

Воронежский институт высоких технологий - автономная некоммерческая образовательная организация высшего образования (ВИВТ-АНОО ВО)

О проблемах идентификации именованных объектов и извлечения связей при анализе научных и технических русскоязычных текстов

аспирант Меняйлов Д.В.
группа А-201

Семантические отношения

- Отношение причинности, например, [взаимодействие высокоэнергетических пучков : деформация]
- Отношение сравнения, например, [реляционные базы данных: объектно-ориентированные базы данных]
- Отношение таксономии, например, [Python : язык программирования]
- Отношение мерономии, например, [модуль : система]
- Отношения синонимии, например, [GPU : графический процессор]
- Отношение использования, например, [метод статистической обработки : анализ текста]

Размеченные связи между терминами

Причинность	27
Сравнение	23
Таксономия	92
Мерономия	79
Синонимия	24
Использование	387

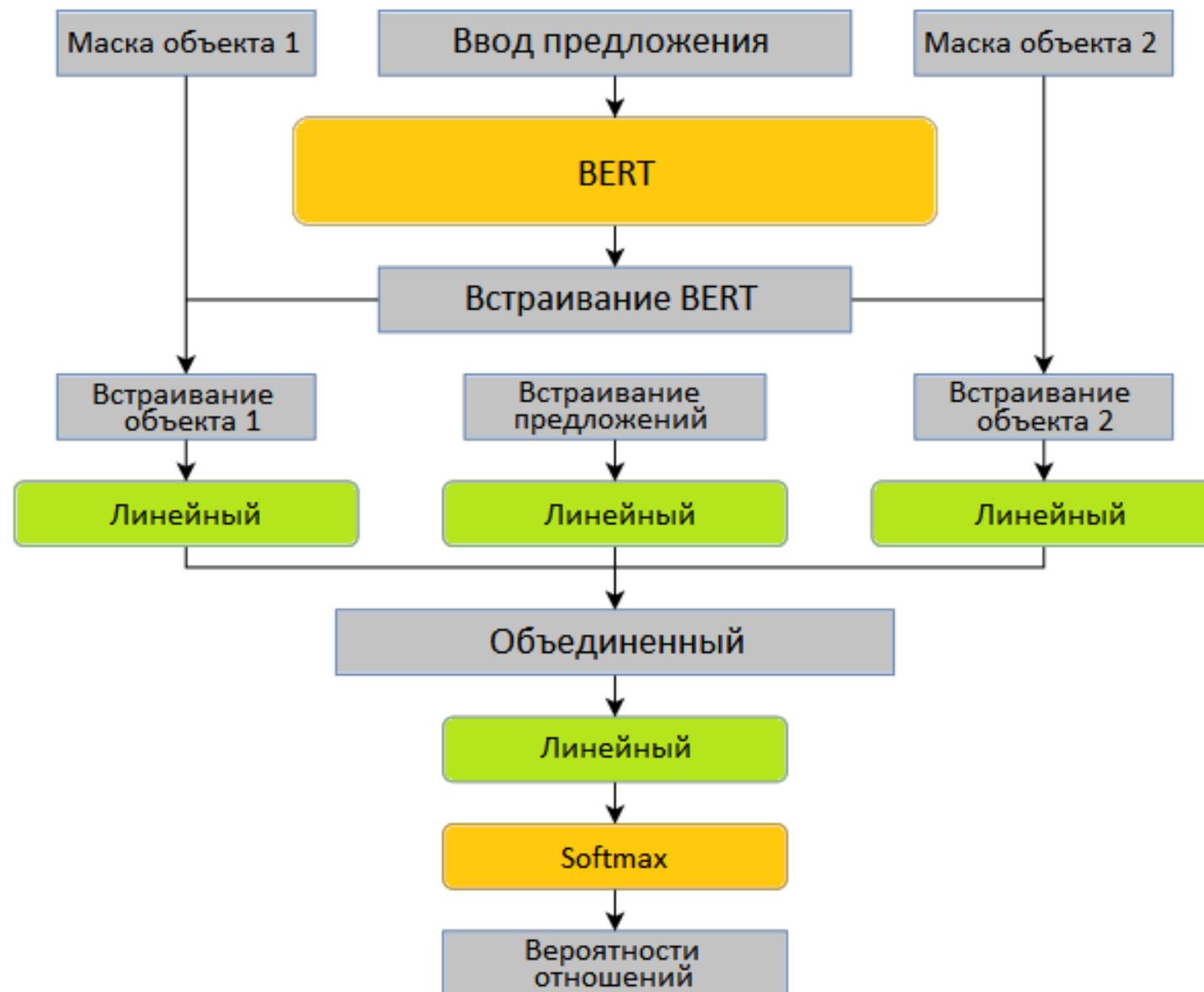
СООТВЕТВИЕ ТОКЕНАМ

Точное сопоставление токенов			
Метод	Точность	Отзыв	F-оценка
По словарю	0,23	0,15	0,18
Комбинированный	0,17	0,11	0,13
RAKE	0,34	0,26	0,3
Оптимизированный RAKE	0,42	0,33	0,37
Нечеткое сопоставление токенов			
По словарю	0,8	0,32	0,46
Комбинированный	0,8	0,26	0,4
RAKE	0,6	0,61	0,61
Оптимизированный RAKE	0,63	0,55	0,59

BERT

- BertLinearER
- BertLstmLinearER, BertCnnLinearER
- BertLstmCrfER, BertCnnCrfER
- BertRC

Архитектура модели для классификации отношений



СРАВНЕНИЕ МОДЕЛЕЙ НА ОСНОВЕ BERT

Модель	F-оценка
LinearER	0,518
LstmLinearER	0,528
CnnLinearER	0,494
LstmCrfER	0,52
CnnCrfER	0,501

ПОКАЗАТЕЛИ КЛАССИФИКАЦИИ ОТНОШЕНИЙ

Отношение	F-оценка
ПРИЧИНА	0,73
СРАВНЕНИЕ	0,65
ISA	0,88
ЧАСТЬ	0,62
СИНОНИМЫ	0,98
ИСПОЛЬЗОВАНИЕ	0,86