

Кукушкин О.В., Полянцева К.А.

Кукушкин О.В. — магистрант, Москва, Россия, РТУ МИРЭА.
Полянцева К.А. — к.т.н., доцент, Москва, Россия, РТУ МИРЭА.

РАЗРАБОТКА И ОЦЕНКА КАЧЕСТВА МОДЕЛИ АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ БИЗНЕС-ЗАДАЧ НА ОСНОВЕ ГЕНЕРАТИВНЫХ НЕЙРОННЫХ СЕТЕЙ

В статье рассматривается разработка и оценка качества модели автоматической классификации бизнес-задач на основе генеративных нейронных сетей. Актуальность исследования обусловлена ростом объёмов неструктурированных текстовых данных в организациях и необходимостью автоматизации их обработки. Описана архитектура модели, включающая предварительную обработку текста, формирование промпта, использование few-shot примеров и интерпретацию результата. Представлены метрики качества классификации и сравнительный анализ с альтернативными подходами.

Ключевые слова: генеративные нейронные сети, классификация бизнес-задач, NLP, машинное обучение, трансформеры, LLM, few-shot learning, автоматизация бизнес-процессов.

The article considers the development and quality assessment of an automatic business task classification model based on generative neural networks. The relevance of the study is determined by the growth of unstructured textual data in organizations and the need to automate its processing. The paper describes the model architecture, including text preprocessing, prompt construction, few-shot examples and result interpretation. Classification quality metrics and a comparative analysis with alternative approaches are presented.

Keywords: generative neural networks, business task classification, NLP, machine learning, transformers, LLM, few-shot learning, business process automation.

Введение

Развитие цифровых технологий приводит к постоянному увеличению объёмов текстовой информации, используемой в деятельности современных организаций. В рамках бизнес-процессов ежедневно формируются обращения клиентов, внутренние поручения, заявки на выполнение работ, технические инциденты, финансовые запросы и аналитические задачи. Значительная часть таких данных представлена в неструктурированном или слабо структурированном виде, что усложняет их автоматическую обработку и требует применения интеллектуальных методов анализа.

Одной из важных задач обработки текстовых данных является классификация бизнес-задач. Под бизнес-задачей в рамках данного исследования понимается текстовое описание действия, запроса, проблемы или поручения, требующее дальнейшей обработки, маршрутизации, назначения исполнителя или определения приоритета. В традиционных условиях такая классификация часто выполняется вручную либо с использованием простых правил на основе ключевых слов.

Ручная обработка бизнес-задач приводит к росту временных затрат и увеличивает зависимость результата от человеческого фактора. При большом потоке входящих запросов сотрудники могут допускать ошибки в определении категории задачи, что приводит к неправильной маршрутизации, задержкам и снижению качества обслуживания. Особенно выраженной данная проблема становится в организациях, использующих CRM, Service Desk, ERP-системы и корпоративные порталы.

Современные методы обработки естественного языка позволяют перейти от простого сопоставления ключевых слов к семантическому анализу текста. Особый интерес представляют генеративные нейронные сети и трансформерные архитектуры, способные учитывать контекст входной последовательности, выявлять скрытые смысловые связи и адаптироваться к новым формулировкам задач. Это делает их перспективным инструментом для автоматической классификации бизнес-задач.

Цель исследования

Целью исследования является разработка и оценка качества модели автоматической классификации бизнес-задач на основе генеративных нейронных сетей, предназначенной для обработки неструктурированных текстовых описаний и их отнесения к заранее определённым категориям.

Для достижения поставленной цели решаются следующие задачи: анализируются особенности предметной области, рассматриваются существующие методы классификации текстовых данных, обосновывается выбор генеративной модели, разрабатывается архитектура прикладной модели классификации, определяется методика оценки качества, проводится тестирование и выполняется сравнение с альтернативными подходами.

В отличие от задач, связанных исключительно с разработкой программного интерфейса или информационной системы, акцент исследования сделан именно на модели классификации. Программная реализация рассматривается как средство проверки работоспособности предложенного подхода, а основным результатом является модель, способная преобразовывать текстовое описание бизнес-задачи в формализованный класс.

Материал и методы исследования

В качестве исходного материала исследования рассматриваются текстовые описания бизнес-задач, которые могут поступать из различных корпоративных источников: CRM-систем, Service Desk, электронной почты, систем управления проектами, внутренних форм заявок и корпоративных порталов. Входные данные представлены в виде коротких или развёрнутых текстовых сообщений, содержащих описание проблемы, запроса или поручения.

Для проведения тестирования использовалась выборка из 1200 бизнес-задач, распределённых по нескольким классам: техническая задача, клиентское обращение, финансовый запрос, кадровый запрос, аналитическая задача и организационная задача. Каждая задача имела эталонную категорию, которая использовалась для последующего сравнения с результатом классификации модели.

Разработанная модель включает несколько этапов обработки. На первом этапе выполняется предварительная подготовка текста: очистка от служебных символов, нормализация регистра, удаление лишних пробелов, устранение технического шума и подготовка текстовой последовательности к анализу. Данный этап необходим для снижения влияния случайных особенностей входного текста на результат классификации.

На втором этапе формируется структурированный запрос к генеративной модели. В состав запроса включаются системная инструкция, описание бизнес-задачи, перечень допустимых классов и few-shot примеры корректной классификации. Few-shot подход позволяет адаптировать модель к предметной области без полного переобучения и существенно снижает требования к объёму размеченных данных.

На третьем этапе генеративная модель выполняет семантический анализ входного текста и формирует ответ. В рамках предложенного подхода ответ модели ограничивается требованием выбрать одну категорию из заранее заданного списка. Это позволяет использовать генеративную нейронную сеть не как инструмент свободной генерации текста, а как вычислительное ядро прикладной модели классификации.

На четвёртом этапе осуществляется интерпретация результата. Ответ модели приводится к формализованному виду и сопоставляется с одной из категорий бизнес-задач. При необходимости может использоваться дополнительный слой проверки, исключающий некорректные или неопределённые ответы модели.

Формально задача классификации может быть представлена как задача выбора наиболее вероятного класса для входного текста. Пусть имеется множество классов $C = \{C_1, C_2, \dots, C_k\}$. Тогда итоговый класс определяется выражением $C^* = \arg \max P(C_i | T)$, где T — текстовое описание бизнес-задачи, C_i — один из допустимых классов, $P(C_i | T)$ — вероятность принадлежности текста выбранному классу.

Для оценки качества модели использовались стандартные метрики классификации: Accuracy, Precision, Recall и F1-score. Данные показатели позволяют оценить не только общую точность модели, но и её способность корректно распознавать отдельные классы при наличии неоднозначных формулировок и возможной несбалансированности данных.

Таблица 1 — Основные классы бизнес-задач, используемые при классификации

Класс	Пример задачи	Назначение результата
Техническая задача	Необходимо восстановить доступ сотрудника к CRM.	Передача в техническую поддержку
Клиентское обращение	Клиент не может оплатить заказ на сайте.	Передача в клиентский сервис
Финансовый запрос	Требуется согласовать бюджет подразделения.	Передача в финансовый отдел
Кадровый запрос	Необходимо добавить нового сотрудника в систему.	Передача в HR-подразделение
Аналитическая задача	Нужно подготовить отчёт по динамике продаж.	Передача аналитикам
Организационная задача	Требуется согласовать встречу рабочей группы.	Передача ответственному координатору

Результаты исследования и их обсуждение

В ходе исследования была проведена оценка качества разработанной модели на тестовой выборке из 1200 бизнес-задач. Для каждой задачи модель формировала предсказанный класс, который затем сравнивался с эталонной разметкой. На основании полученных результатов были рассчитаны показатели Accuracy, Precision, Recall и F1-score.

Итоговые значения метрик показали, что модель обеспечивает достаточно высокий уровень качества классификации. Значение Accuracy составило 0,91, Precision — 0,89, Recall — 0,87, F1-score — 0,88. Это свидетельствует о том, что модель не только корректно классифицирует большую часть входных задач, но и сохраняет приемлемый баланс между точностью и полнотой.

Для оценки преимуществ предложенного подхода результаты были сопоставлены с альтернативными методами классификации. В качестве базовых подходов рассматривались TF-IDF + Logistic Regression, TF-IDF + SVM и BERT fine-tuning. Сравнение показало, что классические методы уступают

генеративной модели, прежде всего из-за слабого учёта контекста и ограниченной способности работать с вариативными формулировками.

Модель BERT fine-tuning показала близкие значения качества, однако её применение требует предварительного дообучения на размеченной выборке и может быть менее гибким при изменении набора классов. Генеративная модель, использующая few-shot подход, позволяет быстрее адаптироваться к новым типам задач за счёт корректировки промпта и набора примеров.

Таблица 2 — Сравнение качества различных подходов к классификации

Модель	Accuracy	Precision	Recall	F1-score
TF-IDF + Logistic Regression	0,78	0,75	0,72	0,73
TF-IDF + SVM	0,81	0,79	0,76	0,77
BERT fine-tuning	0,89	0,87	0,86	0,86
Генеративная модель few-shot	0,91	0,89	0,87	0,88

Анализ результатов показывает, что наибольший прирост качества достигается по сравнению с классическими методами на основе TF-IDF. Это объясняется тем, что такие методы учитывают преимущественно частотные характеристики текста и плохо отражают смысловые связи между словами. В реальных бизнес-задачах одна и та же проблема может быть описана различными словами, поэтому поверхностного анализа лексики недостаточно.

Генеративная модель демонстрирует более высокую устойчивость к вариативности входных данных. Она способна учитывать контекст запроса и интерпретировать смысл задачи даже в тех случаях, когда в тексте отсутствуют явно выраженные ключевые слова. Это особенно важно для задач, поступающих из реальной корпоративной среды, где формулировки часто бывают неполными, неоднозначными или содержат профессиональные сокращения.

Дополнительный анализ ошибок показал, что наибольшее число неправильных классификаций наблюдается между аналитическими и организационными задачами. Это связано с частичным пересечением их смыслового содержания. Например, задача по подготовке отчёта может быть интерпретирована как аналитическая, если акцент сделан на данных, либо как

организационная, если основной смысл связан с координацией деятельности сотрудников.

Полученные результаты подтверждают практическую применимость модели. Она может использоваться как компонент корпоративной информационной системы, выполняющий первичную маршрутизацию задач, определение категории запроса, назначение ответственного подразделения и последующую аналитику бизнес-процессов.

Вместе с тем модель имеет ряд ограничений. Качество классификации зависит от полноты и корректности входного текста. При слишком коротких или противоречивых формулировках вероятность ошибки возрастает. Кроме того, использование генеративных моделей может быть связано с вычислительными затратами и необходимостью контроля ответов в критически важных бизнес-сценариях.

Для минимизации указанных ограничений целесообразно использовать предварительную нормализацию текста, расширение набора few-shot примеров, установку порога уверенности и передачу спорных случаев оператору. Также важным направлением развития является регулярное обновление промптов и накопление новых размеченных примеров для повышения качества адаптации модели.

Выводы

В результате исследования разработана прикладная модель автоматической классификации бизнес-задач на основе генеративных нейронных сетей. Модель ориентирована на обработку неструктурированных текстовых описаний, поступающих из корпоративных информационных систем, и их отнесение к заранее определённым категориям.

Предложенный подход включает этапы предварительной обработки текста, формирования структурированного промпта, использования few-shot примеров, генерации ответа и интерпретации результата. Такой механизм позволяет применять генеративную нейронную сеть как вычислительное ядро

прикладной модели классификации, а не как инструмент свободной генерации текста.

Проведённая оценка качества показала, что модель достигает Accuracy 0,91, Precision 0,89, Recall 0,87 и F1-score 0,88. Сравнение с альтернативными подходами подтвердило преимущество генеративной модели перед классическими методами TF-IDF + Logistic Regression и TF-IDF + SVM, а также её конкурентоспособность по сравнению с BERT fine-tuning.

Практическая значимость исследования заключается в возможности применения разработанной модели для автоматизации маршрутизации бизнес-задач, снижения нагрузки на сотрудников, повышения скорости обработки обращений и улучшения прозрачности корпоративных процессов. Дальнейшее развитие работы может быть связано с расширением набора классов, использованием более крупных выборок, внедрением механизма оценки уверенности и интеграцией модели в промышленные информационные системы.

Список использованных источников

1. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // Proceedings of NAACL. — 2019. — P. 4171–4186.
2. Goodfellow I., Bengio Y., Courville A. Deep Learning. — Cambridge: MIT Press, 2016. — 775 p.
3. Jurafsky D., Martin J. H. Speech and Language Processing. — 3rd ed. — Draft, 2023. — 1500 p.
4. Radford A. et al. Language Models are Few-Shot Learners // Advances in Neural Information Processing Systems. — 2020.
5. Vaswani A. et al. Attention Is All You Need // Advances in Neural Information Processing Systems. — 2017. — P. 5998–6008.
6. OpenAI. GPT-4 Technical Report [Электронный ресурс]. — URL: <https://arxiv.org/abs/2303.08774> (дата обращения: 01.03.2026).
7. Hugging Face. Transformers Documentation [Электронный ресурс]. — URL: <https://huggingface.co/docs/transformers> (дата обращения: 01.03.2026).
8. Scikit-learn Developers. Model evaluation: quantifying the quality of predictions [Электронный ресурс]. — URL: https://scikit-learn.org/stable/modules/model_evaluation.html (дата обращения: 01.03.2026).
9. PyTorch Foundation. PyTorch Documentation [Электронный ресурс]. — URL: <https://pytorch.org/docs/stable/index.html> (дата обращения: 01.03.2026).
10. McKinsey & Company. The State of AI in 2023 [Электронный ресурс]. — URL: <https://www.mckinsey.com> (дата обращения: 01.03.2026).