

ИССЛЕДОВАНИЕ НЕЙРОННЫХ СЕТЕЙ, ИСПОЛЬЗУЕМЫХ В ЧАТ-БОТАХ

Муратова У.Д.

(Университет ИТМО)

Научный руководитель – к.т.н. Балакшин П.В.

(Университет ИТМО)

В работе рассматриваются различные варианты нейронных сетей для обработки текстов. Также выбирается наиболее подходящая по характеристикам нейронная сеть для встраивания в прототип чат-бота.

Введение. Последние несколько лет в мире прослеживается тенденция автоматизации рутинных действий, особенно это заметно в сфере технологий и обслуживания. Все чаще при звонке на горячую линию, общении с компаниями в мессенджерах человеку отвечает не человек, а специально созданная программа, все больше человек использует голосовое управление какой-либо техникой. Это обусловлено простотой и удобством использования, увеличением производительности: например, если в call-центре использовать робота, который будет автоматически принимать звонки и решать основные задачи, только в особых случаях переводя звонок на оператора, загрузку на людей, работающих в таком call-центре, и вообще количество людей можно будет сократить.

В современном мире можно встретить два типа чат-ботов – это чат-боты, в основе алгоритма работы которых используется готовый словарь фраз, с которыми этот бот способен работать, и более умные, использующие нейронную сеть. У каждого из подходов есть достоинства и недостатки.

Первый тип чат-ботов, основанный на готовых словарях, работает по следующему алгоритму: он получает некоторую человеческую фразу, ищет её в словаре, находит соответствующий такому запросу ответ и выводит его пользователю. Плюсы такого подхода – это простота и скорость реализации. Зачастую используется такой подход в тех случаях, когда есть типовый набор вопросов и ответов, и он не меняется, например, для ботов – сборщиков отзывов или ботов, принимающих заказы. Однако, у данного подхода существует два значительных минуса. Первый – это существенное увеличение словарей в случае увеличения объема фраз, что влияет на объемы памяти, которую занимает приложение, и времени, необходимого для поиска фраз. Второй – это ограниченность фраз и конструкций, которыми приложение способно будет оперировать. Для чат-ботов, целью которых является общение с человеком, это недопустимо. Чат-бот должен уметь по-разному отвечать на один и тот же вопрос и понимать суть одного и того же вопроса, заданного по-разному, иначе пользователь потеряет интерес к такому приложению.

Выходом в такой ситуации может являться нейронная сеть, интегрируемая в разрабатываемый чат-бот, которая сможет повысить разнообразие понимаемых запросов пользователя и вариантов ответа самого чат-бота. Плюсом такого подхода будет вариативность диалогов и возможность их приближения к реальности. Проблемами данного варианта могут стать усложнение алгоритма чат-бота и обучение нейронной сети, для чего необходимы ресурсы и время.

Целью данной работы является исследование различных нейронных сетей и выбор оптимальной для встраивания в прототип чат-бота.

Предполагаемые исследования. Для достижения поставленной цели необходимо выполнить следующие задачи:

- выбрать типы нейронных сетей, которые предназначены для работы с текстами;

- на основе характеристик выбранных сетей сделать выводы о наиболее подходящих для встраивания в приложение, работающее на смартфоне, и для использования в облаке;
- определить способы взаимодействия выбранной нейронной сети и системы распознавания речи, используемой в чат-боте;
- на основе характеристик нейронной сети составить список ресурсов, необходимых для обучения выбранной сети;
- составить наборы данных для обучения выбранной нейронной сети.

В результате исследования планируется получить готовое окружение для обучения нейронной сети, нацеленной на работу с текстом, и пути интеграции готовой сети для тестирования в чат-боте.

Муратова У.Д. (автор)

Балакшин П.В. (научный руководитель)