

Метод оценки качества перевода текста в существующих информационных системах

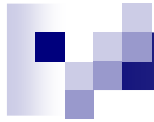
Довгопольый Дмитрий. 604 гр.

2011



Введение

- Актуальность
- Существующие методы
 - BLEU
 - Meteor
 - X-score
- Их недостатки:
 - Разработаны для аналитических языков
 - Необходимы выровненные и размеченные корпуса текстов



Цель работы

Создание и реализация метода оценки качества перевода в автоматических системах для улучшения качества оценивания автоматически генерируемых текстов на русском языке.



Задачи

- Изучение существующих методов, оценивающих качество машинного перевода
- Разработка метода, оценивающего естественность текста, с применением корпуса текстов.
- Создание программного комплекса на основе разработанного алгоритма
- Оценка качества результата и особенностей работы предложенного алгоритма



Метод оценки качества перевода

- Расширенная биграммная модель
- Стоп-слова
- Выбор и анализ параметров метода

$$P(a,b) = \sum_{i=0}^2 \frac{1}{2^i} \frac{k_i(a,b)}{\max(k(a), k(b))}$$

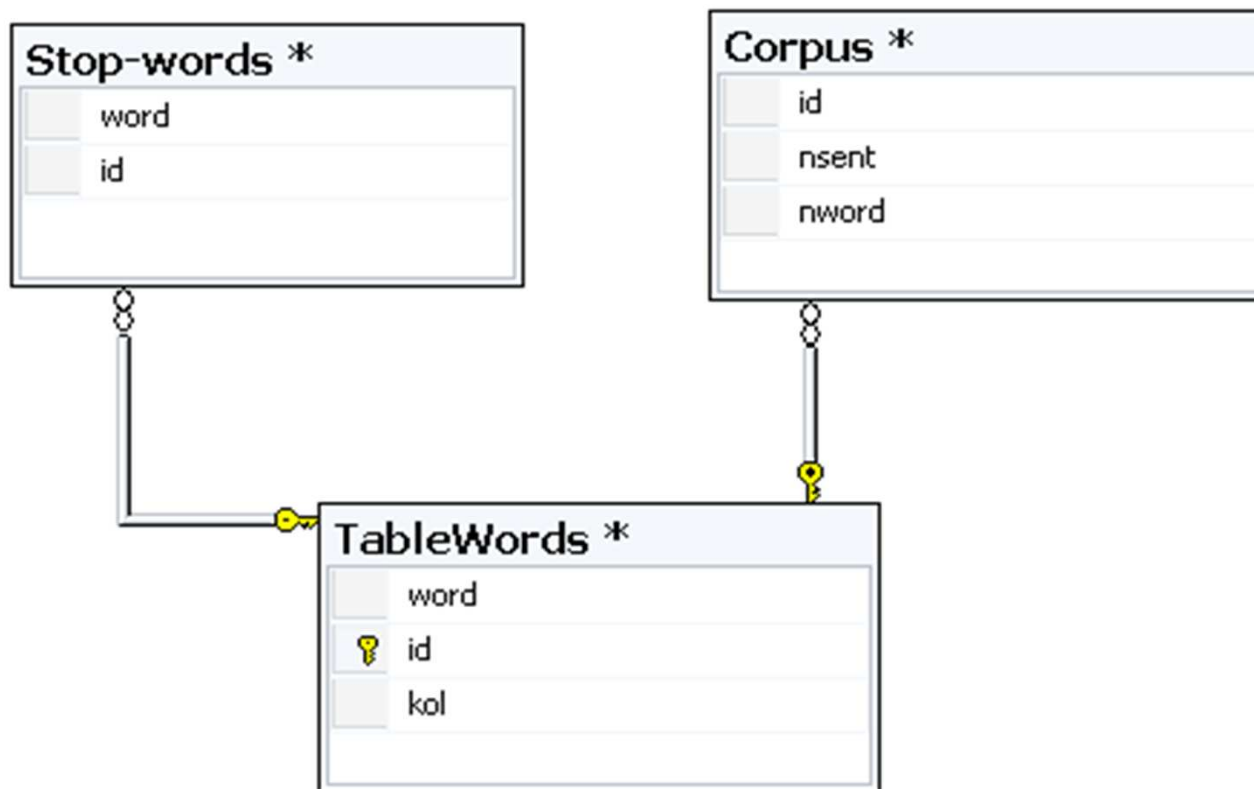


Практическая реализация

Microsoft SQL Server

- надежность и масштабируемость
- высокая производительность
- интеграции с другими продуктами компании Microsoft
- встроенные системы профилирования производительности

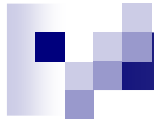
Схема данных





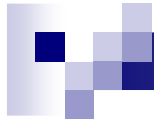
Клиентское приложение

- Ручной ввод данных
- Ввод из файла
- Загрузка большого объема данных для тестирования



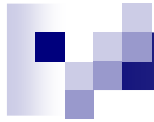
Тестирование

- Корреляция с оценками экспертов
- Нагрузочное тестирование



Результаты

- Метод анализа качества перевода текстов
- Программный комплекс
- Тестирование



Спасибо за внимание