

33. ПРИНЦИП ДИРИХЛЕ И ЕГО ПРИМЕНЕНИЕ В МАШИННОМ ОБУЧЕНИИ

*Спарин Д.Е., Кудрявец А.А., студенты гр.373901, Русина Н. В., аспирант
Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Ефремов А.А. – канд. экон. наук, доцент каф. ЭИ

Аннотация. Данная статья посвящена исследованию принципа Дирихле, применения данного принципа в машинном обучении. Она включает в себя описание принципа Дирихле, его доказательство, применение данного принципа в различных сферах науки. В статье также изложена информация о латентном распределении Дирихле(LDA), описано применение LDA в машинном обучении, анализе данных.

Ключевые слова. Принцип Дирихле, машинное обучение, латентное распределение Дирихле, LDA.

Принцип Дирихле — это фундаментальный математический инструмент, который находит применение в различных областях, включая теорию чисел, комбинаторику, криптографию, и в том числе машинное обучение. Данный принцип был назван в честь немецкого математика Петера Густава Лейбница Дирихле.

Самая популярная формулировка принципа Дирихле такова: «Если в n клетках сидит m зайцев, причем $m > n$, то хотя бы в одной клетке сидят по крайней мере два зайца». На первый взгляд даже непонятно, почему это совершенно очевидное замечание является весьма эффективным методом решения задач. Дело в том, что в каждой конкретной задаче нелегко бывает понять, что же здесь «зайцы» и «клетки» и почему зайцев больше, чем клеток. Выбор зайцев и клеток часто неочевиден; далеко не всегда по виду задачи можно определить, что следует воспользоваться принципом Дирихле.

А главное, этот метод дает неконструктивное доказательство (мы, естественно, не можем сказать, в какой именно клетке сидят два зайца, а знаем только, что такая клетка есть), а попытка дать конструктивное доказательство, т. е. доказательство путем явного построения или указания требуемого объекта, может привести к гораздо большим трудностям.

Как и любое утверждение в математике, принцип Дирихле требует доказательства. В данном случае доказательство строится методом от противного:

Пусть есть N клеток и $N+1$ кролик. Предположим, что в каждой клетке не более одного кролика:

- В первой - хотя бы один;
- Во второй - хотя бы один;
- В клетке с номером N - хотя бы один.

$$\begin{aligned}n_1 + n_2 + n_3 \dots + n_4 &\leq 1 + 1 + 1 + \dots + 1 \\n_1 + n_2 + n_3 \dots + n_4 &\leq N \\N + 1 &\leq N - \text{противоречие}\end{aligned}$$

Рисунок 1 – Неравенства при доказательстве принципа Дирихле

Было получено N неравенств, которые складываются. Слева расположены «кролики», справа расположены «клетки». В итоге мы приходим к противоречию в том, что кроликов меньше, чем клеток, хотя в условии утверждалось обратное.

Принцип Дирихле имеет важные приложения в криптографии. Например, он используется в алгоритме RSA – одном из самых распространенных алгоритмов асимметричного шифрования. В RSA зашифрование сообщения происходит с использованием больших простых чисел p и q . Чтобы получить такие простые числа, генерируют случайное большое число n и разлагают его на множители. Принцип Дирихле гарантирует, что среди множителей числа n обязательно найдутся большие простые числа p и q , необходимые для зашифрования.

Принцип Дирихле также широко применяется в физике, инженерии и других науках для решения задач, связанных с электростатикой, теплопроводностью и другими явлениями, где важную роль играют потенциалы и граничные условия. Он является одним из основных принципов теории потенциала и имеет важное значение для понимания различных физических процессов.

Машинное обучение – это наука о разработке алгоритмов и статистических моделей, которые компьютерные системы используют для выполнения задач без явных инструкций, полагаясь вместо этого на шаблоны и логические выводы. Компьютерные системы используют алгоритмы машинного обучения для обработки больших объемов статистических данных и выявления шаблонов данных. Таким образом, системы могут более точно прогнозировать результаты на основе заданного набора входных данных. Например, специалисты по работе с данными могут обучить медицинское приложение диагностировать рак по рентгеновским изображениям, сохраняя миллионы отсканированных изображений и соответствующие диагнозы.

Машинное обучение помогает компаниям стимулировать рост, открывать новые источники дохода и решать сложные проблемы. Данные являются важной движущей силой принятия бизнес-решений, но традиционно компании использовали данные из различных источников, таких как отзывы клиентов, сотрудников и финансов. Исследования в области машинного обучения автоматизируют и оптимизируют этот процесс. Используя ПО, которое анализирует очень большие объемы данных на высокой скорости, компании могут быстрее достигать результатов.

В контексте машинного обучения принцип Дирихле может быть использован для решения задач регуляризации и обучения моделей. Он может помочь в управлении сложностью модели, предотвращая переобучение и обеспечивая её обобщающую способность.

Принцип Дирихле в машинном обучении может быть проиллюстрирован на примере регуляризации $L1$ (Lasso) и $L2$ (Ridge). Регуляризация $L1$ добавляет штраф к функции потерь, который зависит от суммы абсолютных значений весов модели. Это позволяет уменьшить количество параметров модели и отбросить незначимые признаки.

Регуляризация L2, с другой стороны, добавляет штраф к функции потерь, который зависит от квадрата суммы значений весов модели. Это помогает уменьшить веса признаков и предотвратить переобучение модели.

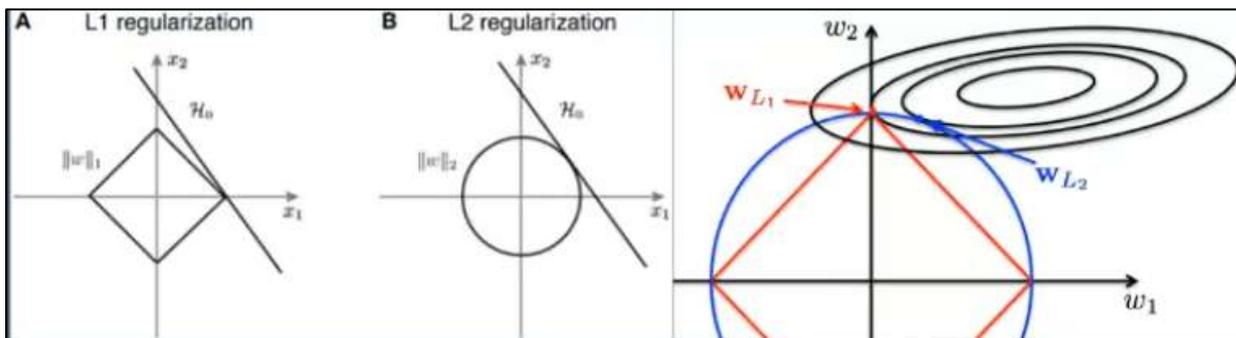


Рисунок 3 – регуляризация L1(A) и L2(B)

В машинном обучении также применяется латентное распределение Дирихле(LDA). Данная модель была предложена Дэвидом Блеем в 2003 году. LDA является одним из методов тематического моделирования и которая позволяет выявить заданное количество тем в корпусе текстов. В основе метода лежит предположение, что каждый документ представляет собой комбинацию ограниченного числа тем, а каждая тема — это распределение вероятностей для слов. При этом, как и в естественном языке, документы могут перекрывать друг друга по темам, а темы — по словам.

Алгоритм латентного распределения Дирихле рассчитывает ассоциацию слов с темами и тем с документами. При этом LDA относится к числу методов машинного обучения без учителя (unsupervised machine learning), то есть не требует предварительной разметки корпуса: система сама «находит» скрытые в корпусе темы и аннотирует каждый документ. Это делает метод востребованным в тех случаях, когда мы сами точно не знаем, что ищем — например, в исследованиях электронных архивов.

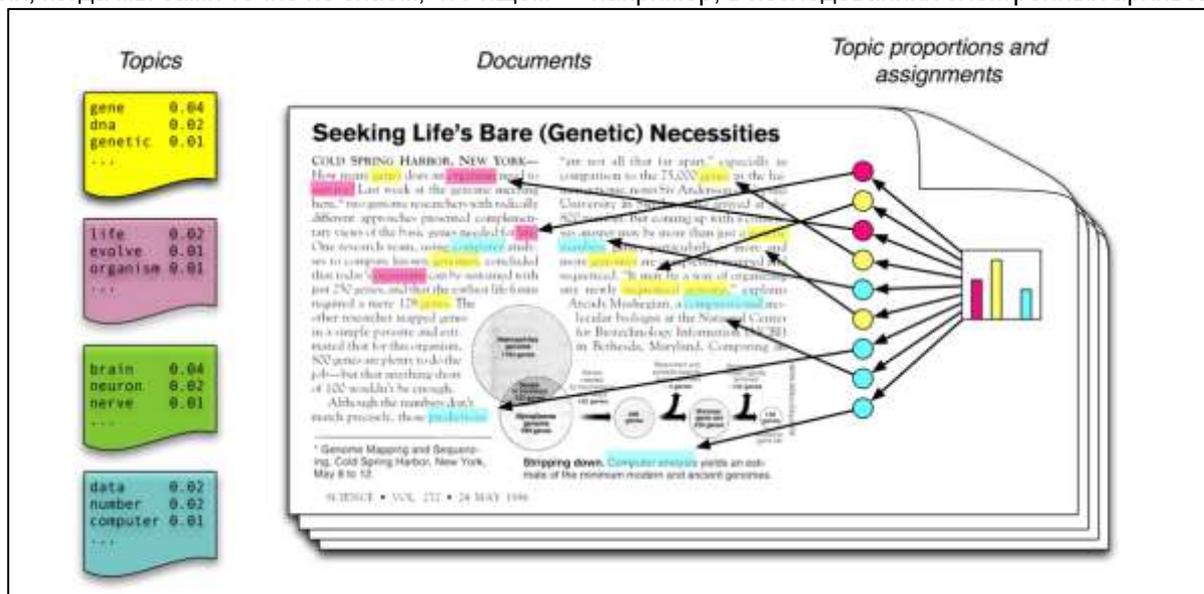


Рисунок 2 – Основные принципы алгоритма латентного размещения Дирихле

На рисунке голубым выделена тема «анализ данных»; розовым — «эволюционная биология», а жёлтым — «генетика». Если разметить все слова в тексте (за исключением «шумовых», таких как союзы, артикли и т.п.), то можно увидеть, что документ представляет собой сочетание нескольких тем. Цветные «окошки» слева — это распределение вероятностей для слов в теме. Гистограмма справа — это распределение вероятностей для тем в документе. Все документы в коллекции представляют собой сочетание одних и тех же тем — но в разной пропорции.

Латентное распределение Дирихле активно применяется в истории науки и философии. Примером применения LDA является публикация 2021 года, вышедшая в журнале Science Education и посвящённая тематике этого журнала в период с 1922 по 2019 года. Авторы этого проекта исследовали

*60-я Юбилейная Научная Конференция Аспирантов, Магистрантов и Студентов БГУИР,
Минск 2024*

около 5,5 тыс. статей, на основе которых они выделили 21 тему и проследили их популярность в диахронной перспективе. Как признают сами авторы, многие из выявленных трендов хорошо знакомы специалистам в отрасли, однако LDA впервые позволяет квантифицировать эти тренды. Похожим образом метод применялся для анализа трендов в онлайн-обучении и в библиотековедении.

Принцип Дирихле служит важным инструментом для точного изучения явлений и процессов в различных областях науки и техники. Он позволяет анализировать поведение функций, опираясь на их значения на границе областей, и находит применение в различных научных дисциплинах.

Список использованных источников:

1. *Introduction to Probabilistic Topic Models'* David M. Blei <https://www.eecis.udel.edu/~shatkay/Course/papers/UIntrotoTopicModelsBlei2011-5.pdf>.
2. *Принцип Дирихле: формулировка, задачи с решениями:* <https://fb.ru/article/16967/2023-2023-printsip-dirihle-formulirovka-zadachi-s-resheniyami>
3. *Латентное распределение Дирихле:* <https://fineproxy.org/ru/wiki/latent-dirichlet-allocation/>