1. Анализ предметной области

1.1. Наименование предметной области

Разработка и внедрение программного обеспечения на основе искусственного интеллекта (ИИ) и машинного обучения (МО) для автоматизации и оптимизации бизнес-процессов в сфере клиентского сервиса и поддержки.

1.2. Компания-исполнитель:

Название компании Aether Algoritms

Вид деятельности разработка решений в сфере искусственного интелекта

Цель: Повышение эффективности бизнеса клиентов через автоматизацию рутинных и интеллектуальных задач.

Миссия: Сделать передовые технологии ИИ доступными для среднего и крупного бизнеса.

1.3. Проблематика предметной области

Современные компании, особенно в секторе ритейла, телекома и финансов, ежедневно получают тысячи обращений от клиентов через различные каналы: email, чаты, мессенджеры, формы на сайте. Обращения носят разнообразный характер:

* Запросы информации (“Какой у тебя баланс?”)
* Жалобы (“Пришёл бракованный товар”)
* Техническая поддержка (“не подключается роутер”
* Благодарности и другие)

Ключевые проблемы:

1. Ручная обработка: Операторы вручную читают каждое обращение, чтобы определить его суть (интент) и направить в нужный отдел. Это медленно и дорого.

2. Задержки в реакции: Клиенты вынуждены долго ждать ответа, что снижает удовлетворенность

3. Ошибки маршрутизации: Человеческий фактор приводит к тому, что обращение попадают не к тем специалистам, что увеличивает время решения проблемы.

4. Невозможность анализа: Большие массивы неструктурированных текстовых данных невозможно анализировать вручную для выявления тенденций и превентивного решения проблем.

1.4. Постановка задачи

Разработать программный модуль, который будет автоматически:

* Классифицировать входящие текстовые обращения по заранее заданным темам (например проблема с доставкой, вопрос по оплате, жалоба на качество).
* Определять срочность обращения на основе тональности текста (сентимент-анализа).
* Маршрутизировать обращение в соответствующую очередь или группу специалистов в CRM-системе компания-клиента.
* Интегрироваться с популярными платформами (например, Jira, Zendesk, Bitrix24) через API.

1.5. Ключевые сущности и понятия предметной области

* Обращение (Ticket/Request): Основная сущность. Имеет атрибуты:
* Текст (основание содержание)
* Источник (email, чат, etc.)
* Иетка\_класса (целевой признак для обучения модели)
* Срочность (высокая,средняя,низкая)
* Статус (новый, в работе, решен)

Модель машинного обучения (ML Model):

Ядро системы. Обучена на размеченных исторических данных для предсказания класса нового обращения.

* Алгоритмы: планируется использовать методы NLP (Natural Language Processing):
  + Классификация: BERT, DistilBERT, Logistic Regression на TF-IDF векторизации.
  + Сентимент-анализ: готовые библиотеки (nltk, spacy, transformers).
* Набор данных (Dataset):

Коллекция рассеченных исторических обращений. Главный ресурс для обучения.

* + X\_features – тексты обращений.
  + Y\_true – ручные метки (категории), проставленные опытыми операторами.
* Интеграция (API): Механизм взаимодействия модуля с внешними системами. Используются RESTful API.
* Метрики качества:
  + Для модели: Accuracy (точность), Precision (точность), Recall (полнота), F1-Score (Гармоническое среднее).
  + Для бизнеса: CSAT, FRT, Обращений, решенных с первого раза.

1.6. Технологический стек

Бэкенд: Python 3.9+, фреймворки FastAPI/Flask.

* Машинное обучение: Библиотеки scikit-learn, transformers (Hugging Face), pandas, numpy.
* Обработка текста (NLP): nltk, spacy, re (для регулярок).
* Развёртывание (Deployment): Docker, Dokcer Compose для конвертизации.
* Хранение данных: Для кэша и логирования Redis, PostgreSQL.
* Система контроля версий: Git,GitHub/GitLab.

1.7. Ограничения и риски:

* Качество данных: Низкое качество разметки исторических данных приведет к плохому качеству модели.
* Конфиденциальность: Обращения могут содержать ПДн (персональные данные) необходимо соблюдение 152-Ф3. Данные должны анонимизироваться.
* Изменяемость языка: Появление новых слов, сленга, изменения контекста (например, пандемия COVID добавила новые типы обращений) требует периодического переобучения модели.
* Интеграция: Риск нестабильности или изменения API со стороны CRM-систем заказчика.

1.8. Ожидаемые результаты и ценность

* Для бизнес-заказчика:
  + Сокращение времени обработки обращения на 40-60%
  + Повышение точности маршрутизации до 95%
  + Высвобождение времени операторов для решения сложных, нетривиальных задач.
  + Улучшение ключевых метрик CSAT и FRT.

Структура компании Aether Algoritms

Общий размер компании: 120-200 сотрудников

1. Руководство

* Задача: Стратегическое управление компанией.
* Ключевые роли: CEO, CTO, CPO, CFO, CDO.
* Примерная численность: 3-5

2. Научно-исследовательский отдел

Задача: Фундаментальные и прикладные исследования генерация идей, прототипирование.

* Ключевые роли: Research Scientist, ML Researcher, Research Engineer.
* Примерная численность: 10-15 человек.

3. Отдел разработки (Engineering Department) — КРУПНЕЙШИЙ ОТДЕЛ

· Задача: Создание, поддержка и развертывание программных продуктов.

· Ключевые роли: ML Engineer, Software Engineer (Backend/Frontend), DevOps/MLOps Engineer, QA Engineer.

· Примерная численность: 50-65 человек (~40% от общего числа).

· Machine Learning Engineering: 15-20 чел.

· Software Development (Backend/Frontend/API): 25-35 чел.

· MLOps/DevOps/QA: 8-10 чел.

4. Отдел данных (Data Department)

· Задача: Обеспечение жизненного цикла данных (сбор, очистка, разметка, анализ).

· Ключевые роли: Data Engineer, Data Analyst, Data Annotator/Labeler.

· Примерная численность: 15-20 человек.

· Data Engineering: 5-7 чел.

· Data Analysis: 5-7 чел.

· Data Annotation: 5-6 чел. (может быть частично аутсорсена).

5. Отдел продукта (Product Department)

· Задача: Определение стратегии продукта, взаимодействие между R&D, разработкой и бизнесом.

· Ключевые роли: Product Manager, Product Designer (UX/UI), UX Researcher.

· Примерная численность: 8-12 человек.

· Product Management: 4-6 чел.

· Product Design: 3-5 чел.

· UX Research: 1-2 чел.

6. Коммерческий отдел (Business & Sales)

· Задача: Привлечение клиентов, монетизация, продвижение бренда.

· Ключевые роли: Sales Manager, Business Development Manager, Marketing Specialist, Content Manager.

· Примерная численность: 12-18 человек.

· Sales & BD: 8-10 чел.

· Marketing: 4-8 чел.

7. Отдел внедрения и поддержки (Implementation & Customer Success)

· Задача: Внедрение продукта у клиентов, техническая поддержка, обучение.

· Ключевые роли: Solutions Architect, Customer Success Manager, Technical Support Engineer.

· Примерная численность: 10-15 человек.

· Численность напрямую зависит от количества и размера клиентов.

8. Юридический отдел и этика (Legal & Ethics)

· Задача: Юридическое сопровождение, обеспечение этичности и соответствия регуляториям.

· Ключевые роли: Lawyer, Compliance Officer, AI Ethics Specialist.

· Примерная численность: 3-5 человек.

· Комментарий: Небольшая, но критически важная для ИИ-компании команда.

9. Вспомогательные отделы (Support Functions)

· Задача: Обеспечение работы компании.

· Ключевые роли: HR-менеджер, Recruiter, Accountant, Office Manager, SysAdmin.

· Примерная численность: 8-12 человек.

· HR & Recruiting: 4-5 чел.

· Finance & Accounting: 2-3 чел.

· Office & IT: 2-4 чел.