

Зимняя школа по параллельному программированию 2019

Применение методов data mining'а к разговорным базам

Участник:

Дроздов Э.К. – НГТУ, ФПМИ, 4 курс

Руководитель:

Городничев М.А. – младший научный сотрудник МО ВВС

03.07.17 – 14.07.17

г. Новосибирск

Содержание

- Проблема
- Актуальность
- Идея решения
- Ожидаемый результат

Проблема

- Большой объём разнородных данных
- Несовершенство производственного конвейера
- Актуальность: сбор и анализ статистики данных позволит понять наиболее часто встречающиеся ошибки на каждом уровне конвейера и принять решение, что именно можно усовершенствовать

Идея решения

- Применить методы data mining'a ко множеству данных об ошибках в разговорах с целью выявления классов ошибок и разработки эффективных методов их устранения.

Результат

Можно выделить следующие классы ошибок:

- stt -- во время разговора фраза человека была неверно обработана stt-распознавателем
- грамматик -- фраза человека верно преобразовалась в текст, но из неё не выделились ожидаемые факты, либо выделились лишние
- реализация скрипта -- верно отработал stt, грамматики выделили верные факты из фразы, но была вызвана неверная реакция
- скриптов -- была вызвана реакция, соответствующая клиентскому скрипту диалога, но семантически неадекватная

Спасибо за внимание!