Основания алгебраического подхода к синтезу корректных алгоритмов

Лектор — Рудаков К.В.
Наборщик — Старожилец В.М.

Оглавление

Лекция 1

Введение

Данные лекции рассматривают общую задачу машинного обучения без привязки к конкретным методам и основы алгебраического подхода к синтезу корректных алгоритмов для её решения. В некотором роде они являются взглядом сверху на задачи машинного обучения и методы их решения.

В первую очередь следует сформулировать задачу машинного обучения в общем виде. По сути это задача построения алгоритма, который реализует отображение из множества начальных информаций в множество конечных информаций. Сразу отметим, что в курсе рассматриваются только такие отображения, для которых существует реализующий их алгоритм.

Определение. *Символом $\mathfrak{I}_\mathfrak{i}$ (читается «И инишл») будем обозначать множество начальных информаций, например, симптомы болезни.*

Определение. *Символом $\mathfrak{I}_\mathfrak{f}$ (читается «И файнэл») будем обозначать множество конечных информаций, например, диагноз.*

Таким образом, на формальном языке нам требуется найти такой алгоритм A, что он осуществляет отображение из множества начальных информаций $\mathfrak{I}_\mathfrak{i}$ в множество конечных информаций $\mathfrak{I}_\mathfrak{f}$:

$$A : \mathfrak{I}_\mathfrak{i} \to \mathfrak{I}_\mathfrak{f}.$$

Пока что задача стоит так, что нам нужно найти некоторое произвольное отображение из одного множества в другое, реализуемое некоторым алгоритмом. При этом свойства этого отображения и алгоритма неважны. В такой постановке у нас нет каких-либо ограничений на искомый алгоритм: даже датчик случайных чисел является решением этой задачу. Поэтому вводятся дополнительные ограничения на допустимые алгоритмы. Итак,

Определение. *Обозначим $\mathfrak{M}^* = \{A |\ A : \ \mathfrak{I}_\mathfrak{i} \to \mathfrak{I}_\mathfrak{f}\}$ множество всех алгоритмов, реализующих отображение из $\mathfrak{I}_\mathfrak{i}$ в $\mathfrak{I}_\mathfrak{f}$.*

Определение. *Обозначим I_{str} структурную информацию, содержащую условия и требования, накладываемые на A.*

Определение. *Обозначим $\mathfrak{M}(I_{str}) \subset \mathfrak{M}^*$ некоторое подмножество $\mathfrak{M}^*$, удовлетворяющее I_{str}.*

Теперь у нас есть дополнительная информация I_{str}, позволяющий накладывать дополнительные ограничения на нашу задачу. Введём определения допустимого отображения и корректного алгоритма.

Определение. *Любое отображение из множества $\mathfrak{M}(I_{str})$ называется допустимым.*

Определение. *Задача Z заключается в построении алгоритма, реализующего допустимое отображение.*

Определение. *Любой алгоритм реализующий любое допустимое отображение называется корректным.*

В такой формулировке необходимым и достаточным условием разрешимости задачи Z является выполнение выражения:

$$\mathfrak{M}(I_{str}) \neq \emptyset,$$

а условием единственности решения — выполнение равенства:

$$|\mathfrak{M}(I_{str})| = 1.$$

Заметим также, что в данной формулировке корректный алгоритм — это алгоритм, не допускающий ни одной ошибки, а множество $\mathfrak{M}(I_{str})$ — множество алгоритмов не допускающих ошибок. Однако можно поставить условия несколько мягче, и дать алгоритмам возможность ошибаться.

Поиск решения задачи

Пусть $\mathfrak{M}(\pi)$ — некоторое параметрическое семейство отображений. После того как мы выбрали некоторое семейство отображений $\mathfrak{M}(\pi)$, попытаемся попасть в $\mathfrak{M}(I_{str})$, взяв в $\mathfrak{M}(\pi)$ какое-нибудь отображение за начальное. Это возможно, если данные семейства пересекаются:

$$\mathfrak{M}(\pi) \cap \mathfrak{M}(I_{str}) \neq \emptyset.$$

Но, с одной стороны, чем *сложнее* наше семейство, тем выше вероятность, что оно пересекается с семейством $\mathfrak{M}(I_{str})$, однако, с другой стороны, достижение этого пересечения может быть *затратно*, если $\mathfrak{M}(\pi)$ сложное. Также всегда остаётся вероятность, что множество $\mathfrak{M}(\pi)$ с $\mathfrak{M}(I_{str})$ не пересекается. Для поиска компромиссного решения используют идею *расширения множества.*

Определение. *Пусть f — некоторая операция над множеством $\mathfrak{M}^*$. Тогда $f(\mathfrak{M}(\pi))$ будем называть расширением множества $\mathfrak{M}(\pi)$.*

Таким образом, мы хотим расширить некоторое «простое» множество до пересечения с $\mathfrak{M}(I_{str})$. Однако, не любая функция f нам подходит, так как «простое» множество может расшириться до слишком «сложного» множества. Важно, что f мы выбираем сами, поэтому можем выбрать его так, чтобы искать нужный алгоритм было не слишком сложно.

Лекция 2

Алгебра, реляционная система и алгебраическая система

Данная лекция посвящена вопросам терминологии, используемой в данном курсе. Поэтому тут будет очень много определений (ещё больше, чем в предыдущей). Для начала определим понятия алгебры, реляционной системы и алгебраической системы.

Определение. *Сигнатура —набор характеристик, однозначно идентифицирующий объект.*

Определение. *Отношение — математическая структура, которая формально определяет свойства различных объектов и их взаимосвязи. В нашем курсе будет обозначаться буквой R.*

Определение. *Алгеброй называется структура*

$$\begin{pmatrix} A & Op_1 & Op_2 & ... & Op_k \\ & n_1 & n_2 & ... & n_k \end{pmatrix}$$

где A — множество, Op_i — операции на этом множестве, n_i — сигнатуры.

Определение. *Реляционной системой называется структура*

$$\begin{pmatrix} A & R_1 & R_2 & ... & R_k \\ & n_1 & n_2 & ... & n_k \end{pmatrix}$$

где A — множество, R_i — отношения на этом множестве, n_i — сигнатуры.

Определение. *Алгебраической системой называется структура*

$$\begin{pmatrix} A & R_1 & R_2 & ... & R_k & Op_1 & Op_2 & ... & Op_l \\ & n_{1,1} & n_{1,2} & ... & n_{1,k} & n_{2,1} & n_{2,2} & ... & n_{2,l} \end{pmatrix}$$

где A — множество, R_i — отношения на этом множестве, Op_i — операции на этом множестве, n_i — сигнатуры.

Первичные свойства фанкции

Теперь поговорим о функциях. Первичные свойства функций — это иньективность, сурьективность и биективность. Все остальные свойства требуют задать некоторую структуру на тех множествах, на которых они действуют (например, метрику). Пусть функция f действует из A в B. То есть:

$$\begin{cases} f : A \to B \\ f(A) \subseteq B \end{cases}$$

Определим также понятие отношения эквивалентности на множестве A:

Определение. *Отношением эквивалентности π_f на множестве A называется бинарное отношение, которое обладает свойствами транзитивности, симметричности и рефлексивности.*

Данное определение приводит нас к определению фактор-множества A_{π_f} и ядерной эквивалентности отображения f.

Определение. *Фактор-множество A_{π_f} — это множество всех классов эквивалентности заданного множества A, по заданному отношению π_f.*

Определение. *Ядерная эквивалентность отображения f:*

$$(a_1 \equiv a_2) \equiv (f(a_1) = f(a_2)),$$

где эквивалентность понимается в смысле π_f. Следует понимать, что мы выбираем π_f так, чтобы это свойство было выполнено. Это выполнено не для любого отношения эквивалентности.

Саня, тут надо как нибудь переписать... непонятно что откуда идет. Я так понимаю что мы π_f выбираем по f, но чёрт его знает.

Таким образом, мы можем, например, показать, что любое отображение из A в B раскладывается в суперпозицию суръекции, инъекции и биекции. Данный факт легко понять с помощью рисунка 2.1:

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "f_s"'] & B \\
A_{\pi_f} \arrow[r, "f_{1-1}"] & f(A) \arrow[u, "f_i"']
\end{tikzcd}$$

Рис. 2.1: Иллюстрация представления отображения f как суперпозиции суръекции, инъекции и биекции. f_S - суръективное отображение, f_{1-1} - биекция, f_i - инъекция

На данном рисунке изображены четыре множества: A, B, A_{π_f} —фактормножество, и $f(A)$. Множество A суръективно отображается в свое фактормножество A_{π_f}, из-за ядерной эквивалентности отображения f, A_{π_f} биективно отображается в $f(A)$. В свою очередь $f(A)$ иньективно вкладывается в B как его подмножество.

О декартовом произведении множеств

Поговорим о декартовом произведении множеств и том, как его можно представить через другие операции с множествами. Итак, пусть есть множество индексов $\mathfrak{A} = \{\alpha\}$ и соответствующий этому множеству индексов набор множеств $\{A_\alpha |\ \alpha \in \mathfrak{A}\}$. Чему тогда равно произведение $\prod_{\alpha\in\mathfrak{A}} A_\alpha$? По крайней мере оно не нулевое так как любое декартово произведение произвольного семейства непустых множеств в непустом количестве непусто (об этом свидетельствует теорема выбора). Оказывается, что

$$\prod_{\alpha\in\mathfrak{A}} A_\alpha = \{f|\ f:\ \mathfrak{A} \to \bigcup_{\alpha\in\mathfrak{A}} A_\alpha, \forall\alpha \in \mathfrak{A} : f(\alpha) \in A_\alpha\}$$

Например, пусть

$$A_\alpha = A_{(x,y)} = \{(x', y')|\rho((x, y), (x', y')) \leq 1\}$$

Тогда,

$$\prod_{\alpha\in\mathfrak{A}} A_\alpha = \{f|f : R^2 \to R^2,\ \forall(x, y) : \rho((x, y), f((x, y))) \leq 1\}$$