

# Прогнозирование финансовых показателей заёмщиков с использованием методов предиктивной аналитики

РОМАН ТИНЯЕВ



Капитал





## Основные цифровые технологии (индустрия 4.0)



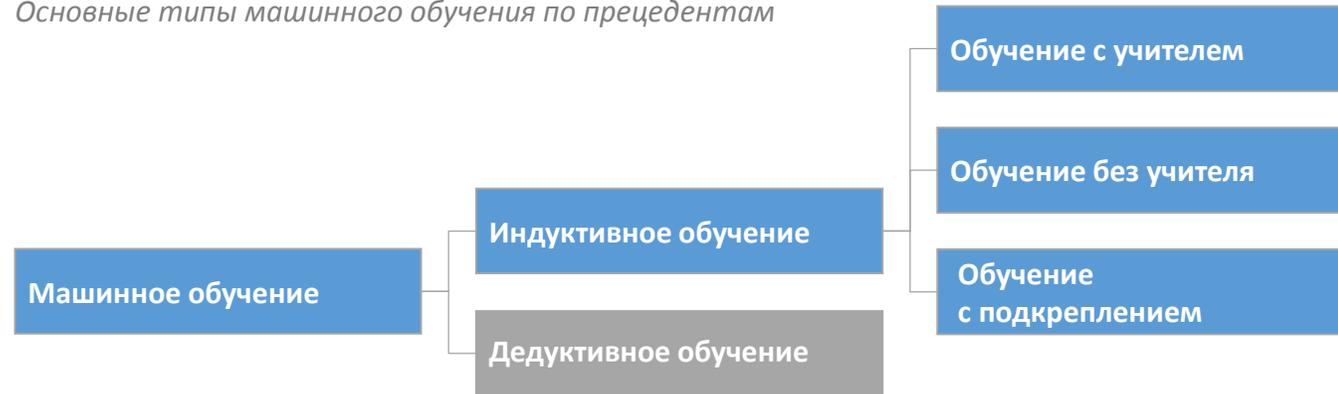
## Определение

**Машинное обучение** (Machine Learning, ML) - это обширный подраздел прикладной математики, находящийся на стыке математической статистики, оптимизации, искусственного интеллекта, изучающий методы построения алгоритмов, способных обучаться по эмпирическим данным.

## Типы машинного обучения

- **Обучение по прецедентам** (индуктивное обучение), основано на выявлении общих закономерностей по частным эмпирическим данным
- **Дедуктивное обучение** - формализация знаний экспертов и их перенос в компьютер в виде базы знаний

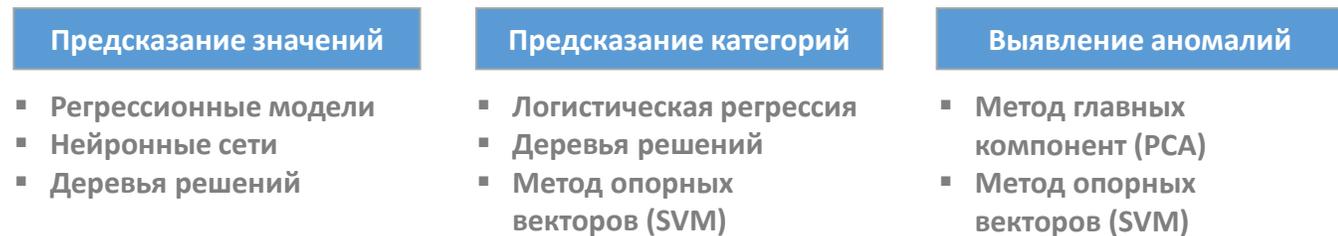
Основные типы машинного обучения по прецедентам



## Алгоритмы машинного обучения

**Алгоритм машинного обучения** - вычислительная процедура, которая по обучающей выборке производит настройку модели. Выходом алгоритма обучения является функция, аппроксимирующая неизвестную (восстанавливаемую) зависимость. Алгоритм обучения ищет такой набор параметров модели, при котором функционал качества на заданной обучающей выборке принимает оптимальное значение.

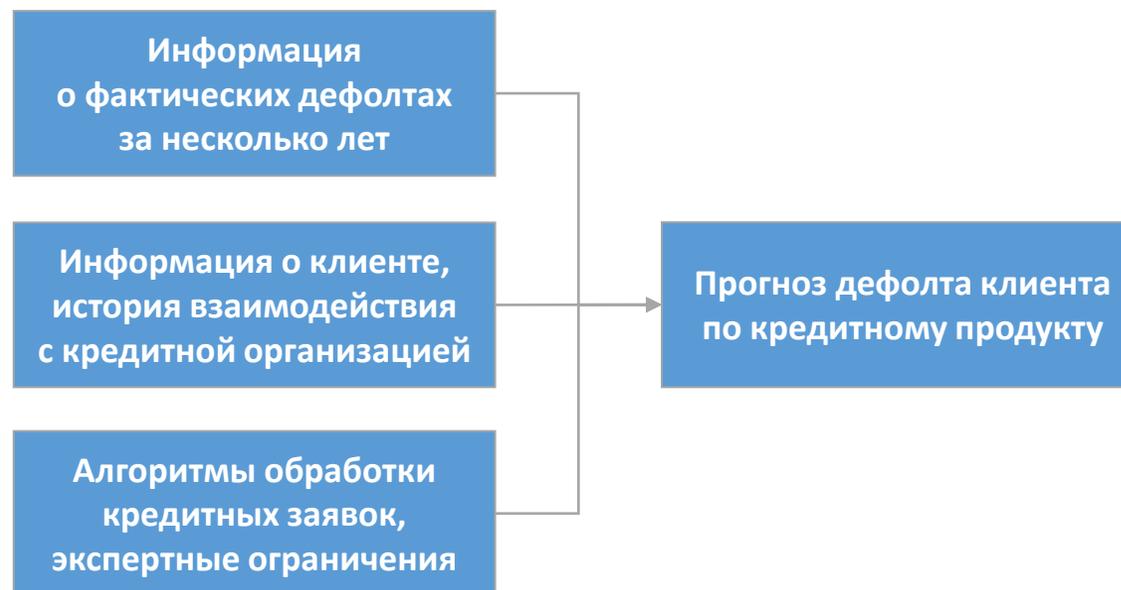
Примеры алгоритмов машинного обучения



## Приложение ML в экономике

- Управление активами (инвестиционный менеджмент)
- Биржевая торговля
- Кредитный скоринг, анализ существующих кредитных портфелей
- Маркетинговый анализ (например, прогнозирование спроса, сегментация клиентов, предсказание ухода клиентов, оптимизация ценовой политики)
- Расчет страховой премии
- Обнаружение мошенничества, защита от киберугроз
- Финансовое планирование

Пример использования машинного обучения (модель прогнозирования дефолта по кредитному продукту)



## Этапы типового проекта по созданию предиктивной модели средствами машинного обучения

### 1. Бизнес-анализ

- Определение бизнес-требований (цели проекта, рамки проекта / объём работ, ресурсы и т.д.)
- Подготовка технического задания

### 2. Подготовка данных

- Выбор перечня и источников данных (внутренних и внешних)
- Получение данных
- Обработка данных (очистка шумов, выбросов, нормализация)

### 3. Поиск и выбор факторов

- Поиск скрытых закономерностей
- Выявление значимых закономерностей
- Выбор предикторов

### 4. Обучение машины

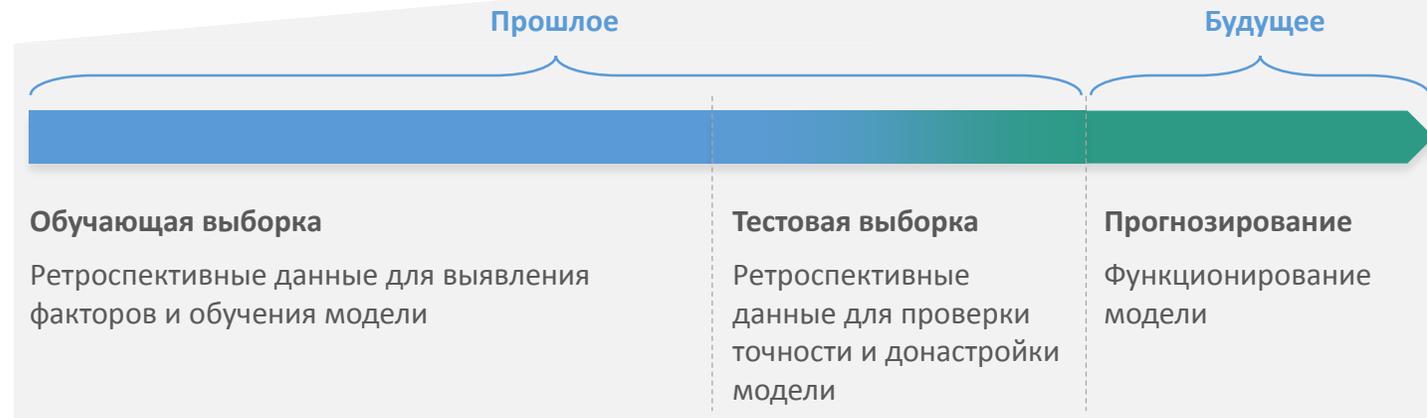
- Построение модели алгоритмами машинного обучения
- Тюнинг модели

### 5. Тестирование

- Тестирование модели на реальных данных

### 6. Эксплуатация

- Использование модели для прогноза показателей
- Регулярное переобучение модели на новых данных



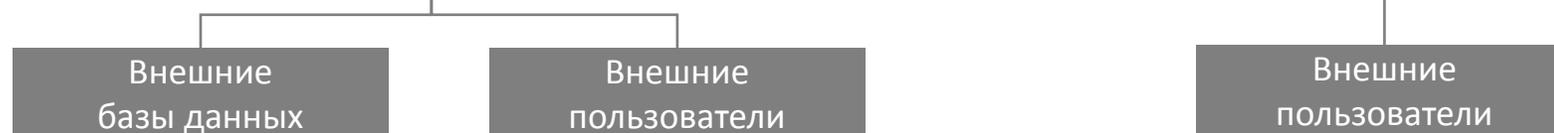
# Система с функционалом предиктивной аналитики включает в себя 3 блока (подсистемы): сбор и хранение данных, аналитика, визуализация и интерфейс

## Упрощенный пример ИТ-архитектуры системы с функционалом предиктивной аналитики и веб-интерфейсом

### Контур кредитной организации (КО)

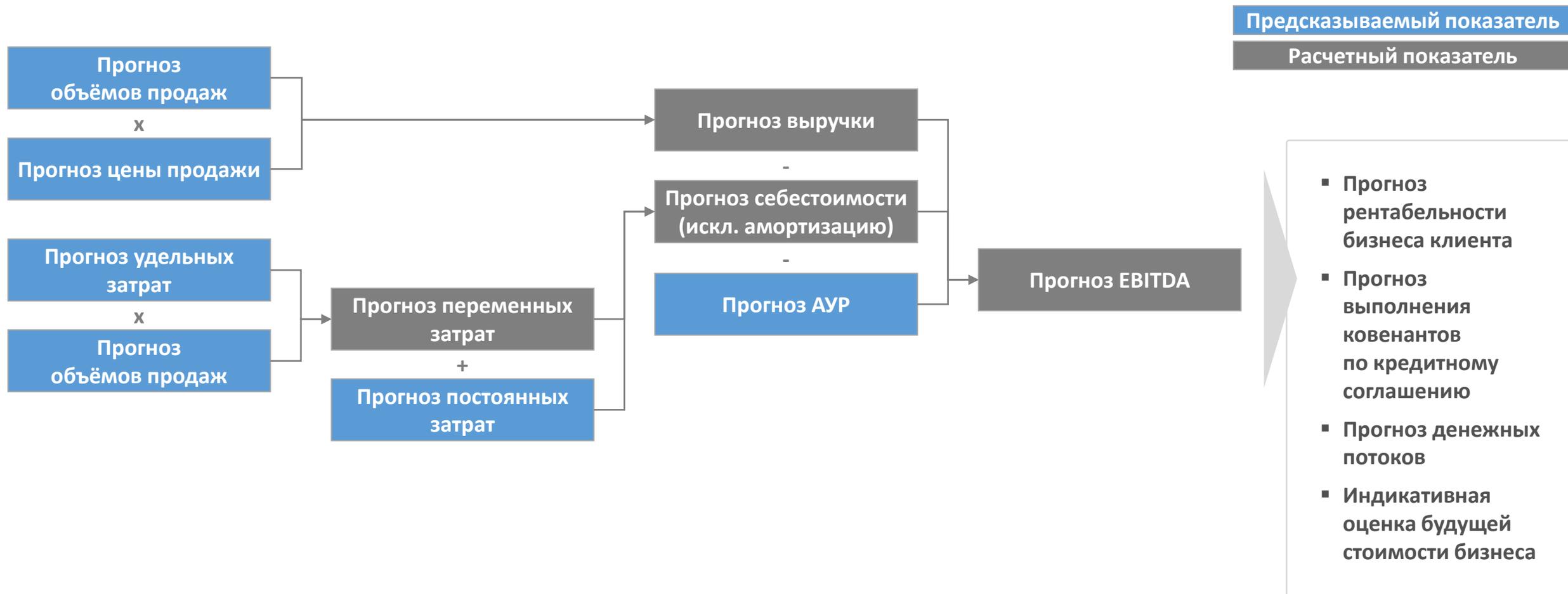


1. Выгрузка данных из внешних баз данных SQL-запросами по VPN-каналу.
2. Загрузка данных внешними пользователями на SFTP сервер КО.



### Внешний контур

## Пример расчета прогнозного значения EBITDA на основе предсказываемых финансовых и операционных показателей



# Базой для расчета финансовых показателей является прогноз основных внутренних и внешних факторов, влияющих на доходы и расходы клиента

## Пример подхода к прогнозированию выручки торгового центра (ТЦ)

Обучение системы на ретроспективных данных, построение прогноза с учетом ключевых факторов:

- сезонность (изменение показателя внутри дня / недели, праздничные дни, сезонные обновления коллекций, сезонные распродажи, сезонные различия в стоимости товаров и т.д.)
- учет планируемых маркетинговых мероприятий
- учет прогноза погоды (температура, осадки)



Обучение системы на ретроспективных данных, построение прогноза с учетом сезонных факторов



Предсказываемый показатель

Расчетный показатель

Выручка ТЦ

# С развитием технологий будет расти объем информации, которую можно будет использовать для выявления закономерностей и построения более точных предиктивных моделей



## Существующие вызовы

- Кредитные организации накапливают большие объёмы информации, которые всё сложнее обрабатывать традиционными методами
- На обработку (анализ) информации тратится много времени / ресурсов
- Часто анализируются ретроспективные данные, когда на ситуацию уже нельзя повлиять
- Часть информации не учитывается вообще
- Решения принимаются в условиях неопределённости или без учета всех значимых фактов

## Решение

**Применение предиктивной аналитики средствами машинного обучения позволяет:**

- автоматизировать работу с большими объемами информации, минимизировать «человеческий фактор»
- компенсировать недостаток опыта, отраслевой экспертизы
- более точно предсказывать отклонения по ключевым параметрам, своевременно реагировать на них
- быстрее принимать обоснованные решения

## Сопутствующие условия

- Нарращивание компетенций в области цифровых технологий (в первую очередь, интеллектуальная обработка данных)
- Создание условий для кооперации специалистов по анализу и обработке данных с бизнес-аналитиками и отраслевыми экспертами
- Обеспечение бесперебойного автоматизированного доступа к первичной информации в режиме онлайн
- Обеспечение информационной безопасности

Спасибо за внимание!