

# Вербализированное машинное обучение: пересмотр машинного обучения с языковыми моделями

Дата: 2025-02-14 00:00:00

Ссылка на исследование: <https://arxiv.org/pdf/2406.04344>

Рейтинг: 75

Адаптивность: 80

## Ключевые выводы:

Исследование представляет концепцию вербализованного машинного обучения (VML), где параметры модели представлены в виде естественного языка, а не числовых значений. Основная цель - создать интерпретируемые модели машинного обучения, где большие языковые модели (LLM) используются как аппроксиматоры функций, параметризованные текстовыми промптами. Результаты показывают, что VML может эффективно решать классические задачи машинного обучения, обеспечивая при этом полную интерпретируемость.

## Объяснение метода:

Исследование предлагает революционный подход к машинному обучению через вербализацию параметров моделей в естественном языке. Высокая концептуальная ценность и интерпретируемость делают метод полезным для широкой аудитории. Ограничения связаны с необходимостью доступа к API или локальным LLM для полной реализации и сложностями работы с высокоразмерными данными.

## Ключевые аспекты исследования: 1. **Концепция вербализованного машинного обучения (VML)** - исследование предлагает новую парадигму машинного обучения, где параметры модели представлены в виде естественного языка, а не числовых значений. Это делает процесс обучения и модели полностью интерпретируемыми.

**Двухкомпонентная архитектура** - VML использует две языковые модели: модель-ученик (learner LLM), которая делает предсказания на основе вербализованных параметров, и модель-оптимизатор (optimizer LLM), которая обновляет эти параметры на основе ошибок предсказания.

**Итеративное обучение** - процесс обучения происходит итеративно: оптимизатор анализирует ошибки предсказания ученика и обновляет текстовое описание модели для улучшения точности.

**Задание индуктивных предпочтений** - VML позволяет легко внедрять предварительные знания о задаче и желаемой структуре модели через

естественный язык.

**Автоматический выбор класса модели** - оптимизатор может самостоятельно выбирать подходящий класс модели (линейная, полиномиальная и т.д.) на основе данных и обновлять его во время обучения.

## Дополнение:

### Действительно ли для работы методов этого исследования требуется дообучение или API?

Нет, методы этого исследования **не требуют** дообучения языковых моделей. Вся работа проводилась с использованием предобученных моделей без дополнительной настройки. Доступ к API или локальным моделям необходим в основном для возможности запуска двух параллельных экземпляров LLM (learner и optimizer) и для удобства автоматизации процесса.

### Концепции и подходы, применимые в стандартном чате:

**Вербализация моделей** - пользователь может попросить LLM описать модель для конкретной задачи на естественном языке.

**Итеративное улучшение** - пользователь может показать LLM результаты предсказания, указать на ошибки и попросить улучшить модель.

**Включение предварительных знаний** - пользователь может включить свои знания о предметной области в промпт.

**Автоматический выбор модели** - пользователь может попросить LLM рассмотреть разные типы моделей и выбрать наиболее подходящую.

**Интерпретация результатов** - LLM может объяснить, почему модель делает определенные предсказания.

### Ожидаемые результаты адаптации:

Более интерпретируемые и понятные модели для неспециалистов  
Возможность использования предметных знаний без программирования  
Лучшее понимание процесса обучения и принятия решений моделью  
Возможность создавать простые предсказательные модели в диалоговом режиме  
Повышение доверия к результатам благодаря прозрачности процесса  
Главное ограничение при работе в стандартном чате - необходимость вручную проводить итерации обучения и ограниченный контекст чата, что может затруднять работу с большими объемами данных.

## Анализ практической применимости: 1. **Концепция вербализованного машинного обучения** - Прямая применимость: Высокая для пользователей, работающих с LLM. Позволяет создавать интерпретируемые модели без программирования. - Концептуальная ценность: Исключительно высокая. Даёт

понимание того, как LLM могут "рассуждать" о моделях и обучении. - Потенциал для адаптации: Высокий. Подход может быть адаптирован для широкого спектра задач, от простой регрессии до классификации изображений.

**Двухкомпонентная архитектура** Прямая применимость: Средняя. Требуется доступа к API или самостоятельного хостинга LLM, что не всегда доступно обычным пользователям. Концептуальная ценность: Высокая. Разделение "учёбы" и "оптимизации" помогает понять процесс обучения моделей. Потенциал для адаптации: Средний. Архитектура может быть упрощена для использования одной LLM в стандартном чате.

### **Итеративное обучение**

Прямая применимость: Средняя. Итеративный процесс может быть реализован вручную пользователем через диалог с LLM. Концептуальная ценность: Высокая. Показывает, как LLM могут улучшать свои решения через итерации. Потенциал для адаптации: Высокий. Пользователи могут применять этот подход в обычных разговорах с LLM.

### **Задание индуктивных предпочтений**

Прямая применимость: Очень высокая. Пользователи могут легко сообщать свои предварительные знания через обычный язык. Концептуальная ценность: Исключительная. Демонстрирует, как естественный язык может служить универсальным способом передачи предметных знаний. Потенциал для адаптации: Очень высокий. Подход применим в любых взаимодействиях с LLM.

### **Автоматический выбор класса модели**

Прямая применимость: Средняя. Требуется специальной настройки промптов. Концептуальная ценность: Высокая. Показывает способность LLM к метарассуждению о подходящих моделях. Потенциал для адаптации: Высокий. Может быть адаптирован для разных задач через соответствующие промпты.

## **Prompt:**

Использование концепции VML в промптах для GPT ## Ключевые принципы VML для промптов

Исследование вербализованного машинного обучения (VML) предлагает несколько ценных подходов, которые можно адаптировать для создания более эффективных промптов:

**Параметризация через естественный язык** - использование текстовых описаний вместо числовых значений **Итеративное улучшение** - пошаговая оптимизация ответов **Включение индуктивных предположений** - явное указание ожидаемых паттернов в данных **Интерпретируемость** - требование объяснения каждого шага рассуждения ## Пример промпта с применением принципов VML

[=====] # Задача анализа временного ряда продаж

## Контекст и данные Вот ежемесячные данные продаж за последние 2 года: [данные продаж]

## Инструкции с применением VML 1. Проанализируй данные, предполагая, что в них может присутствовать сезонность и долгосрочный тренд (индуктивное предположение)

Для каждого шага анализа: Опиши своё текущее понимание данных в виде словесной модели Объясни, почему ты выбрал именно эту модель Предложи прогноз на следующие 3 месяца Оцени точность прогноза и предложи улучшения модели

Итеративно улучшай свою словесную модель минимум 3 раза, каждый раз объясняя:

Что не работало в предыдущей версии Какие новые паттерны ты заметил Как новая модель учитывает эти паттерны

Финальный ответ должен содержать:

Окончательную словесную модель данных Прогноз на следующие 3 месяца с обоснованием Ограничения твоей модели [=====] ## Как это работает

Данный промпт применяет ключевые принципы VML:

**Вербализация параметров** - мы просим GPT описывать свою модель словами, а не числами, что делает рассуждения прозрачными и понятными

**Итеративная оптимизация** - требуем минимум 3 итерации улучшения, аналогично тому, как в VML оптимизатор постепенно улучшает модель

**Индуктивные предположения** - явно указываем на возможность сезонности и тренда, направляя модель на поиск этих паттернов

**Прозрачность рассуждений** - требуем объяснения каждого шага и обоснования изменений, что делает процесс полностью интерпретируемым

Такой подход позволяет получить не только конечный результат, но и понять логику его формирования, что повышает доверие к ответам модели и их практическую применимость.