Декомпозировать планируемую систему, определить ее основные функциональные части

Декомпозиция антифрод-системы позволяет выделить её основные функциональные компоненты, каждая из которых решает конкретную задачу в общем процессе обнаружения и предотвращения мошенничества.

Ниже представлена структурированная декомпозиция системы с описанием ключевых модулей.

**🧩** Декомпозиция антифрод-системы: основные функциональные части

**1.** Модуль сбора и обогащения данных (Data Ingestion & Enrichment)

Назначение: Получение и подготовка данных о транзакции в реальном времени.  
Функции:

* Приём транзакционных событий от платёжного шлюза (в реальном времени).
* Обогащение транзакции контекстными данными:
  + История поведения клиента (частота, суммы, геолокация, устройства).
  + Статус счёта/карты (лимиты, блокировки).
  + Внешние источники (чёрные списки, IP-репутация, данные о рисковых странах).
* Нормализация и валидация данных.

Технологии: Kafka / RabbitMQ, streaming-процессинг (Flink, Spark Streaming), lookup-сервисы.

**2.** Модуль генерации признаков (Feature Engineering)

Назначение: Преобразование «сырых» данных в признаки для ML-модели.  
Функции:

* Расчёт агрегатов в скользящих окнах:
  + Количество транзакций за 1/5/60 мин.
  + Средняя/максимальная сумма за час.
  + Отклонение от типичного поведения клиента (z-score).
* Категориальные эмбеддинги (например, MCC-коды, типы получателей).
* Геопространственные признаки (расстояние между последними транзакциями).
* Онлайн-кэширование агрегатов для низкой latency.

Особенность: Должен работать в реальном времени (< 20 мс на транзакцию).

**3.** Модуль машинного обучения (ML Inference Engine)

Назначение: Принятие решения о риске мошенничества.  
Функции:

* Загрузка предобученной модели (например, LightGBM, XGBoost).
* Вычисление вероятности мошенничества на основе признаков.
* Применение калиброванного порога (настроенного под recall ≥ 95%, FPR ≤ 5%).
* Поддержка нескольких моделей (A/B-тестирование, fallback-модель).
* Возврат не только решения, но и объяснения (топ-3 признака риска).

Требования: Низкая задержка (< 50 мс), высокая доступность.

**4.** Модуль принятия решений (Decision Engine)

Назначение: Преобразование ML-прогноза в бизнес-действие.  
Функции:

* Автоматическая блокировка — если риск > высокого порога.
* Отправка на ручную проверку — если риск в средней зоне.
* Пропуск без проверки — если риск низкий.
* Учёт бизнес-правил (например: «все транзакции > 100 000 ₽ — на ручную проверку»).
* Интеграция с системой уведомлений (SMS/email клиенту).

Гибкость: Правила должны настраиваться без перезапуска системы.

**5.** Модуль обратной связи и разметки (Feedback & Labeling)

Назначение: Сбор «истинных» меток для переобучения модели.  
Функции:

* Фиксация результатов ручной проверки («мошенничество» / «ложная тревога»).
* Автоматическое подтверждение мошенничества по заявкам клиентов.
* Сбор отложенных меток (мошенничество может выявляться через 1–7 дней).
* Формирование датасета для переобучения.

Важно: Поддержка временной согласованности (label leakage prevention).

**6.** Модуль обучения и переобучения моделей (Model Training & CI/ML)

Назначение: Обеспечение актуальности и качества ML-моделей.  
Функции:

* Автоматическая подготовка обучающих выборок.
* Обучение и валидация моделей с учётом дисбаланса (F2-score, PR-AUC).
* Тестирование на исторических сценариях (backtesting).
* A/B-тестирование новых моделей.
* Деплой новой версии модели в продакшен (с откатом при падении метрик).

Инструменты: MLflow, Kubeflow, Airflow, DVC.

**7.** Модуль мониторинга и аналитики (Monitoring & Observability)

Назначение: Контроль качества, производительности и рисков.  
Функции:

* Мониторинг бизнес-метрик: ущерб, FPR, recall, отток.
* Мониторинг ML-метрик: дрейф признаков, деградация PR-AUC.
* Мониторинг инфраструктуры: latency, throughput, ошибки.
* Алертинг при отклонении от порогов.
* Дашборды для аналитиков и fraud-операторов.

Инструменты: Grafana, Prometheus, Evidently AI, WhyLabs.

**8.** Модуль интерпретации и поддержки (Explainability & Support)

Назначение: Обеспечение прозрачности решений для клиентов и операторов.  
Функции:

* Генерация объяснений (например, через SHAP или LIME).
* API для службы поддержки: «Почему была заблокирована транзакция?»
* Формирование отчётов для регуляторов.

**📐 Схема взаимодействия (упрощённо)**

1

2

3

4

5

6

7

8

9

10

11

Платёжный шлюз

↓

[Сбор и обогащение данных]

↓

[Генерация признаков]

↓

[ML Inference Engine] → [Decision Engine] → Действие (блок/проверка/пропуск)

↑ ↓

[Модуль обучения] ← [Обратная связь и разметка]

↓

[Мониторинг и аналитика]

**💡 Заключение**

Антифрод-система — это не просто ML-модель, а сложный инженерный продукт, включающий:

* потоковую обработку данных,
* low-latency inference,
* гибкую логику принятия решений,
* цикл непрерывного обучения,
* прозрачность и соответствие регуляторным требованиям.

Каждый из выделенных модулей может развиваться и масштабироваться независимо, что упрощает поддержку и развитие системы.