0.1 Колмогоровская сложность моделей

Определение 1. *Пусть задано отображение из множества бинарных слов в себя:*

$$D : \{0,1\}^* \to \{0,1\}^*.$$

Колмогоровской сложностью бинарной строки x назовем минимальную длину описания относительно D:

$$K_D(x) = \min_{p \in \{0,1\}^*} \{|p| : D(p) = x\},$$

Перечислим некоторые свойства колмогоровской сложности [**?**].

Независимости от способа написания. Пусть заданы отображения D_1, D_2, такие что существуют константы c_1, c_2 такие что для любого другого отображения D' и для любой строки x:

$$K_{D_1} \leq K_{D'} + c_1, \quad K_{D_2} \leq K_{D'} + c_2.$$

Тогда $K_{D_1} = K_{D_2} + O(1)$.

В дальнейшем будем фиксировать некоторое отображение D и положим $K(x) = K_D(x)$.

Невычислимость

Теорема 1. *Пусть k — произвольная вычислимая функция. Если $k(x) \leq K(x)$ для всех x, для которых определена k, то k — ограничена.*

Из теоремы следует, что колмогровская сложность в общем случае невычислима: любая оценка сложности будет ограничена.

Определение 2. *Пусть задано отображение из декартового произведения двух множеств бинарных слов в себя:*

$$D : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*.$$

Условной колмогоровской сложностью бинарной строки y при условии x назовем минимальную длину описания относительно D:

$$K_D(y|x) = \min_{p \in \{0,1\}^*} \{|p| : D(p, y) = x\},$$

Оценка условной Колмогоровской сложности [?]

$$K(x, y) \leq K(x) + K(y|x) + O(\log K(x, y)).$$

Разность $I(x : y) = K(y) - K(y|x)$ задает количество информации в x об объекте y. **Количество информации в паре x,y симметрично с точностью до константы**:

$$I(x : y) = I(y : x) + O(\log K(x, y)).$$

Отметим, что схожими свойствами обладает взаимная информация, определение которой дано ниже.

Определение 3. *Пусть задано дискретная случайная величина x с вероятностным распределением p, принимающая значения $x_1, \dots, x_n$, Энтропией распределения случайной величины x назовем:*

$$H(x) = -\sum_{i=1}^{n} p(x = x_i) \log\ p(x = x_i).$$

Взаимной информацией I двух случайных величин x, y назовем следующее выражение:

$$I(x, y) = H(x) - H(x|y), \quad H(x) = -\sum_{i} p_x(x_i) \log p_x(x_i)$$

$$I(x, y) = I(y, x).$$

Теорема 2. *Пусть задано семейство частично-определенных отображений* $\mathfrak{D} = \{D : \{0,1\}^* \to \{0,1\}^*\}$*, такое что для любого отображения* $D \in \mathfrak{D}$ *и элемента из области определения* D *в области определения не содержится префиксов этого элемента.*

Тогда для минимальной средней длины описания слова:

$$L = \min_{D \in \mathfrak{D}} \sum_i |D(x_i)| p(x = x_i)$$

справедливо неравенсвто:

$$H(x) \leq L \leq H(x) + 1.$$

Теорема 3. *Пусть* f *— вычислимое распределение на бинарных словах. Тогда справедливы следующие оценки:*

$$2^{K(f)+O(1)-K(x)} > f(x),$$

$$-\log m(x) = K(x) \pm O(1).$$

Теорема 4. *[?] Пусть* f *— вычислимая функция вероятности на пространстве бинарных векторов произвольной длины. Тогда*

$$0 \leq \big(\mathsf{E}_f K(X) - H(x)\big) \leq K(f) + O(1).$$

Доказательство. K(X) — это длина префиксного кода для x, по noiseless coding theorem:

$$H(X) \leq \mathsf{E}_f K(X).$$

$$f(x) \leq 2^{K(f)+O(1)}(K(x) + O(1)) :$$

$$\log \frac{1}{f(x)} \geq K(x) - K(f) - O(1) :$$

$$\sum_x f(x) K(X) \leq H(X) + K(f) + O(1).$$

□

0.2 Принцип минимальной длины описания

Рассмотрим префиксную сложность, которая описывается машиной Тьюринга специального вида, имеющих однонаправленную ленту для чтения и двунаправленную рабочую ленту. Будем полагать что машина Тьюринга T останавливается на p с выводом x: $T(p) = x$, если вся зхапись p осталась слева от читающей каретки, x осталась слева от пишущщей каретки и T остановлена.

Определение 4. *Префиксная Колмогоровская сложность:*

$$K(x) = \min_{p \in \{0,1\}^*, i \in \mathcal{N}} \{|i| + |p| : T_i(p) = x\},$$

0.3 Вероятностная интерпретация минимальной длины описания

Список литературы

[1] Успенский В., Шень А., Верещагин Н. Колмогоровская сложность и алгоритмическая случайность. – Litres, 2017

[2] Grunwald P., Vitányi P. Shannon information and Kolmogorov complexity //arXiv preprint cs/0410002. – 2004.