

Алгоритмическое обеспечение персонального виртуального ассистента для автоматизации обработки клиентских запросов

К.С. Добратулин
НИТУ «МИСиС»
Москва, Россия
dobratulin@yahoo.com

М.И. Нежурина
НИТУ «МИСиС»
Москва, Россия
min@misis.ru

Аннотация—В данной статье описывается процесс создания алгоритмического обеспечения для функционирования персонального виртуального ассистента, позволяющего автоматизировать обработку запросов клиентов. Целью исследования является сокращение ошибок и времени обработки запроса клиента в бизнес-системах – текстовых чатах или голосовых каналах с использованием системы транскрибации текста. Приводятся результаты разработки алгоритмического обеспечения и оценка качества работы на синтетических данных.

Ключевые слова— виртуальный ассистент, обработка естественного языка, автоматизация обработки запросов, алгоритмическое обеспечение.

1. ВВЕДЕНИЕ

Персональные виртуальные ассистенты прошли путь развития длиной в несколько десятилетий – от систем распознавания и транскрибации текста до многофункциональных программных агентов с применением технологических достижений из области искусственного интеллекта [1]. В 2022 году персональные виртуальные ассистенты выполняют задачи для конечного пользователя и используются в бизнес-сервисах, таких как: виртуальные операторы, системы автоматического ответа, умные устройства, голосовые ассистенты, интерактивное голосовое меню [2]. Персональные виртуальные ассистенты, ориентированные на конечного потребителя, внедряются в программные комплексы бизнес-систем. Использование персональных виртуальных ассистентов в системах автоматизированной обработки клиентских запросов позволяет бизнесу строить новые сценарии взаимодействия с пользователем, улучшать показатели бизнес-метрик на основе оптимизации ресурсов и влияния на факторы технологические, временные, экономические и поведенческие [3]. Такие возможности использования послужили обоснованием необходимости алгоритмического обеспечения персонального виртуального ассистента.

2. ОПИСАНИЕ ЛИНГВИСТИЧЕСКИХ ПРИЗНАКОВ

Лингвистические признаки – основные при классификации запроса клиента. Так как специфика взаимодействия с персональным виртуальным ассистентом подразумевает отправку пользователем запроса в текстовом или голосовом канале, становится возможным на этапе обработки запроса выделить признаки из текстового представления запроса пользователя. Такими признаками являются признаки

лексические, текстовые и лексико-грамматические [4]. По таким признакам становится возможно классифицировать пользовательские данные, например, присвоив некоторую оценку сложности запроса на основе лингвистических признаков, попытаться проанализировать настроение пользователя на основе употребляемых в контексте запроса слов или фраз, удалить повторы или незначимые слова по текстовым признакам для статистического анализа. При разработке персонального виртуального ассистента следует внимательно учитывать потенциал анализа лингвистических признаков, так как это позволяет применять подходы из области искусственного интеллекта [5].

3. АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

Для прогнозирования цели клиентского запроса на основе текста пользователя, в данной работе используются модели машинного обучения – логистическая регрессия и библиотека fastText, разработанная Facebook [6, 7]. Выбранные модели хорошо показали себя при выполнении задач, связанных с классификацией текста [8]. Также выбор в пользу данных моделей был сделан по причине их быстрой работы и сравнительно низких требований к вычислительным ресурсам, что является важными факторами при промышленной эксплуатации [9].

В качестве языка программирования для интеграции компонентов загрузки данных и алгоритмического обеспечения был выбран язык Python [10]. Строгих ограничений по техническому обеспечению для применения выбранных технологий в данной статье не заявлено, они зависят от количественной нагрузки пользовательских запросов и должны подбираться индивидуально, исходя из задач. В качестве результата сравнения моделей, помимо качества классификации пользовательских запросов алгоритмическим обеспечением, получены время выполнения классификации одного запроса и размер модели, занимаемый на жестком диске

4. РЕЗУЛЬТАТЫ

По результатам работы прототипа получены метрики на подготовленных наборах данных. Выборкой из синтетических данных, зачастую применяемых для оценки качества работы алгоритмов машинного обучения, стал набор данных под названием «The 20 Newsgroups data set» [11]. Представленный набор данных содержит разделенные на 20 классов пользовательские тексты из обезличенной переписки между людьми, что

позволяет провести аналогию с взаимодействием пользователя с персональным виртуальным ассистентом. Выбранный набор данных является независимым от данных бизнес-заказчика и позволяет провести первичную оценку корректности и адекватности работы алгоритмического обеспечения до создания набора промышленных данных и получения их от бизнес-заказчика. Предварительная оценка качества для анализа метрик производилась с применением алгоритмического обеспечения, используемого в готовом для внедрения в систему автоматизированной обработки запросов прототипе персонального виртуального ассистента.

В качестве метрик для анализа качества работы прототипа использовались метрики точность, полнота и F1-мера. Эксперимент по оценке качества алгоритмического обеспечения проводился по схеме перекрестной проверки с разбиением набора данных на 5 непересекающихся частей. В таблице 1 представлены результаты классификации синтетических данных «The 20 Newsgroups data set».

ТАБЛИЦА 1. РЕЗУЛЬТАТЫ КЛАССИФИКАЦИИ ДАННЫХ

Класс	Точность	Полнота	F1-мера	Размер данных
alt.atheism	0.752	0.734	0.743	319
comp.graphics	0.680	0.781	0.727	389
comp.os.ms-windows.misc	0.740	0.731	0.736	394
comp.sys.ibm.pc.hardware	0.716	0.735	0.725	392
comp.sys.mac.hardware	0.828	0.816	0.822	385
comp.windows.x	0.845	0.729	0.783	395
misc.forsale	0.764	0.897	0.825	390
rec.autos	0.905	0.891	0.898	396
rec.motorcycles	0.940	0.945	0.942	398
rec.sport.baseball	0.881	0.932	0.906	397
rec.sport.hockey	0.947	0.945	0.946	399
sci.crypt	0.931	0.889	0.910	396
sci.electronics	0.764	0.784	0.774	393
sci.med	0.895	0.838	0.866	396
sci.space	0.898	0.919	0.908	394
soc.religion.christian	0.788	0.925	0.851	398
talk.politics.guns	0.711	0.901	0.795	364
talk.politics.mideast	0.964	0.854	0.906	376
talk.politics.misc	0.780	0.594	0.674	310
talk.religion.misc	0.814	0.470	0.596	251
Доля правильных ответов	0.825	0.825	0.825	–
Среднее	0.827	0.815	0.817	7532
Среднее взвешенное	0.829	0.825	0.823	7532

В результате анализа метрик можно сделать окончательный вывод о значительной точности работы алгоритмического обеспечения и его возможности обобщать применение на различных наборах данных.

5. ЗАКЛЮЧЕНИЕ

В результате выполнения работы описан процесс создания алгоритмического обеспечения для функционирования персонального виртуального ассистента. Получены результаты оценки качества классификации пользовательских запросов на синтетических данных, а также получено время работы моделей машинного обучения. Усредненное взвешенное значение F1-метрики при работе двух моделей машинного обучения составило 82.3% при времени работы моделей логистической регрессии в 189 мс и модели fastText в 214 мс соответственно. Результаты свидетельствуют о применимости выбранных моделей в промышленной среде.

ЛИТЕРАТУРА

- [1] Maher, T. History of Speech & Voice Recognition and Transcription Software / T. Maher // Dragon Systems. – Centaur Academic Media DBA Centaur Multimedia, 2003.
- [2] Mishra, A.N. Business value appropriation roadmap for artificial intelligence / A.N. Mishra, A.K. Pani // VINE Journal of Information and Knowledge Management Systems. – 2020. – P. 353-368. DOI: 10.1108/VJKMS-07-2019-0107.
- [3] Zumalt, J.R. Voice recognition technology: Has it come of age? / J.R. Zumalt // Information technology and libraries. – 2005. – Vol. 24(4). – P. 180-185. DOI: 10.6017/ital.v24i4.3382.
- [4] Mahajan, R.S. Machine learning based paraphrase identification system using lexical syntactic features / R.S. Mahajan, M.A. Zaveri // IEEE International Conference on Computational Intelligence and Computing Research (ICIC). – 2016. – P. 1-5. DOI: 10.1109/ICIC.2016.7919721.
- [5] Khan, H.U. Using Machine Learning Techniques for Subjectivity Analysis based on Lexical and Nonlexical Features / H.U. Khan, A. Daud // International Arab Journal of Information Technology (IAJIT). – 2017. – Vol. 14(4).
- [6] LaValley, M.P. Logistic regression / M.P. LaValley // Circulation. – 2008. – Vol. 117(18). – P. 2395-2399. DOI: 10.1161/CIRCULATIONAHA.106.682658
- [7] fastText [Electronic resource]. — Access mode: <https://fasttext.cc/> (20.02.2022).
- [8] Pratiwi, N.I. Hate speech detection on Indonesian Instagram comments using FastText approach / N.I. Pratiwi, I. Budi, I. Alfina // International Conference on Advanced Computer Science and Information Systems (ICACSIS). – 2018. – P. 447-450. DOI: 10.1109/ICACSIS.2018.8618182.
- [9] Tiun, S. Classification of functional and non-functional requirement in software requirement using Word2vec and fast Text / S. Tiun, U.A. Mokhtar, S.H. Bakar, S. Saad // Journal of Physics: Conference Series. – 2020. – Vol. 1529(4). DOI: 10.1088/1742-6596/1529/4/042077.
- [10] Welcome to Python.org [Electronic resource]. — Access mode: <https://www.python.org/> (20.02.2022).
- [11] Home Page for 20 Newsgroups Data Set [Electronic resource]. — Access mode: <http://qwone.com/~jason/20Newsgroups/> (20.02.2022).