНАНСИРОВАНИЯ

Текстовые данные для проведения исследования подготовлены по гранту Российского научного фонда в рамках реализации проекта №22-15-00135 «Научное обоснование, разработка и внедрение новых технологий диагностики коморбидных йододефицитных и аутоиммунных заболеваний щитовидной железы с использованием возможностей искусственного интеллекта»

БИБЛИОГРАФИЯ

- [1] Raffel, C. et al. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. JMLR, 2020.
- [2] [Liu, P. et al. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. ACM CSUR, 2023.
- [3] Y. Luo, Z. Yang, F. Meng, Y. Li, J. Zhou, and Y. Zhang, "An empirical study of catastrophic forgetting in large language models during continual fine-tuning," arXiv preprint arXiv:2308.08747, 2023.
- [4] Kirkpatrick, J., Pascanu, R., Rabinowitz, N. C., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., Hassabis, D., Clopath, C., Kumaran, D., and Hadsell, R. Overcoming catastrophic forgetting in neural networks. Proceedings of the National Academy of Sciences, 114:3521 – 3526, 2016.
- [5] Zhicheng Wang, Yufang Liu, Tao Ji, Xiaoling Wang, Yuanbin Wu, Congcong Jiang, Ye Chao, Zhencong Han, Ling Wang, Xu Shao, et al. 2023. Rehearsalfree continual language learning via efficient parameter isolation. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 10933–10946.
- [6] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen. Lora: Low-rank adaptation of large language models. arXiv preprint arXiv:2106.09685, 2021
- [7] Zhang, Y., Jiang, S., Zhao, M., Li, Y., Fan, Y., Wu, X., & Chen, Q. (2025). Gere: Towards efficient anti-forgetting in continual learning of llm via general samples replay. *arXiv preprint arXiv:2508.04676*.
- [8] Sanyal, S., Prairie, H., Das, R., Kavis, A., & Sanghavi, S. (2025). Upweighting easy samples in fine-tuning mitigates forgetting. *arXiv preprint arXiv:2502.02797*.
- [9] Zhizhong Li and Derek Hoiem. 2017. Learning without forgetting. IEEE transactions on pattern analysis and machine intelligence, 40(12):2935–2947.
- [10] Y. Chen, S. Zhang, G. Qi, and X. Guo. Parameterizing context: Unleashing the power of parameter-efficient fine-tuning and in-

- context tuning for continual table semantic parsing. *Advances in Neural Information Processing Systems*, 36, 2024.
- [11] Tarvainen, A., & Valpola, H. Mean Teachers are Better Role Models: Weight-Averaged Consistency Targets Improve Semi-Supervised Deep Learning Results. *NeurIPS*, 2017
- [12] Ali Behrouz, Meisam Razaviyayn, Peilin Zhong, and Vahab Mirrokni. *Nested Learning: The Illusion of Deep Learning Architecture* [Электронный ресурс] - <https://research.google/blog/introducing-nested-learning-a-new-ml-paradigm-for-continual-learning/> (22.12.2025).
- [13] Qin, Y., Qian, C., Yi, J., Chen, W., Lin, Y., Han, X., ... & Zhou, J. (2022). Exploring mode connectivity for pre-trained language models. *arXiv preprint arXiv:2210.14102*.
- [14] Ren, W., Li, X., Wang, L., Zhao, T., & Qin, W. (2024). Analyzing and reducing catastrophic forgetting in parameter efficient tuning. *arXiv preprint arXiv:2402.18865*.
- [15] Ali S, Cibas E. *The Bethesda System for Reporting Thyroid Cytopathology*. (Ali SZ, Cibas ES, eds.). Cham: Springer International Publishing; 2018. doi: <https://doi.org/10.1007/978-3-319-60570-8>
- [16] Ali SZ, Baloch ZW, Cochand-Priollet B, Schmitt FC, Vielh P, VanderLaan PA. *The 2023 Bethesda System for Reporting Thyroid Cytopathology*. *Thyroid*. July 2023. doi: <https://doi.org/10.1089/thy.2023.0141>
- [17] [Электронный ресурс] - <https://unsloth.ai/>
- [18] [Электронный ресурс] - <https://github.com/BY571/sft-kl-lora-trainer>
- [19] [Электронный ресурс] - https://github.com/EugeneCS/mephi_nlp/tree/sviridenko
- [20] Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002, July). Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics* (pp. 311-318).
- [21] Chin-Yew, L. (2004). Rouge: A package for automatic evaluation of summaries. In *Proceedings of the Workshop on Text Summarization Branches Out, 2004*.
- [22] Banerjee, S., & Lavie, A. (2005, June). METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In *Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization* (pp. 65-72).
- [23] Popović, M. (2017, September). chrF++: words helping character n-grams. In *Proceedings of the second conference on machine translation* (pp. 612-618).
- [24] Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2019). BERTscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675*.
- [25] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)* (pp. 4171-4186).
- [26] Liu, Y., Maier, W., Minker, W., & Ultes, S. (2021). Naturalness evaluation of natural language generation in task-oriented dialogues using bert. *arXiv preprint arXiv:2109.02938*.
- [27] [Электронный ресурс] - <https://huggingface.co/datasets/cais/mmlu>
- [28] Kornblith, S., Norouzi, M., Lee, H., & Hinton, G. (2019, May). Similarity of neural network representations revisited. In *International conference on machine learning* (pp. 3519-3529). PMIR.

Статья получена 17 марта 2026.

Свириденко Дмитрий Константинович, Национальный Исследовательский Ядерный Университет МИФИ, магистрант dmitrii.sviridenko@yandex.ru

Боброва Елизавета Витальевна, Национальный Исследовательский Ядерный Университет МИФИ, аспирант, EVBobrova@mephi.ru

Зайцев Константин Сергеевич, Национальный Исследовательский Ядерный Университет МИФИ, профессор, email: KSZaytsev@mephi.ru

Adapting Large Language Models for narrow domains using the exponential moving average method

D.K. Sviridenko, E.V. Bobrova, K.S. Zaytsev, E.V. Dyuldin, B.M. Shifman

Abstract - Adapting Large Language Models (LLMs) to specialized subject areas requires fine-tuning on profile data corpora, which is inevitably associated with the risk of catastrophic forgetting of previously acquired knowledge. This paper presents a comparative study of four fine-tuning strategies: (1) standard supervised learning based on cross-entropy loss (CE), (2) method with L1-regularization of weights (CE-L1), (3) regularization method using a static "teacher" model (KL-CE), and (4) the proposed approach using a Teacher Exponential Moving Average (TEMA), in which the "teacher" weights are updated dynamically via exponential smoothing of the trainee model weights. Experimental validation was performed on Qwen2-0.5B and Qwen2-1.5B models using 4-bit quantization and LoRA adapters on a medical corpus comprising over 27,000 cytological reports based on the Bethesda System. A comprehensive evaluation of generation quality was conducted using lexical (BLEU, ROUGE, METEOR, ChrF) and semantic (BLEURT, BERTScore) metrics, as well as the MMLU benchmark (5-shot) to control for the preservation of general cognitive abilities. The results showed that the KL-CE method limits model adaptation to the new domain, while the L1-regularization method (CE-L1) demonstrates low efficiency in both generation and knowledge retention. At the same time, standard fine-tuning (CE), while providing high quality on new data, reduces generation quality on "general" data not included in the training domain. The proposed TEMA method provides the best balance between plasticity and stability, improves the semantic quality of generation, and minimizes the degradation of general knowledge. The obtained data allow recommending TEMA as an effective tool for adapting LLMs for highly specialized tasks, such as medical diagnostics.

Keywords: *LLM, catastrophic forgetting, Qwen2, Bethesda system, text generation, regularization, Teacher Exponential Moving Average, PEFT.*

REFERENCES

- [1] Raffel, C. et al. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. JMLR, 2020.
- [2] [Liu, P. et al. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. ACM CSUR, 2023.
- [3] Y. Luo, Z. Yang, F. Meng, Y. Li, J. Zhou, and Y. Zhang, "An empirical study of catastrophic forgetting in large language models during continual fine-tuning," arXiv preprint arXiv:2308.08747, 2023.
- [4] Kirkpatrick, J., Pascanu, R., Rabinowitz, N. C., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., Hassabis, D., Clopath, C., Kumaran, D., and Hadsell, R. Overcoming catastrophic forgetting in neural networks. Proceedings of the National Academy of Sciences, 114:3521–3526, 2016.
- [5] Zhicheng Wang, Yufang Liu, Tao Ji, Xiaoling Wang, Yuanbin Wu, Congcong Jiang, Ye Chao, Zhencong Han, Ling Wang, Xu Shao, et al. 2023. Rehearsalfree continual language learning via efficient parameter isolation. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 10933–10946.
- [6] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen. Lora: Low-rank adaptation of large language models. arXiv preprint arXiv:2106.09685, 2021
- [7] Zhang, Y., Jiang, S., Zhao, M., Li, Y., Fan, Y., Wu, X., & Chen, Q. (2025). Gere: Towards efficient anti-forgetting in continual learning of llm via general samples replay. *arXiv preprint arXiv:2508.04676*.
- [8] Sanyal, S., Prairie, H., Das, R., Kavis, A., & Sanghavi, S. (2025). Upweighting easy samples in fine-tuning mitigates forgetting. *arXiv preprint arXiv:2502.02797*.
- [9] Zhizhong Li and Derek Hoiem. 2017. Learning without forgetting. IEEE transactions on pattern analysis and machine intelligence, 40(12):2935–2947.
- [10] Y. Chen, S. Zhang, G. Qi, and X. Guo. Parameterizing context: Unleashing the power of parameter-efficient fine-tuning and in-context tuning for continual table semantic parsing. Advances in Neural Information Processing Systems, 36, 2024.
- [11] Tarvainen, A., & Valpola, H. Mean Teachers are Better Role Models: Weight-Averaged Consistency Targets Improve Semi-Supervised Deep Learning Results. NeurIPS, 2017
- [12] Ali Behrouz, Meisam Razaviyayn, Peilin Zhong, and Vahab Mirrokni. Nested Learning: The Illusion of Deep Learning Architecture [Электронный ресурс] - <https://research.google/blog/introducing-nested-learning-a-new-ml-paradigm-for-continual-learning/> (22.12.2025).
- [13] Qin, Y., Qian, C., Yi, J., Chen, W., Lin, Y., Han, X., ... & Zhou, J. (2022). Exploring mode connectivity for pre-trained language models. arXiv preprint arXiv:2210.14102.
- [14] Ren, W., Li, X., Wang, L., Zhao, T., & Qin, W. (2024). Analyzing and reducing catastrophic forgetting in parameter efficient tuning. arXiv preprint arXiv:2402.18865.
- [15] Ali S, Cibas E. The Bethesda System for Reporting Thyroid Cytopathology. (Ali SZ, Cibas ES, eds.). Cham: Springer International Publishing; 2018. doi: <https://doi.org/10.1007/978-3-319-60570-8>
- [16] Ali SZ, Baloch ZW, Cochand-Priollet B, Schmitt FC, Vielh P, VanderLaan PA. The 2023 Bethesda System for Reporting Thyroid Cytopathology. Thyroid®. July 2023. doi: <https://doi.org/10.1089/thy.2023.0141>
- [17] [Электронный ресурс] - <https://unsloth.ai/>
- [18] [Электронный ресурс] - <https://github.com/BY571/sft-kl-lora-trainer>
- [19] [Электронный ресурс] - https://github.com/EugeneCS/mephi_nlp/tree/sviridenko
- [20] Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002, July). Bleu: a method for automatic evaluation of machine translation. In Proceedings of the 40th annual meeting of the Association for Computational Linguistics (pp. 311-318).
- [21] Chin-Yew, L. (2004). Rouge: A package for automatic evaluation of summaries. In Proceedings of the Workshop on Text Summarization Branches Out, 2004.
- [22] Banerjee, S., & Lavie, A. (2005, June). METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization (pp. 65-72).
- [23] Popović, M. (2017, September). chrF++: words helping character n-grams. In Proceedings of the second conference on machine translation (pp. 612-618).
- [24] Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2019). Bertscore: Evaluating text generation with bert. arXiv preprint arXiv:1904.09675.

- [25] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers) (pp. 4171-4186).
- [26] Liu, Y., Maier, W., Minker, W., & Ultes, S. (2021). Naturalness evaluation of natural language generation in task-oriented dialogues using bert. *arXiv preprint arXiv:2109.02938*.
- [27] [Электронный ресурс] -
<https://huggingface.co/datasets/cais/mmlu>
- [28] Kornblith, S., Norouzi, M., Lee, H., & Hinton, G. (2019, May). Similarity of neural network representations revisited. In *International conference on machine learning* (pp. 3519-3529). PMIR.

