

# Мультимоделирование, привилегированное обучение\*

Нечепуренко И. О.<sup>1</sup>, Нейчев Р. Г.<sup>2</sup>, Стрижов В. В.<sup>2</sup>

author@site.ru

<sup>1</sup>Организация; <sup>2</sup>Организация

Цель данной работы - опробовать различные методы построения моделей оптимальной вычислительной сложности. Затраты по времени и памяти для работы модели играют очень важную роль в большинстве сфер человеческой жизнедеятельности. Носимая электроника и защищенные устройства для решения задач биометрии, устройства автоматической обработки телеметрических данных, системы потоковой аналитики результатов коллизий Большого Адронного Коллайдера — это лишь малая доля случаев, когда требуются достаточно быстрые алгоритмы машинного обучения. Существует широкий спектр способов уменьшения сложности модели. Один из них - это разбиение объектов на подобласти, и описание данных в каждой своей собственной моделью. Этот способ называется мультимоделированием (в иноязычной литературе Mixture experts). Другой способ - мета-обучение (Metalearning), метод, основанный на парадигме учителя-ученика, но при этом в качестве учителя могут выступать не только ответы на экспериментальных данных, но и результаты работы другой модели учителя. В случае добавления дополнительной, априорной информации, можно значительно улучшить результаты сходимости происходящего во время обучения оптимизационного процесса, снизить сложность модели, а также повысить итоговое качество модели. Эти свойства были проверены на реальных данных.

**Ключевые слова:** *машинное обучение, мета-обучение, мультимоделирование*.

## 1 Введение

Основное сообщение — чему посвящена работа (одна-две фразы) Рассматривается задача построения алгоритмов анализа данных максимальной точности с ограниченными затратами ресурсов. Конкретно работа посвящена самому распространенному типу машинного обучения - обучению с учителем. Обычно в промышленности и науке машинное обучение связано с обработкой больших массивов данных и итерационными процессами оптимизации, а это влечет за собой сильные затраты по ресурсам, которые не могут себе позволить даже крупные компании. Рассматриваемая нами задача применима практически во всех сферах науки и бизнеса.

Существуют различные подходы к решению задачи. Один из них - мультимоделирование. Нам интересен один конкретный [1] алгоритм: смесь экспертов. В реализации этого алгоритма специальная функция, именуемая шлюзовой, определяет ценность предсказаний конкретного эксперта. Одной из полезных особенностей способа можно назвать возможность отфильтровывать "слабые" модели.

Другой подход - модификация алгоритма обучения с учителем, в котором в роли учителя могут выступить не только экспериментальные данные, но и ответы машины, обученной на более полных данных либо более сложным алгоритмом. Два метода, основанных на парадигме "машина учит машину": дистилляция[2] и контроль сходства[3], были обобщены[4].

Современное состояние области (два-четыре абзаца)

---

\*Работа выполнена при финансовой поддержке РФФИ, проект №00-00-00000. Научный руководитель: Стрижов В. В. Задачу поставил: Эксперт И. О. Консультант: Консультант И. О.

Рассматриваемые нами алгоритмы достаточно молодые: например, контроль сходства был представлен Вапником в 2009 г., а метод дистилляции - в 2015 -м году Д. Хинтоном. Тем не менее, эти алгоритмы уже используются на практике: например, не так давно они использовались для использования энергии датацентрами Яндекса.

Что предлагается (два абзаца)

Предлагается на реальных данных: данных о ценах энергопотребления в Польше в зависимости от времени протестировать известные алгоритмы, сравнить сложности вычисления и достигаемые качества при различных реализациях.

После аннотации, но перед первым разделом, располагается введение, включающее в себя описание предметной области, обоснование актуальности задачи, краткий обзор известных результатов, и т. п. [?, ?, ?, ?, ?, ?].

## 2 Название раздела

Данный документ демонстрирует оформление статьи, подаваемой в электронную систему подачи статей <http://jmla.org/papers> для публикации в журнале «Машинное обучение и анализ данных». Более подробные инструкции по стилевому файлу `jmla.sty` и использованию издательской системы  $\text{\LaTeX}$  2 $\epsilon$  находятся в документе `authors-guide.pdf`. Работу над статьёй удобно начинать с правки  $\text{\TeX}$ -файла данного документа.

### 2.1 Название параграфа.

Нет ограничений на количество разделов и параграфов в статье. Разделы и параграфы не нумеруются.

### 2.2 Теоретическую часть работы

желательно структурировать с помощью окружений `Def`, `Axiom`, `Hypothesis`, `Problem`, `Lemma`, `Theorem`, `Corollary`, `State`, `Example`, `Remark`.

**Определение 1.** Математический текст хорошо структурирован, если в нём выделены определения, теоремы, утверждения, примеры, и т. д., а неформальные рассуждения (мотивации, интерпретации) вынесены в отдельные параграфы.

**Утверждение 1.** Мотивации и интерпретации наиболее важны для понимания сути работы.

**Теорема 1.** Не менее 90% коллег, заинтересовавшихся Вашей статьёй, прочитают в ней не более 10% текста.

**Доказательство.** Причём это будут именно те разделы, которые не содержат формул. ■

**Замечание 1.** Выше показано применение окружений `Def`, `Theorem`, `State`, `Remark`, `Proof`.

## 3 Некоторые формулы

Образец формулы:  $f(x_i, \alpha^\gamma)$ .

Образец выключной формулы без номера:

$$y(x, \alpha) = \begin{cases} -1, & \text{если } f(x, \alpha) < 0; \\ +1, & \text{если } f(x, \alpha) \geq 0. \end{cases}$$

Образец выключной формулы с номером:

$$y(x, \alpha) = \begin{cases} -1, & \text{если } f(x, \alpha) < 0; \\ +1, & \text{если } f(x, \alpha) \geq 0. \end{cases} \quad (1)$$

Образец выключной формулы, разбитой на две строки с помощью окружения align:

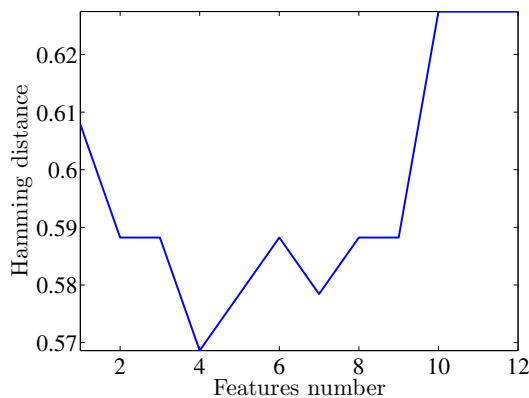
$$R'_N(F) = \frac{1}{N} \sum_{i=1}^N \left( P(+1 | x_i) C(+1, F(x_i)) + \right. \\ \left. + P(-1 | x_i) C(-1, F(x_i)) \right). \quad (2)$$

Образцы ссылок: формулы (1) и (2).

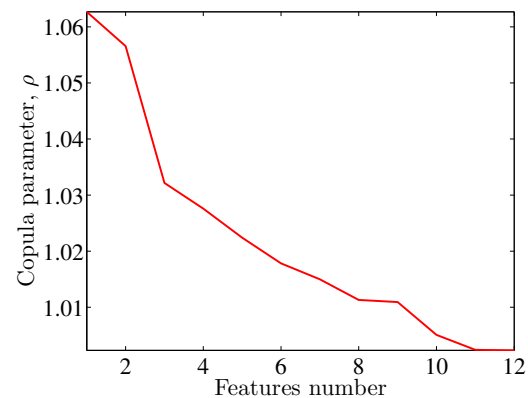
## 4 Пример иллюстрации

Рисунки вставляются командой `\includegraphics`, желательно с выравниванием по ширине колонки: `[width=\linewidth]`.

Практически все популярные пакеты рисуют графики с подписями, которые трудно читать на бумаге и на слайдах из-за малого размера шрифта. Шрифт на графиках (подписи осей и цифры на осях) должны быть такого же размера, что и основной текст.



(а) Первый рисунок



(б) Второй рисунок

**Рис. 1** Подпись должна размещаться под рисунком.

При значительном количестве рисунков рекомендуется группировать их в одном окружении `{figure}`, как это сделано на рис. 1.

## 5 Пример таблицы

Подпись делается *над таблицей*, см. таблицу 1.

## 6 Заключение

Желательно, чтобы этот раздел был, причём он не должен дословно повторять аннотацию. Обычно здесь отмечают, каких результатов удалось добиться, какие проблемы остались открытыми.

**Таблица 1** Подпись размещается над таблицей.

Задача	CCEL	boosting
<b>Cancer</b>	<b>3.46</b> $\pm$ 0.37 (3.16)	4.14 $\pm$ 1.48
<b>German</b>	<b>25.78</b> $\pm$ 0.65 (1.74)	29.48 $\pm$ 0.93
<b>Hepatitis</b>	18.38 $\pm$ 1.43 (2.87)	19.90 $\pm$ 1.80

[1] Yuksel Seniha Esen, Wilson Joseph N., Gader Paul D. Twenty Years of Mixture of Experts // IEEE Transactions on Neural Networks and Learning Systems. 2012. T. 23, № 8. С. 1177–1193.

[2] Hinton Geoffrey E., Vinyals Oriol, Dean Jeffrey. Distilling the Knowledge in a Neural Network // CoRR. 2015. T. abs/1503.02531. URL: <http://arxiv.org/abs/1503.02531>.

[3] Vapnik V., Vashist A. A new learning paradigm: Learning using privileged information. // Neural Networks. 2009.

[4] David Lopez-Paz, Léon Bottou, Bernhard Schölkopf, Vladimir Vapnik Unifying distillation and privileged information

<https://arxiv.org/abs/1511.03643>