

УДК 626:0627

https://doi.org/10.33619/2414-2948/101/48

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ОБРАБОТКА ЕСТЕСТВЕННОГО ЯЗЫКА КАК ОСНОВА ЧАТ-БОТОВ

©*Лиманова Н. И.*, SPIN-код: 9799-8380, д-р техн. наук, Поволжский государственный университет телекоммуникаций и информатики, г. Самара, Россия

©*Ковтун Д. С.*, Поволжский государственный университет телекоммуникаций и информатики, г. Самара, Россия, dkpguti@gmail.com

ARTIFICIAL INTELLIGENCE AND NATURAL LANGUAGE PROCESSING AS THE BASIS OF CHAT BOTS

©*Limanova N.*, SPIN-code: 9799-8380, Ph.D., Povolzskiy State University of Telecommunications and Informatics, Samara, Russia

©*Kovtun D.*, Povolzskiy State University of Telecommunications and Informatics, Samara, Russia, dkpguti@gmail.com

Аннотация. Рассматривается роль искусственного интеллекта и обработки естественного языка в создании чат-ботов. Также, в статье обсуждаются основные технологии, лежащие в основе этих систем, и анализируются их применение в практике. В статье представлен обзор существующих подходов к созданию чат-ботов. На основе полученных данных делаются выводы о практической значимости и перспективах развития искусственного интеллекта и обработки естественного языка в чат-ботах.

Abstract. The article examines the role of artificial intelligence and natural language processing in the creation of chatbots. The article also discusses the main technologies underlying these systems and analyzes their application in practice. The article provides an overview of existing approaches to creating chatbots. Based on the data obtained, conclusions are drawn about the practical significance and prospects for the development of artificial intelligence and natural language processing in chatbots.

Ключевые слова: искусственный интеллект, чат-боты, машинное обучение, нейронные сети, алгоритмы обработки текста.

Keywords: artificial intelligence, chatbots, machine learning, neural networks, text processing algorithms.

Сегодня искусственный интеллект и обработка естественного языка являются главными технологиями, которые лежат в основе современных чат-ботов. Эти системы способны понимать и обрабатывать естественный человеческий язык. Такой подход позволяет чат-ботам общаться с людьми на естественном для них языке. Благодаря этому чат-боты умеют выполнять различные задачи, от простого ответа на вопросы до более сложных задач, таких как бронирование билетов или оказание помощи при покупке на сайте. Однако, даже с учетом прогресса в этом направлении, создание по-настоящему умных чат-ботов все еще представляет собой серьезную проблему. Одна из главных причин этого – специфика человеческого общения. Дело в том, что наш язык сложен и обладает множеством нюансов и особенностей, трудных для понимания и обработки роботом.

Помимо этого, чат-боты также должны уметь быстро и на ходу адаптироваться к уникальным контекстам и ситуациям, что в свою очередь требует использования сложных алгоритмов и технологий машинного обучения. Сюда также добавляется проблема, которая заключается в том, что некоторые люди говорят невнятно и пишут с ошибками, как логическими, так и пунктуационными. Вот и получается, что создание качественного чат-бота — это непростая задача. В работе рассматриваются технологии, которые лежат в основе искусственного интеллекта (ИИ). Именно они и позволяют обрабатывать естественный язык чат-ботами в процессе их общения с людьми. Чат-боты представляют из себя программы, которые используют искусственный интеллект для понимания и обработки естественного языка, чтобы общаться с людьми на естественном языке. Как правило, в основе чат-ботов находятся технологии ИИ, например машинное обучение, глубокое обучение, нейронные сети и алгоритмы обработки текста [1].

Нейронные сети (нейросети) представляют из себя обученные алгоритмы, которые умеют моделировать и имитировать работу человеческого мозга. Они используются для обработки сложных и больших данных. Нейронные сети используются в чат-ботах для обучения моделей распознавания естественного языка, и других задач, связанных с ними. Машинное обучение является разделом ИИ, который создает алгоритмы, помогающие работать с информацией. Такие алгоритмы могут учиться на основе данных и улучшать свою эффективность по мере получения новой информации. Машинное обучение используется в чат-ботах для более точного понимания естественного языка и генерации осмысленных ответов на вопросы пользователей. И последнее важное понятие — глубокое обучение. Оно представляет из себя вид машинного обучения, в котором активно применяются нейросети. Глубокое обучение, как правило, является более точной и «естественной» программой для имитации живого общения, но требует большего времени на обучение. Для распознавания смысла и намерений человека, используются алгоритмы обработки текста. Они призваны обрабатывать данные и извлекать смысловую нагрузку из них. Такие алгоритмы обработки текста используются в чат-ботах для понимания естественного языка, а также для таких задач, как токенизация, лемматизация, частично-речевая разметка и других [2].

Таким образом, основные технологии искусственного интеллекта, которые чаще всего используются в чат-ботах, следующие: машинное обучение; глубокое обучение; нейронные сети; алгоритмы обработки текста. Чаще всего эти технологии используются в комбинации друг с другом для создания эффективных чат-ботов, способных понимать и обрабатывать естественный язык и общаться с людьми так, чтобы они порой даже не поняли, что разговаривают с роботом. В работе рассматривается роль и значение обработки естественного языка в чат-ботах. *Обработка естественного языка* (ОЕЯ) играет ключевую роль в работе чат-ботов, поскольку именно она позволяет им понимать вопросы пользователей и генерировать логичные и целевые ответы. Обработка естественного языка является разделом ИИ, который занимается созданием алгоритмов и моделей для понимания и естественной речи людей. Как правило, данные алгоритмы призваны решать те проблемы, которые мы озвучивали в начале статьи, а именно: особенность языка, индивидуальность запросов, речевые и логические ошибки пользователей [3].

ОЕЯ включает в себя такие задачи, как токенизация, лемматизация, частично-речевая разметка, синтаксический анализ, семантический анализ и другие. Остановимся на всем этом подробнее.

Токенизация. Представляет собой разделение текста на отдельные «токены», или слова. Это помогает распознавать смысл запроса.

Лемматизация. Это процесс приведения слов к их базовой форме, или лемме. Например: слово «читала» приводится к базовой форме «читать».

Частично-речевая разметка. Представляет собой процесс присвоения токенам грамматических меток, таких как существительное, глагол, прилагательное и другие. Это также упрощает понимание контекста и смысла запроса, поскольку ИИ разделяет текст так, как это удобно для его алгоритмов распознавания речи.

Синтаксический анализ. Это процесс определения структуры предложения и отношений между его элементами (например, словами и предлогами).

Семантический анализ. Он позволяет определить смысловое значение запроса и извлечь из него смысл. Происходит на основе того, что в ИИ заложена база слов и их возможных смыслов с учетом контекста [4].

В чат-ботах обработка текста используется для обучения моделей понимания естественного человеческого языка, которые способны понимать вопросы пользователей и генерировать осмысленные ответы. ОЕЯ также используется для таких задач, как распознавание намерений пользователя, то есть, того, чего он хочет в своем запросе. Например, купить что-то, или забронировать себе место в отеле. При этом очень важно, чтобы ИИ распознавал все намерения пользователей правильно.

Существует несколько подходов к реализации ОЕЯ в чат-ботах. Обычно используются такие: правило-основанный подход, статистический подход и гибридный подход. Каждый из них основан на своих принципах:

1. Правило-основанный подход основан на использовании набора правил, описывающих грамматику и семантику языка. То есть на том, как в нашем языке принято строить предложения и выражать свои мысли.

2. Статистический подход основан на использовании статистических данных, которые используются для обучения моделей. Как правило, такие модели не генерируют ответ, а просто достают готовый вариант, немного изменяя его под конкретный контекст.

3. Гибридный подход сочетает в себе элементы правило-основанного и статистического подходов. Как правило, при таком подходе информация обрабатывается дольше, зато ответ более качественный [5].

Таким образом, становится понятно, что подходов к распознаванию естественного языка становится все больше. Кроме того, модели совершенствуются и становятся все лучше с каждым днем. Из этого можно сделать вывод, что искусственный интеллект и распознавание естественного языка — это не только технологии нашего времени, а и перспективные направления в будущем.

Список литературы:

1. Осадчук П. О. Чат-боты для автоматизации внутренних коммуникаций // Молодой ученый. 2018. №27. С. 12-16.
2. Тугушева Н. А. Использование чат-ботов в различных сферах повседневной жизни // Молодой ученый. 2017. №21. С. 36-39.
3. Кодиров Э. С. У., Халилов З. Ш. Возможности и преимущества искусственного интеллекта (ИИ) и логических вычислений // Universum: технические науки. 2020. №6-1 (75). С. 18-21.
4. Голованова О. С., Масюк Н. Н. Чат-бот-digital помощник рекрутера // Территория новых возможностей. Вестник Владивостокского государственного университета экономики и сервиса. 2020. Т. 12. №4. С. 118-128.

5. Пантелеева Т. А., Арустамов Э. А., Максаев А. А. Возможности искусственного интеллекта в управлении кадровыми ресурсами в условиях свободного предпринимательства // Отходы и ресурсы. 2019. Т. 6. №3. С. 6-6.

References:

1. Osadchuk, P. O. (2018). Chat-boty dlya avtomatizatsii vnutrennikh kommunikatsii. *Molodoi uchenyi*, (27), 12-16. (in Russian).
2. Tugusheva, N. A. (2017). Ispol'zovanie chat-botov v razlichnykh sferakh povsednevnoi zhizni. *Molodoi uchenyi*, (21), 36-39. (in Russian).
3. Kodirov, E. S. U., & Khalilov, Z. Sh. (2020). Vozmozhnosti i preimushchestva iskusstvennogo intellekta (II) i logicheskikh vychislenii. *Universum: tekhnicheskie nauki*, (6-1 (75)), 18-21. (in Russian).
4. Golovanova, O. S., & Masyuk, N. N. (2020). Chat-bot-digital pomoshchnik rekrutera. Territoriya novykh vozmozhnostei. *Vestnik Vladivostokskogo gosudarstvennogo universiteta ekonomiki i servisa*, 12(4), 118-128. (in Russian).
5. Panteleeva, T. A., Arustamov, E. A., & Maksaev, A. A. (2019). Vozmozhnosti iskusstvennogo intellekta v upravlenii kadrovymi resursami v usloviyakh svobodnogo predprinimatel'stva. *Otkhody i resursy*, 6(3), 6-6. (in Russian).

*Работа поступила
в редакцию 10.03.2024 г.*

*Принята к публикации
17.03.2024 г.*

Ссылка для цитирования:

Лиманова Н. И., Ковтун Д. С. Искусственный интеллект и обработка естественного языка как основа чат-ботов // Бюллетень науки и практики. 2024. Т. 10. №4. С. 426-429. <https://doi.org/10.33619/2414-2948/101/48>

Cite as (APA):

Limanova, N., & Kovtun, D. (2024). Artificial Intelligence and Natural Language Processing as the Basis of Chat Bots. *Bulletin of Science and Practice*, 10(4), 426-429. (in Russian). <https://doi.org/10.33619/2414-2948/101/48>