



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

Метод построения поисковых индексов в реляционной базе данных на основе глубоких нейронных сетей

Студент: Маслова Марина Дмитриевна ИУ7-83Б
Руководитель: Оленев Антон Александрович

Москва, 2023

Цель и задачи

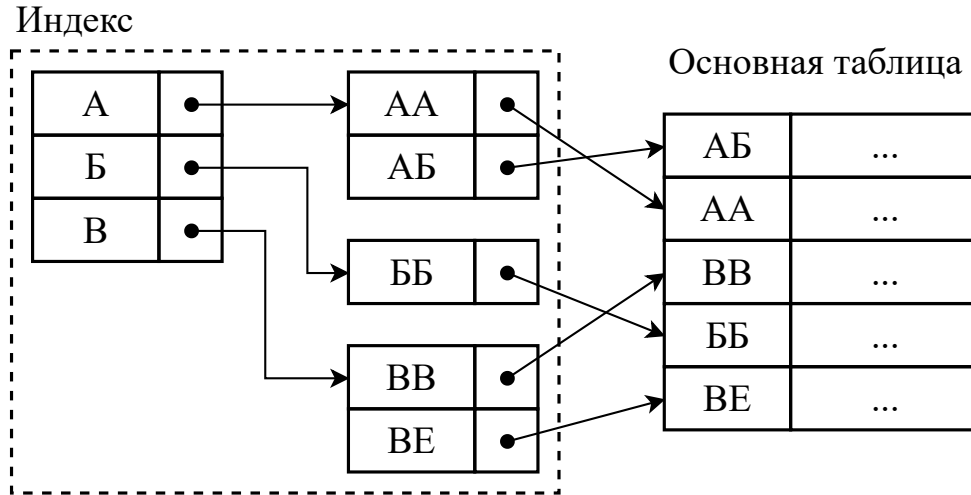
Цель: разработка метода построения поисковых индексов в реляционной базе данных на основе глубоких нейронных сетей.

Задачи:

- рассмотреть и сравнить известные методы построения индексов;
- привести описание построения индексов с помощью нейронных сетей;
- разработать метод построения индексов в реляционной базе данных на основе глубоких нейронных сетей;
- разработать программное обеспечение, реализующее данный метод;
- провести исследование (по времени и памяти) операций поиска и вставки с использованием индекса, построенного разработанным методом, при различных объемах данных.

Классические структуры индексов

Индекс — это некоторая структура, обеспечивающая быстрый поиск записей в базе данных за счет определяет соответствие ключа поиска конкретной записи с положением этой записи.

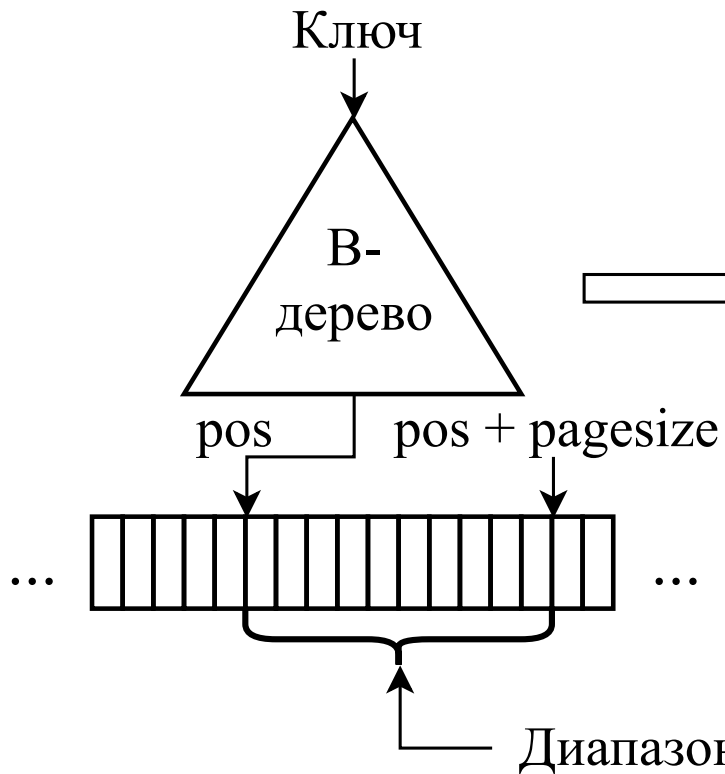


Основные типы структур:

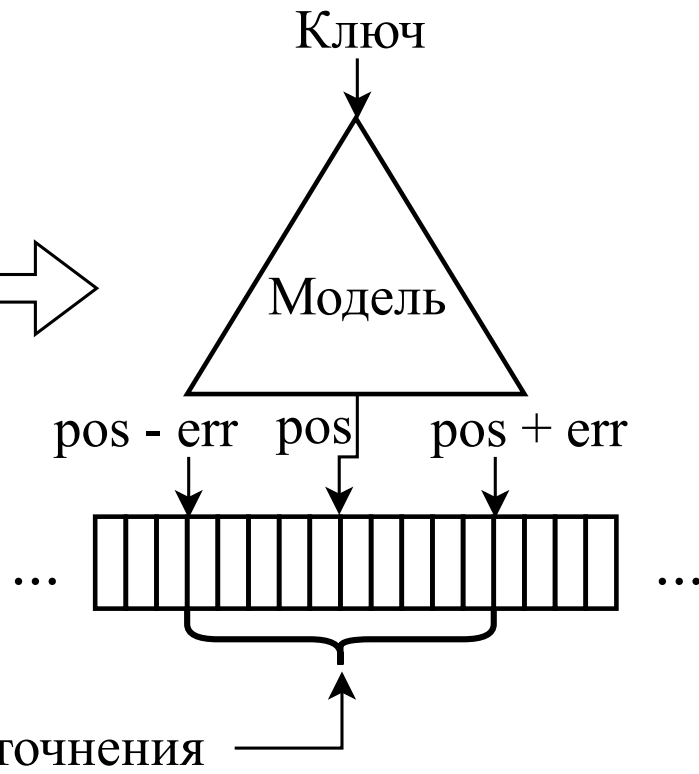
- В-деревья;
- хеш-таблицы;
- битовые карты.

Обученные индексы

Индекс на основе В-дерева



Обученный индекс

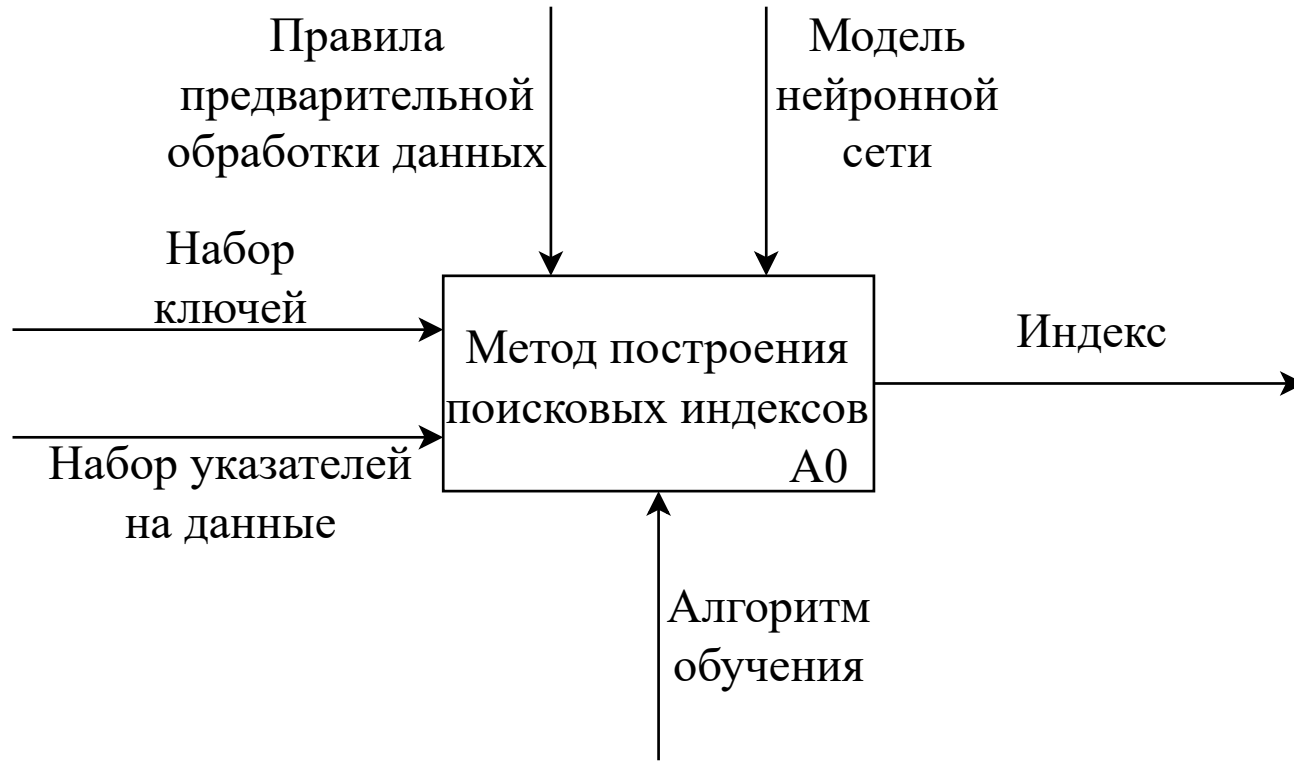


Сравнение методов построения индексов

Метод		Классические индексы			Обученные индексы
		В-дерево	Хеш-таблица	Фильтр Блума	
Временная сложность	поиска	$O(\log N)$	$O(1) / O(N)$	$O(k)$	$O(1) / O(N)$
	вставки	$O(\log N)$	$O(1) / O(N)$	$O(k)$	(*)
Память		Высокая	Средняя	Низкая	Средняя
Поиск в диапазоне		+	-	-	+
Поиск единичного ключа		+	+	-	+
Проверка существования		+	+	+	+

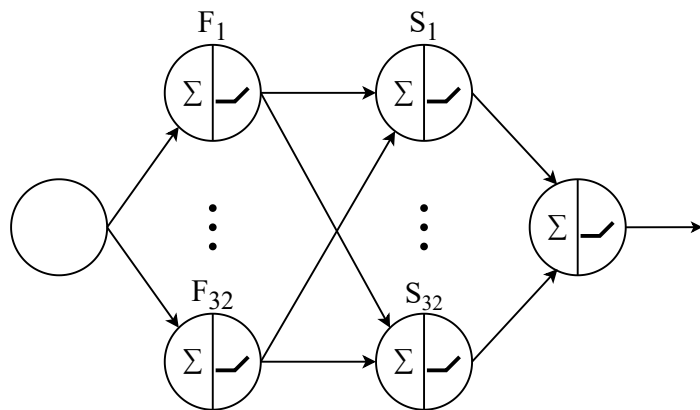
(*) — вставка в обученный индекс требует переобучения, сложность которого зависит от архитектуры используемой модели машинного обучения.

Постановка задачи



Ограничение: ключи — целые уникальные числа.

Архитектура нейронной сети

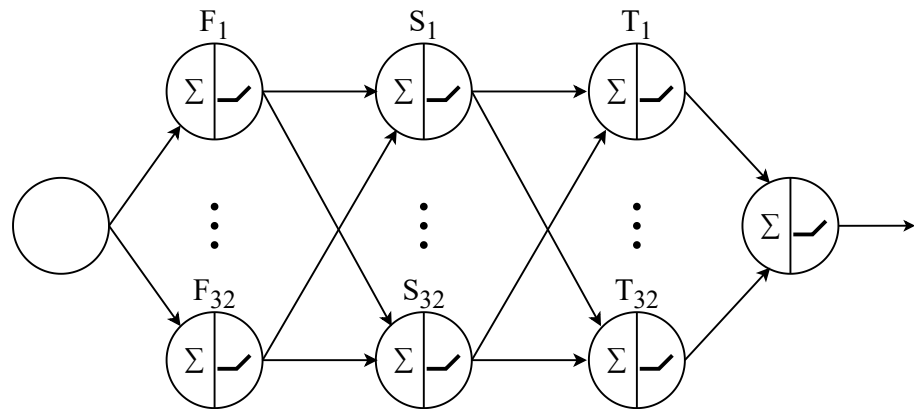


1 вход: нормализованный ключ K .
1 выход: значение функции распределения $F(K)$.
2-3 скрытых слоя по 32 нейрона.
Функция активации: ReLU.

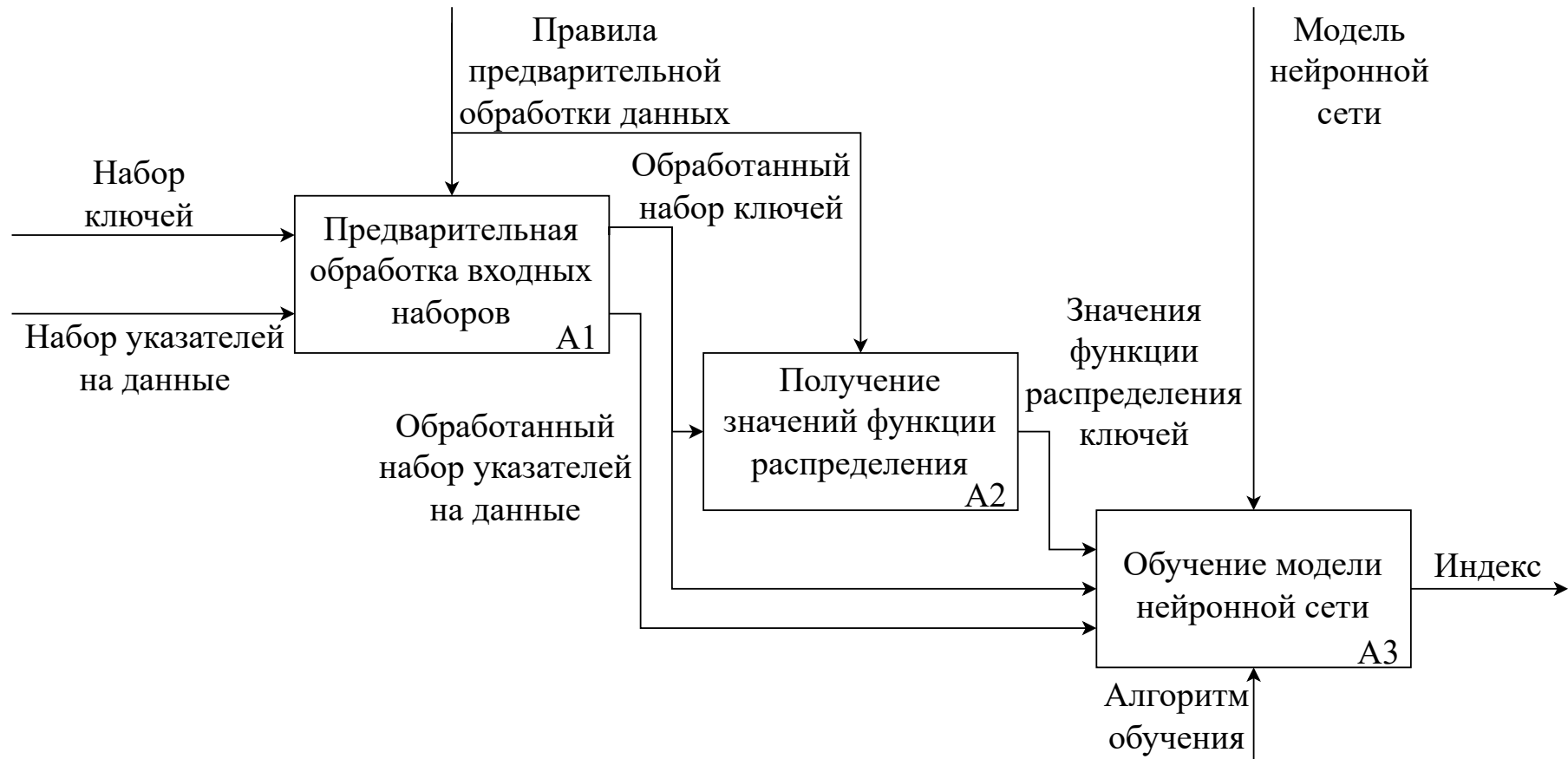
Определение значения функции распределения по позиции ключа:

$$F(K) = \frac{p}{N},$$

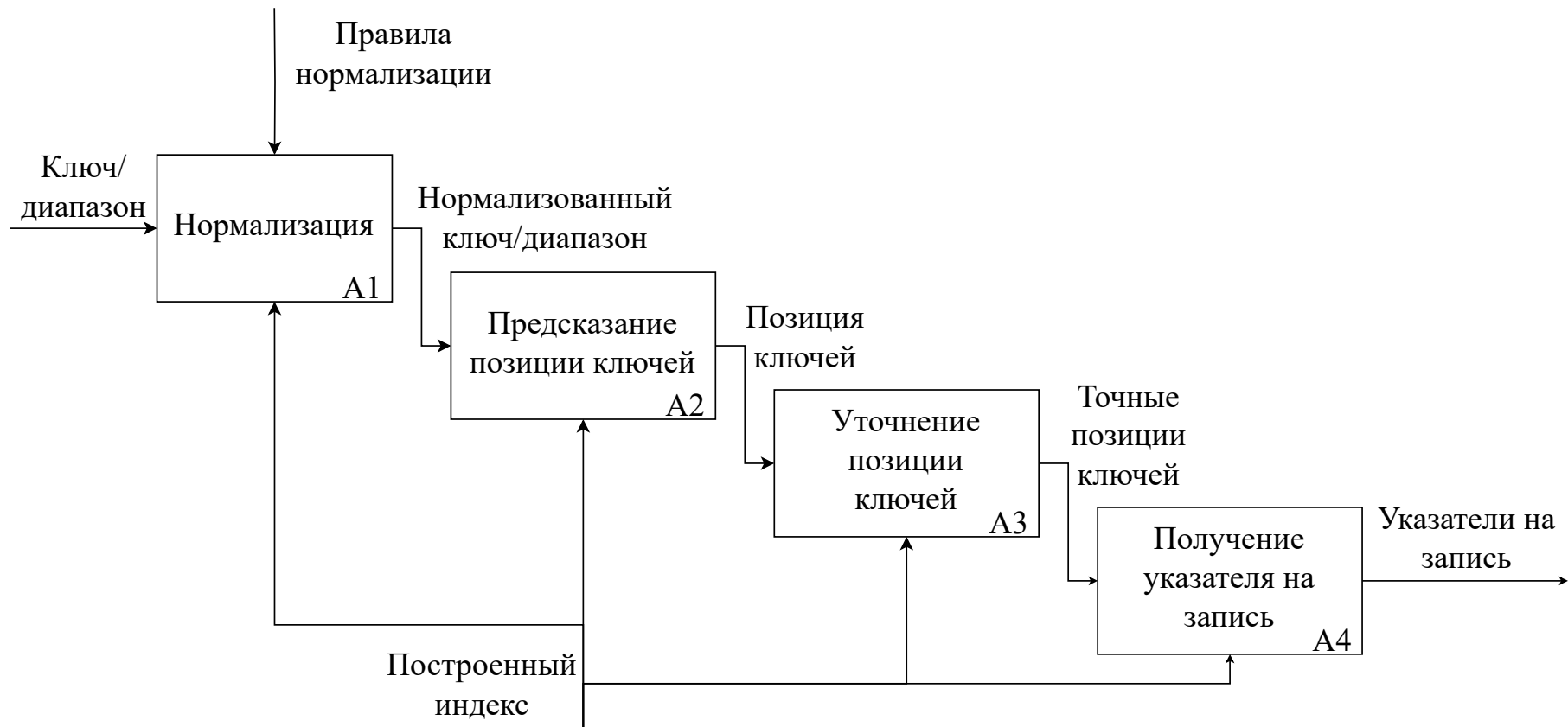
где p — искомая позиция;
 K — ключ поиска;
 $F(K)$ — функция распределения;
 N — количество ключей.



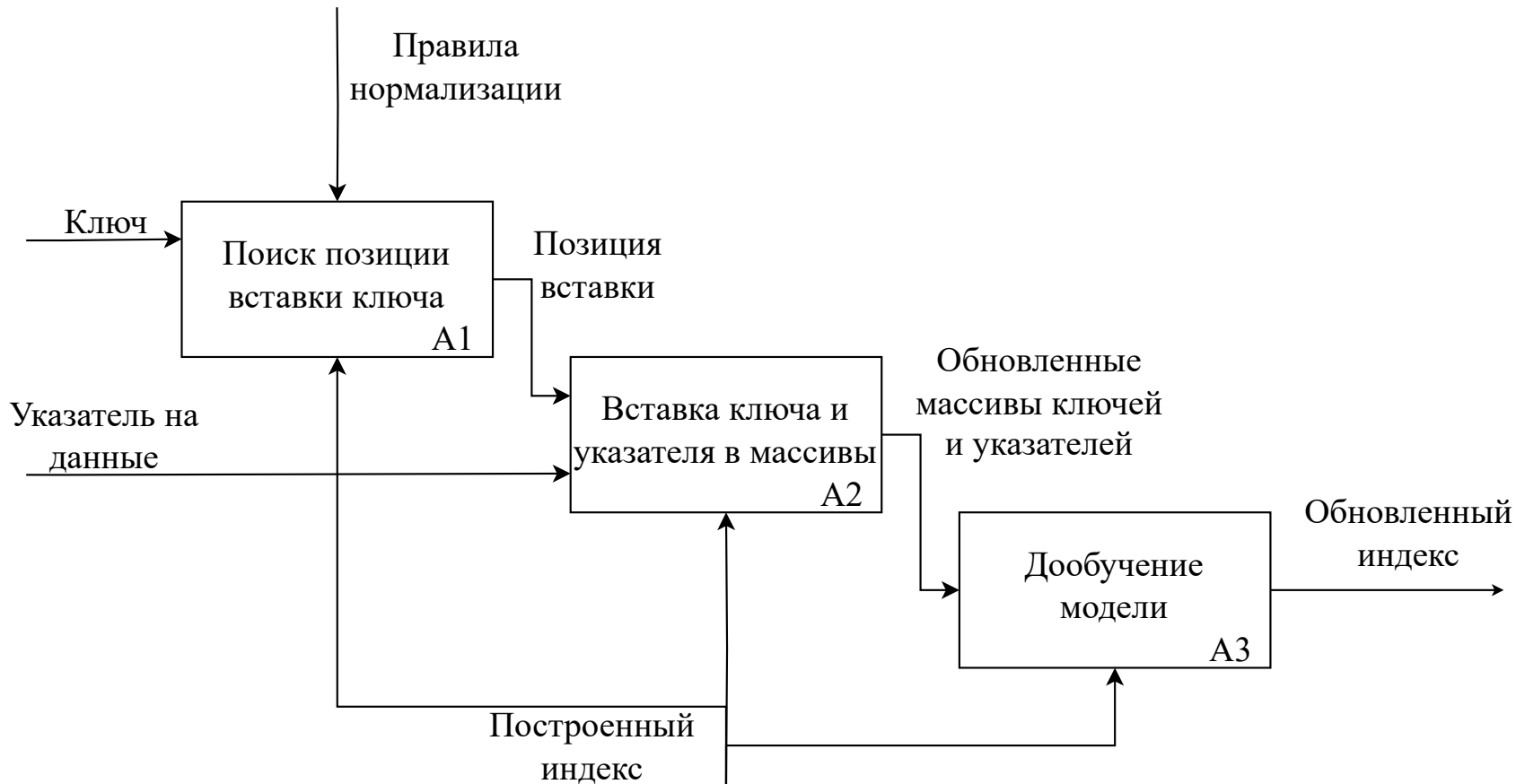
Функциональная схема построения индекса



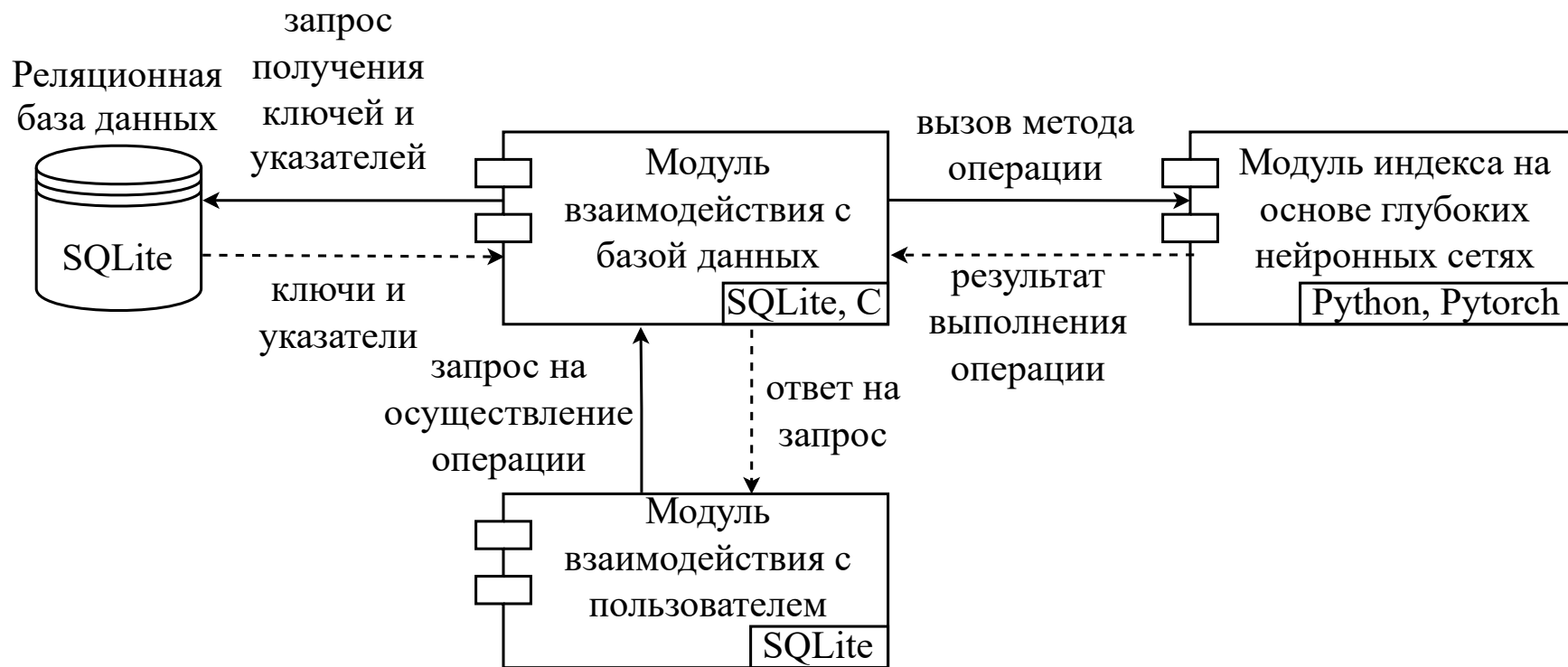
Функциональная схема поиска



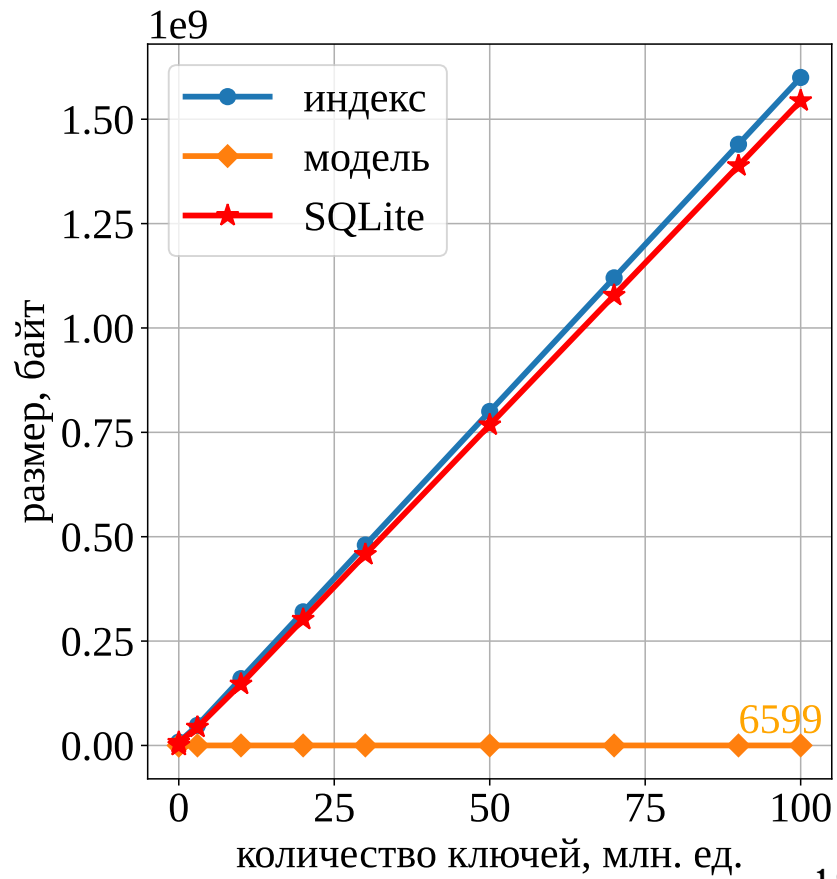
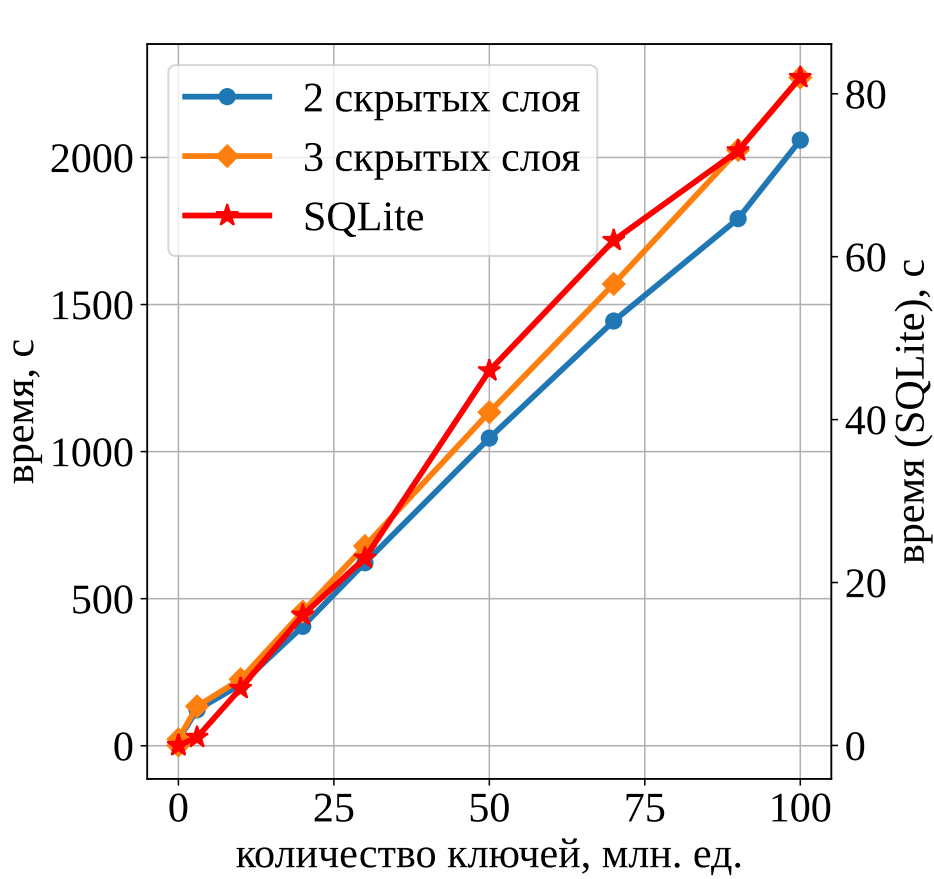
Функциональная схема вставки



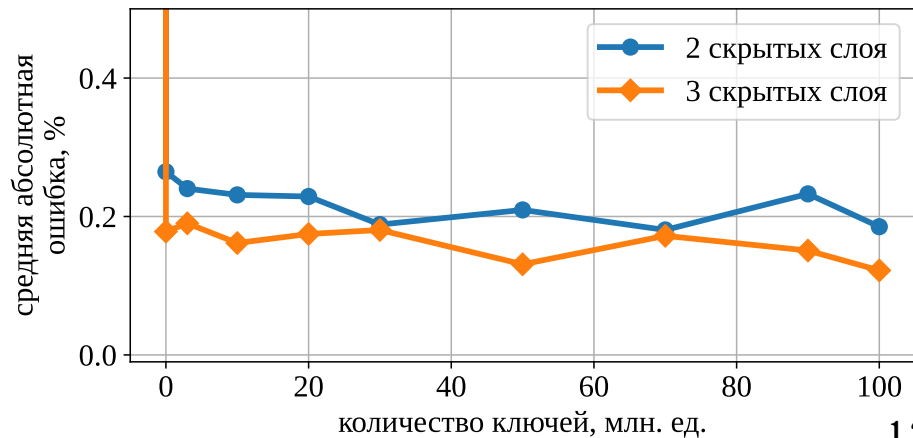
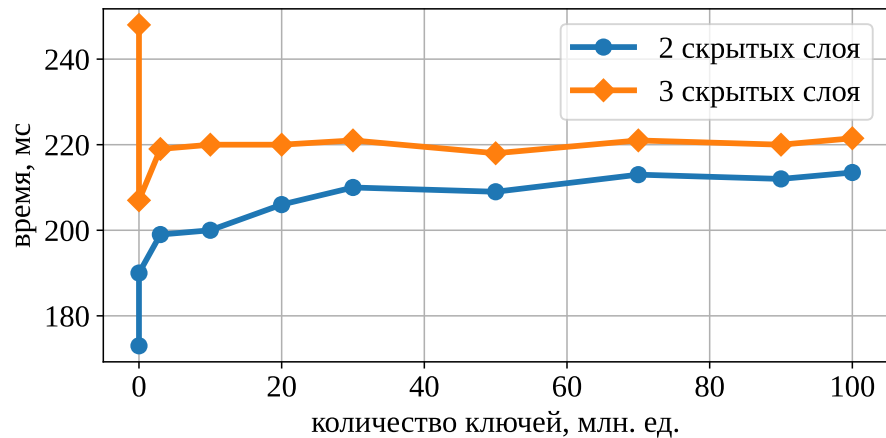
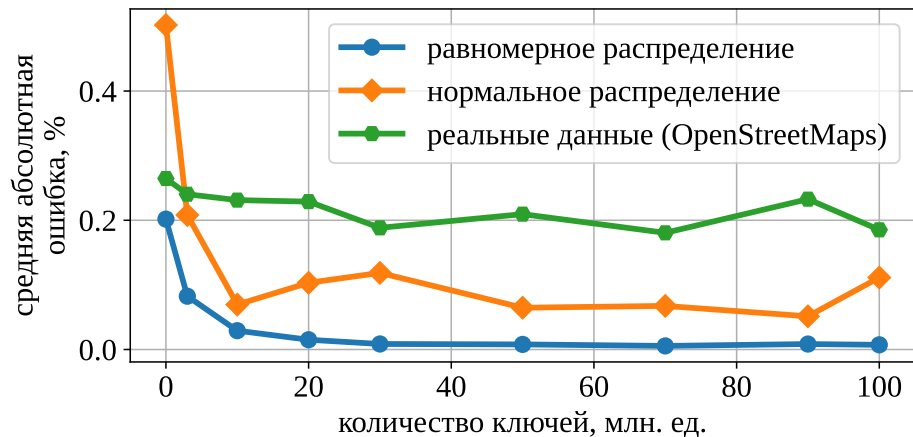
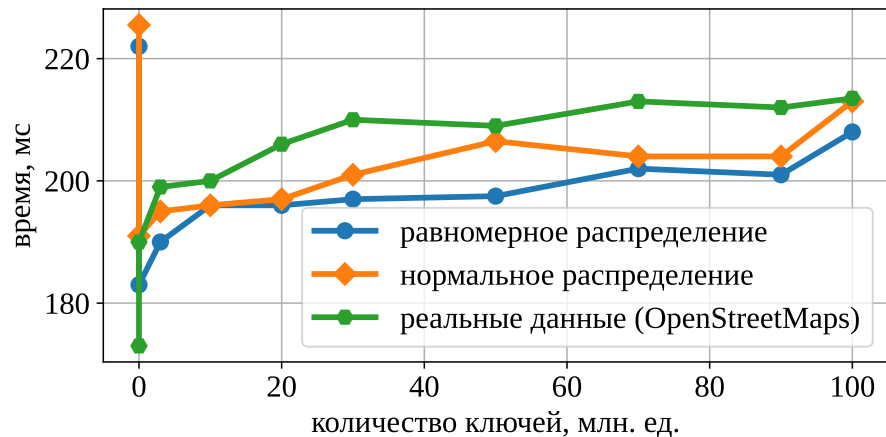
Структура программного обеспечения



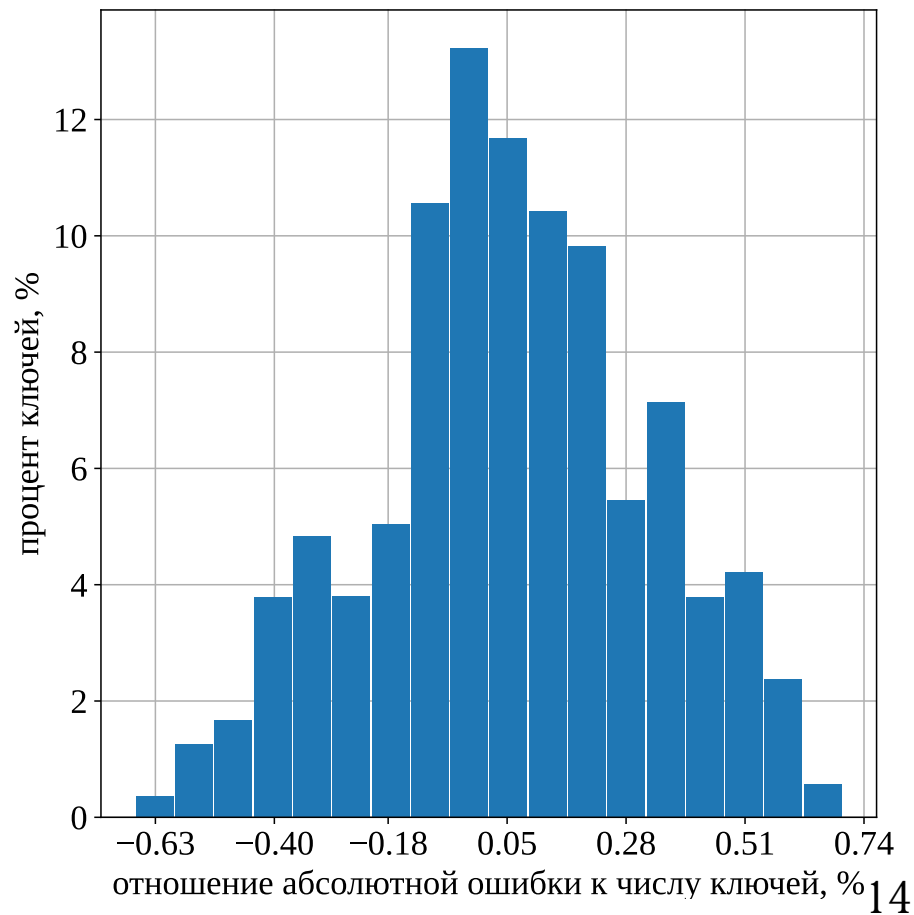
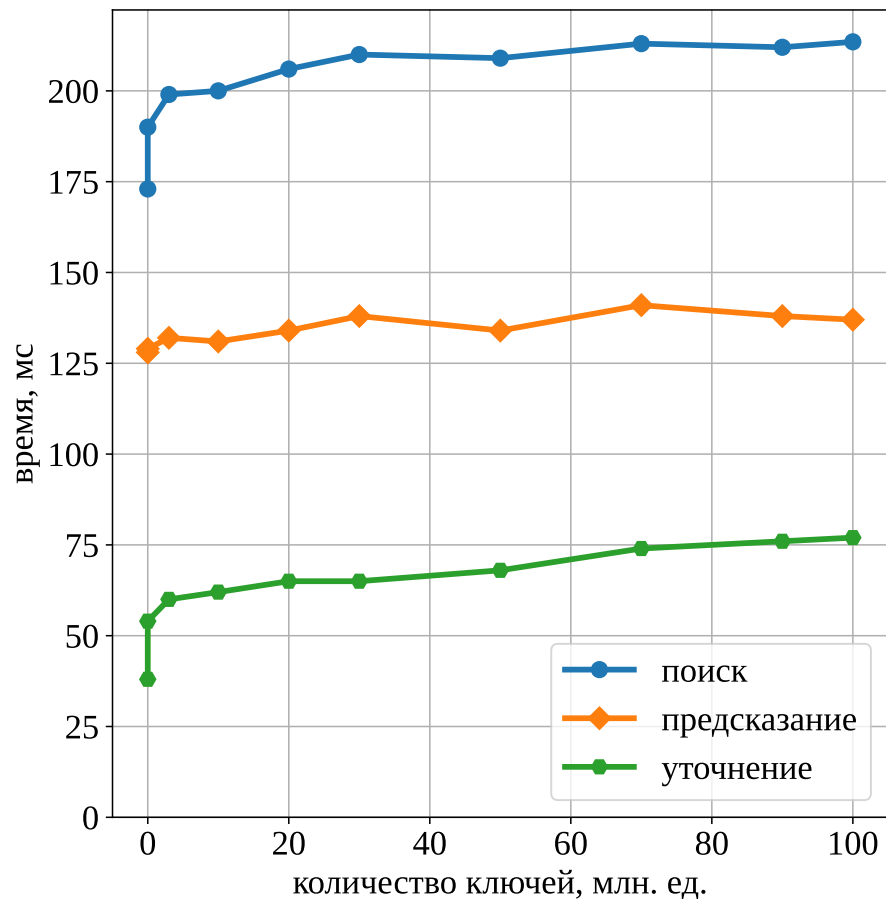
Исследование времени построения и размера индекса



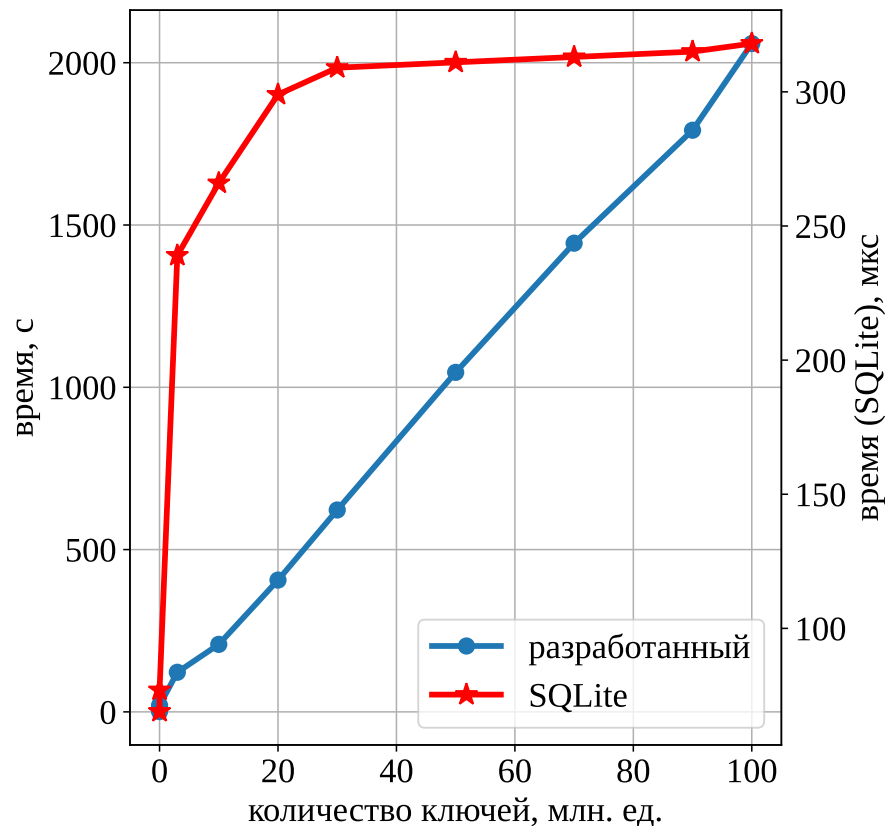
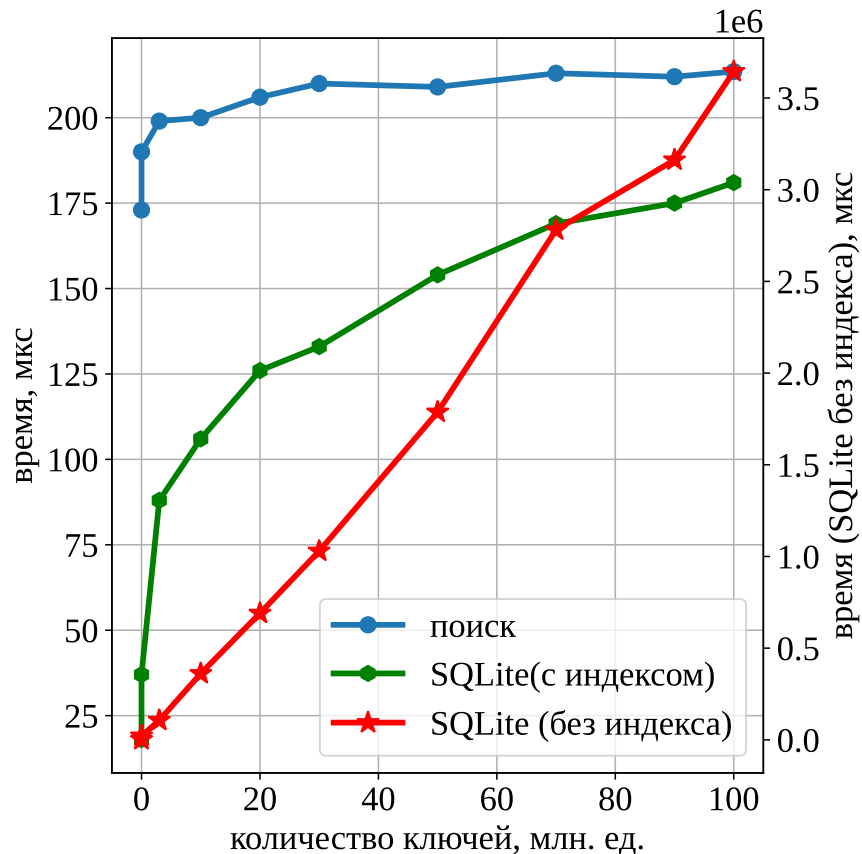
Исследование времени поиска



Исследование времени поиска



Исследование времени поиска и вставки



Заключение

В ходе данной работы:

- проанализированы известные методы построения индексов;
- приведено описание построения индексов с помощью нейронных сетей;
- разработан метод построения индексов в реляционной базе данных на основе глубоких нейронных сетей;
- разработано программное обеспечение, реализующее данный метод;
- проведено исследование (по времени и памяти) операций поиска и вставки с использованием индекса, построенного разработанным методом, при различных объемах данных.

Поставленная цель достигнута.

Дальнейшее развитие

1. Оптимизация алгоритма вставки с учетом распределения ключей.
2. Добавление возможности построения индекса по ключам других типов данных.
3. Построение многомерных обученных индексов.