

Министерство образования Республики Беларусь

Учреждение образования

Белорусский государственный университет

информатики и радиоэлектроники

УДК 004.891.2

Клепец

Андрей Владимирович

Обработка естественного языка на примере виртуального собеседника,
разработанного на основе языка разметки AIML

АВТОРЕФЕРАТ

на соискание степени магистра информатики и вычислительной техники
по специальности 1-40 80 04 «Математическое моделирование численные
методы и комплексы программ»

Научный руководитель

Пешков Анатолий Тимофеевич

кандидат технических наук, доцент

Минск 2015

ВВЕДЕНИЕ

Искусственный интеллект является одним из перспективных направлений развития науки. Можно выделить два направления развития ИИ: первая – решение проблем, связанных с приближением специализированных систем ИИ к возможностям человека, и их интеграции, которая реализована природой человека; вторая – создание искусственного разума, представляющего интеграцию уже созданных систем ИИ в единую систему, способную решать проблемы человечества.

Но в настоящий момент в области искусственного интеллекта наблюдается вовлечение многих предметных областей, имеющих скорее практическое отношение к ИИ, а не фундаментальное. Многие подходы были опробованы, но к возникновению искусственного разума ни одна исследовательская группа пока так и не подошла. На фоне вышеперечисленного стоит отметить создание экземпляров ИИ наиболее приблизившихся к понятию разума. Как пример можно привести IBM Watson – перспективная разработка IBM, способная воспринимать человеческую речь и производить вероятностный поиск, с применением большого количества алгоритмов. Кроме того к образцам ИИ можно отнести чатботов участвующих ежегодно в премии Лебнера. Они могут общаться с людьми и поддерживать беседу так, что собеседник не догадывается что общается с машиной. Некоторый аналог такого чатбота рассматривается в данной работе. Чатбот, созданный в этой работе, не может претендовать на прохождение теста Тьюринга, но может выступать в качестве консультанта на онлайн ресурсах. В качестве исходных данных используется существующая платформа для создания чатботов и база знаний вопросов – ответов по теме краткосрочных кредитов. Целью работы является проверка возможности создания чатбота, способного выступать в качестве консультанта при общении с людьми на веб ресурсах. Результатом будет сам чатбот, результат его обучения и тестирования.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Целью работы ставилась проверка на практике возможности создания чат-бота онлайн консультанта способного отвечать на вопросы пользователей касающиеся определенной темы.

В качестве задач можно выделить выбор платформы и обоснование сделанного выбор, преимущества и недостатки выбранной платформы. создание экземпляра чатбота, выбор содержания базы знаний по-умолчанию, загрузка имеющейся базы знаний, проведение тестирования, корректировка и доработка базы знаний, проведение тестирования созданной сущности, подведение итогов.

Чатботы довольно распространены в интернете, на сайтах онлайн магазинов и часто используются в качестве консультантов.

В данной работе производится создание чатбота, выбор и обоснование программных средств, тестирование созданного экземпляра и приводится заключение о результате проделанной работы.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой части работы описываются основные направления использования аналогичных программных продуктов, дается характеристика каждому из направлений.

Во второй части приведена история исследования темы человеческого и машинного интеллекта, исследований Тьюринга и других работ.

В третьей части приводится обоснование выбора платформы, ее краткое описание. Приводятся основы языка разметки AIML, описывается его алгоритм поиска соответствия, разбираются основные структурные единицы: теги, предикаты, свойства и их предназначение. Рассмотрены возможности платформы такие как обучение через диалог и добавление новых категорий. В работе описано создание экземпляра чатбота, изменение его свойств, загрузка базы знаний, проведено тестирование созданного экземпляра. Проведена доработка база знаний для улучшения релевантности ответа, произведено повторное тестирование.

В заключении проанализированы достоинства и недостатки созданного экземпляра.

ЗАКЛЮЧЕНИЕ

В данной магистерской работе создан экземпляр чатбота онлайн-консультанта способного отвечать на вопросы пользователей. Чатбот имеет открытое API и при необходимости может быть встроено в любое веб-приложение. Задачей было создание консультанта, который способен релевантно отвечать на введенные запросы. Чатбот консультанта был создан протестирован и обучен. Результатом обучения стал невысокий процент выдачи ответа «по умолчанию», в итоге меньше двадцати процентов от общего количества вопросов из имеющейся базы. Выбранную платформу и чатбота в частности можно использовать для аналогичных консультационных целей в других сферах. Для этого нам необходима база знаний с набором информации из этой сферы. Также будет необходимо «дообучение» бота для лучшей релевантности ответов пользователю.

В качестве недостатка продукта можно указать неспособность отвечать на простые вопросы не касающиеся темы. Это возможно исправить загрузив полную базу знаний, которая доступна на платформе pandorabots, но в таком случае ответы на тематические вопросы станут менее информативными, так как у общей базы шире покрытие и пользователи будут получать «осмысленный», но не несущий полезной информации ответ. В целом созданный программный продукт соответствует поставленной задаче.