Исследование методов восстановления частично заданных схем из функциональных элементов

Трубицын Юрий Алексеевич

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМ. М.В. ЛОМОНОСОВА
ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И КИБЕРНЕТИКИ
КАФЕДРА МАТЕМАТИЧЕСКОЙ КИБЕРНЕТИКИ

Москва, 2017

Постановка задачи

1. Выделить набор признаков СФЭ, которые будут использоваться для решения задачи распознавания, а также реализовать и протестировать алгоритмы вычисления признаков СФЭ;
2. Построить регрессионную модель;
3. Протестировать построенную модель на примере класса мультиплексорных функций.

Обозначим $X_\Sigma = \{x_1, \ldots, x_n\}$ – множество входов схемы Σ, $Z_\Sigma = \{z_1, \ldots, z_m\}$ – множество выходов схемы Σ.
Введем также $X^* = \{x_1^*, \ldots, x_n^*, \ldots\}$ – счетный упорядоченный алфавит заходов удаленных контактов, $Z^* = \{z_1^*, \ldots, z_n^*, \ldots\}$ – счетный упорядоченный алфавит исходов удаленных контактов.

Исследование методов восстановления частично заданных схем из функциональных элементов

Определение

Частично заданной СФЭ (замаскированной СФЭ) Σ *будем называть такую схему* $\Sigma^{'}$, *которая получается путем удаления одного или нескольких ребер из исходной схемы* Σ. *При этом вершины, инцидентные удаленным ребрам помечаются некоторой переменной из множеств* X^* *и* Z^* *в зависимости от того, было ли ребро заходящим или исходящим.*

Определение

Одновыходной СФЭ будем называть такую СФЭ, у которой множество выходных вершин содержит всего одну вершину.

Исследование методов восстановления частично заданных схем из функциональных элементов

В общем виде восстановить функциональность частично заданной СФЭ невозможно, так как мы не знаем распределния на пространстве всевозможных восстановлений. Поэтому мы априорно предполагаем, что для сокрытия использовался определенный алгоритм и возникает задача классификации.
Таким образом, формальная постановка задачи звучит так: реализовать алгоритмы, на вход которых подается частично заданная СФЭ. Выходом алгоритма должно быть решение о принадлежности объекта заданному классу.
Для тестирования алгоритма в качестве идентифицируемого класса взят класс схем, реализующих мультиплексорные функции.

Исследование методов восстановления частично заданных схем из функциональных элементов

Были выделены следующие признаки:

1. доля каждого возможного функционального элемента;
2. максимальная полустепень исхода вершин, нормированная на количество контактов в СФЭ;
3. максимальная полустепень захода вершин;
4. минимальная полустепень исхода/захода вершин;
5. средняя полустепень исхода/захода вершин;
6. средняя глубина, нормированная на максимальную глубину;
7. среднее количество присоединенных переменных, нормированное на общее количество переменных;

Определение

Регрессионная модель $f(\mathbf{w}, \mathbf{x})$ *– это параметрическое семейство функций, задающее отображение*

$$f : W \times X \longrightarrow Y, \tag{1}$$

где $\mathbf{w} \in W$ *– пространтсво параметров,* $\mathbf{x} \in X$ *– пространство свободных переменных,* Y *– пространство зависимых переменных.*
Модель является настроенной (обученной) когда зафиксированы её параметры, то есть модель задаёт отображение

$$f : X \longrightarrow Y \tag{2}$$

для фиксированного значения $\bar{\mathbf{w}}$.

Исследование методов восстановления частично заданных схем из функциональных элементов

Пусть у нас множество X представлено пространством всевозможных векторов, размерность которых равной количеству признаков схем, выделенных для решения задачи распознавания.
Множество $Y = \{0, 1\}$.
Если настроенная регрессионная модель возвращает 0, значит вектор признаков, полученный из некоторой СФЭ, не попадает в область, в которой предположительно находятся вектора признаков объектов класса. Следовательно, согласно регрессионной модели, данная схема не входит в класс. В случае, когда регрессионная модель возвращает 1 означает, что вектор признаков, полученный из некоторой СФЭ, попадает в область, в которой предположительно находятся вектора признаков объектов класса.

Исследование методов восстановления частично заданных схем из функциональных элементов

Использовались следующие алгоритмы машинного обучения:

- метод опорных векторов (поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве);
- метод ближайших соседей (простейший метрический классификатор, основанный на оценивании сходства объектов; классифицируемый объект относится к тому классу, которому принадлежат ближайшие к нему объекты обучающей выборки.);
- случайный лес (алгоритм машинного обучения, заключающийся в использовании комитета (ансамбля) решающих деревьев);
- логистическая регрессия (метод построения линейного классификатора, позволяющий оценивать апостериорные вероятности принадлежности объектов классам).

Исследование методов восстановления частично заданных схем из функциональных элементов

Тестирование построенной модели проводилось на классе мультиплексорных функций.
Для проверки точности полученной модели использовался скользящий контроль.
Скользящий контроль или кросс-проверка или кросс-валидация (cross-validation, CV) – процедура эмпирического оценивания обобщающей способности алгоритмов, обучаемых по прецедентам.

Исследование методов восстановления частично заданных схем из функциональных элементов

На рис. 1 показаны результаты скользящего контроля.

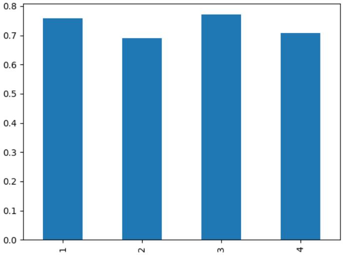

Рис. 1: 1 - случайный лес, 2 - метод ближайших соседей, 3 - логистическая регрессия, 4 - метод опорных векторов.

Исследование методов восстановления частично заданных схем из функциональных элементов

По результатам скользящего контроля можно заметить, что лучше всего себя показывают алгоритмы случайного леса и логистичекой регрессии. На них и проходило обучение контрольной выборки.

Полученные результаты

Класс	Процент удаленных проводов	Случайный лес	Логистическая регрессия
Мультиплексоры	5%	0.997996	1.0
	10%	0.997996	1.0
	15%	0.997996	1.0
	20%	0.998024	1.0
	25%	0.997996	1.0
	30%	0.998004	1.0
	35%	0.998008	1.0
	40%	0.998016	1.0
	45%	0.998028	1.0
	50%	0.998043	1.0
	55%	0.998047	1.0
	60%	0.998047	0.998054
	65%	0.998058	0.997665
	70%	0.998058	0.997005
	75%	0.998095	0.996076
Не мультиплексоры	-	0.994616	0.913862

Полученные результаты

1. Выделен набор признаков СФЭ, которые использовались для решения задачи распознавания, а также реализованы и протестированы алгоритмы вычисления признаков СФЭ;
2. Построена регрессионную модель;
3. Построенная модель протестирована на примере класса мультиплексорных функций.

Исследование методов восстановления частично заданных схем из функциональных элементов

Трубицын Юрий Алексеевич

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМ. М.В. ЛОМОНОСОВА
ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И КИБЕРНЕТИКИ
КАФЕДРА МАТЕМАТИЧЕСКОЙ КИБЕРНЕТИКИ

Москва, 2017