KMY VITMO

ХІІ КОНГРЕСС МОЛОДЫХ УЧЕНЫХ



АРХИТЕКТУРА МОДУЛЯ ЯЗЫКОВОЙ МОДЕЛИ ДЛЯ ПРЕДСКАЗАНИЯ ПРЕДВАРИТЕЛЬНОГО ДИАГНОЗА ПАЦИЕНТА ПРИ ВЗАИМОДЕЙСТВИИ С ИСПОЛЬЗОВАНИЕМ ДИАЛОГОВОГО АГЕНТА

Подготовили: Егоров М.П., Погребной Д.А.

Руководитель: Чижик А. В.

Предварительный диагноз пациента

- Первый прием пациента в медицинской организации
 - Сбор анамнеза
 - Постановка предварительного диагноза
- Оба процесса требуют много ресурсов
- Большой поток пациентов перегружает врачей
- Оба процесса могут быть автоматизированы





Существующие решения

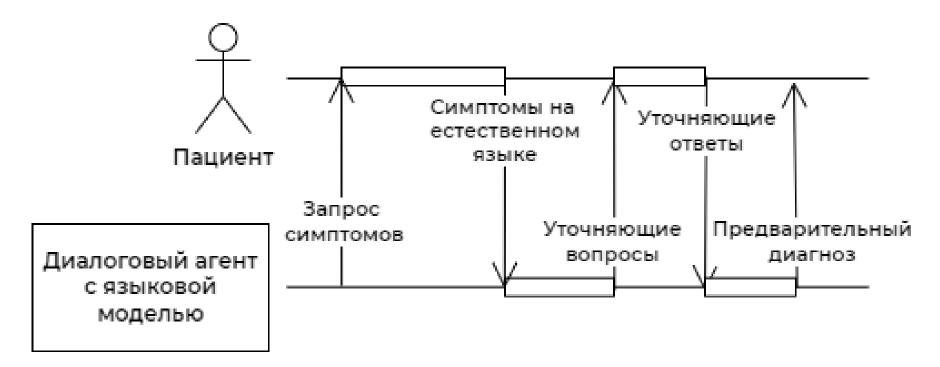
- Справочные системы для поиска диагноза по заданным симптомам
 - Карта симптомов Mail.ru
- Примитивные диалоговые агенты на правилах и ветвлениях
 - Чат-бот Сбер Здоровье
- Проприетарные решения для определения диагноза по списку симптомов
 - Сервис «Симптомчекер» от Webiomed
- Существующие решения не поддерживают работу с естественным языком





Языковая модель с диалоговым агентом

• **Цель**: Разработать языковую модель и диалоговый агент, которые позволят в режиме диалога собирать анамнез у пациента и предоставлять предварительный диагноз на базе собранных данных.



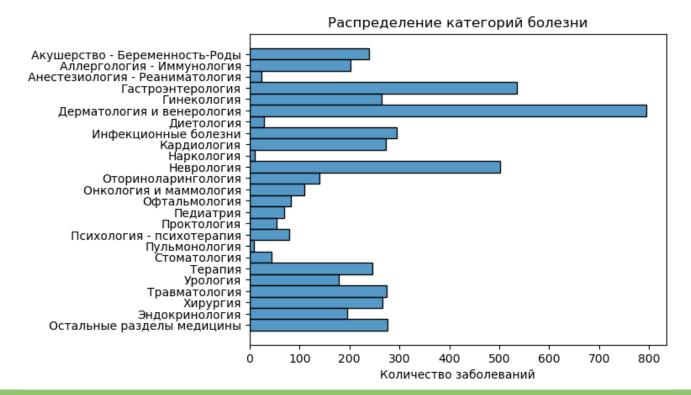






Датасет

- Данные для обучения и валидации
 - Априорные знания о болезнях и их симптомах
 - Собранный датасет симптомов и болезней на <u>Kaggle</u> 5193 записи

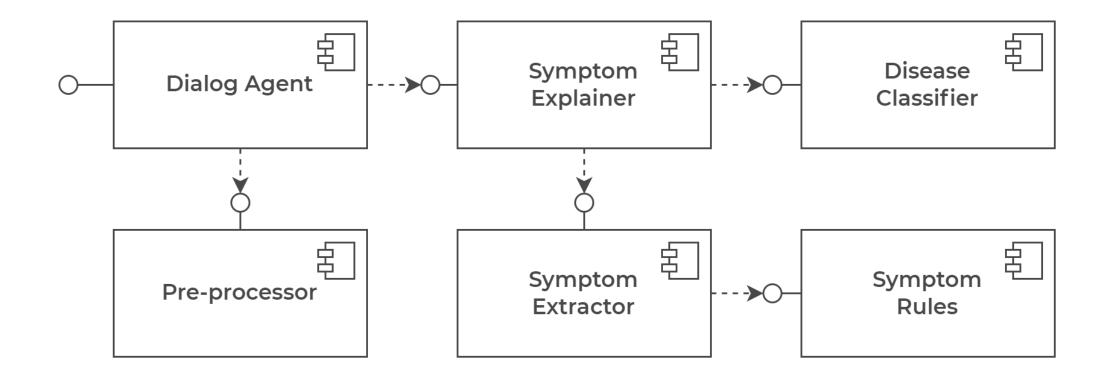








Архитектура









Шаг 1 - Сбор симптомов пациента

- 1. Запрос симптомов у пациента на естественном языке
- 2. Предварительная обработка
 - Удаление знаков пунктуации
 - Лемматизация слов
- 3. Выделение симптомов из текста
 - 1. SpaCy библиотека
 - 2. Выделение симптомов по заданным правилам 5213 симптома
- 4. Обработка отрицаний симптомов с помощью библиотеки Negspacy
- 5. Получаем векторное представление симптомов пациента







Шаг 2 – Определение предварительного диагноза

- ML модель предсказывает предварительный диагноз
- При предсказании диагноза важна интерпретируемость
 - Логистическая регрессия
 - Что-то более сложное
- Если модель не уверена в предсказании, то агент уточняет наличие необходимых симптомов
- Первые метрики:
 - Accuracy предсказания болезни без дополнительных запросов 0,25







Заключение

- Разработан прототип языковой модели для определения предварительного диагноза пациента по симптомам на естественном языке
- Опубликована альфа версия рір пакета с инструментом
 - pypi.org/project/distool/
- Открытый исходный код
 - github.com/NIRMA-PATIENT-INTAKE/disease/





Дальнейшие планы

- Замерить метрики качества инструмента
- Поддержать работу с большем количеством симптомов
- Улучшить точность распознавания симптомов
- Провести эксперименты с более сложными моделями







KMY VITMO

kmu.itmo.ru

