



Искусственный Интеллект: мифы, рифы, инструкция по применению

Юрий Сирота, старший вице-президент

- Руководитель департамента Искусственного Интеллекта и Анализа Данных
- Руководитель центра компетенций VI

2020-02-18, ПАО УралСиб

Цифровизация → генерирование
данных

Данные необходимо, но
недостаточно сгенерировать,
складировать в озеро/хранилище

ИИ = способ извлечения знаний из
данных →
правильные управленческие
действия → **МОНЕТИЗАЦИЯ**

1. не только **автоматизация** рутинных задач
2. улучшать подход к **принятию решения**: к примеру, ИИ может классифицировать и прогнозировать лучше, чем человек, особенно на больших объёмах информации
3. поставлять не только данные, а **вводы на основе данных**

Искусственный интеллект: необходимые условия успешности проекта

ориентирован на потребности заказчика и сосредоточен на конкретных целях

➤ просчитан экономически:

себестоимость текущего процесса <?>

приобретение + внедрение + поддержание

➤ оценены выгоды:

- сокращение операционных затрат, ошибок, цикла обслуживания, текучести персонала
- увеличение доступности сервиса

➤ не для стремления следовать хайпу!!!

Необходимы

- экспертиза в предметной области
- тестирование

Рекомендации

- учите экспертов в предметной области основам ИИ
- учите специалистов ИИ предметной области

Замечания

- ИИ - не волшебная палочка:
- человек задает правильные вопросы и требования
- разрабатывается под решение конкретных задач

Проблема 1: цифровизация бизнеса

- внедрение ПО
- установка оборудования
- изменение процессов

Проблема 2:

подбор квалифицированных специалистов в ИИ

Проблема 3:

- получение для обучения ИИ больших, всеобъемлющих, репрезентативных наборов данных
- неполные, некачественные данные
- машинное обучение может потребовать большое количество человеческих усилий для разметки данных для обучения

Проблема 4:

Модели машинного обучения, имеют репутацию неинтерпретируемых «**черных ящиков**»

Следствие:

- прогнозы моделей могут игнорироваться пользователями
- упущенные возможности

Рекомендации:

- выделение наиболее важных факторов в предсказании
- валидация моделей

Проблема 5:

Гиперпараметры моделей машинного обучения

- глубина деревьев в случайном лесе
 - количество слоев в глубоких нейронных сетях
- должны быть определены до обучения модели. Их значения не вытекают из имеющихся данных

Решение

Методы проб, перебора

Следствия

- сложность и неинтерпретируемость моделей
- переученность моделей
- ошибочная спецификация моделей

Проблема 6:

нестабильность моделей на

Out-of-sample/out-of-time

Проблема 7:

Модели ML необходимо тюнинговать, чтобы улавливать новые закономерности в данных

Способы решения

- периодическое ручное обновление модели
- динамическая калибровка: самообучающиеся алгоритмы, методы Байеса

Риски: чрезмерный акцент на краткосрочных закономерностях в данных может нанести ущерб точности модели в будущем

Решение: валидация

Проблема 8:

Встраивание рекомендации и алгоритмов ИИ (ИИ) в бизнес процесс - задача технологическая, инфраструктурная, организационная.

Эффективное и быстрое взаимодействие бизнеса и IT

Модели ML требуют больших вычислений - это требование может упускаться из виду в процессе разработки

Проблемы:

Разработчики строят сложные прогностические модели, но IT системы банка не способны их интегрировать или «эксекьютить» с приемлемой задержкой

Риски:

- большого объема данных проходящих через модель
- остаться в пилоте
- нестыковки с IT архитектурой
- задержки (проскальзывания)

- руководство организации и бизнес-подразделений разделяют стратегическое видение аналитической функции, приверженность аналитической культуре проникает во все уровни организации
- созданы кросс-функциональные команды: нет барьеров между IT, центром аналитической компетенций, бизнесом.

➤ инвестиции в инфраструктуру, технологии, компетенции:

- создан стек инструментов по сбору, профилированию, аналитике данных
- использование продвинутых математических подходов, наличие методологии моделирования
- наличие компетенции в областях: науки о данных, инженерии данных, управлении данными

Аналитический проект: секреты успеха

- правильная постановка задачи от заказчика и корректные ожидания
- приоритет аналитических инициатив «быстрая победа»
- единая команда разработки, внедрения в продакшн, поддержки: остается неизменной
- результаты аналитики интегрируются в бизнес процессы

Аналитический проект: признаки провала

- отсутствует понимание разницы между традиционной аналитикой (BI, отчетность) и продвинутой аналитикой (прогнозирование, предписание)
- опора бизнес-лидеров на интуицию и традиционную практику принятия решений, а также сопротивление изменениям
- стремление иметь технологию, а не решить задачу

- финансовый эффект аналитической инициативы не оценен
- некачественны данные: нехватка, однобокость, проблемы интеграции, неучтенная динамика, пренебрежение неструктурированными данными
- иллюзия того, что ИИ может решить любую задачу и без применения усилий, как волшебная палочка

- центр аналитической компетенции изолирован от бизнеса:
 - отсутствует бизнес-аналитик/специалист по монетизации, гармонизирующий взаимодействие заказчика и аналитиков данных.
 - математики не знают предметной области
 - невостребованность бизнесом/ заказчиками аналитических моделей

Цель: обеспечение в реальном времени

- автоматизации
- поддержку принятия решений

Механизм:

- обрабатывает текущие и исторические данные в ответ на события
- использует потоковую обработку, оптимизацию, управление бизнес-правилами
- интегрирован в бизнес-процессы

Аналитические запросы генерируются через

- поиск
- обработку естественного языка (НЛП)
- голос (Voice Recognition)

Интерфейс поиска или разговор с виртуальным
ПОМОЩНИКОМ

ИИ работает как человеческий мозг

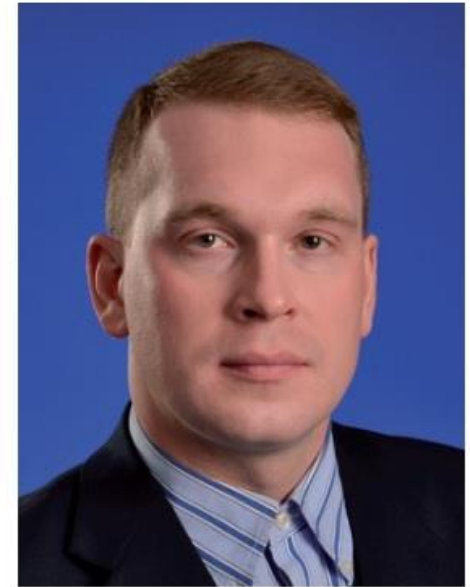
- некоторые алгоритмы ML вдохновлены структурой человеческого мозга
- некоторые алгоритмы эффективнее возможностей человека: computer vision может работать точнее, чем человеческий глаз
- ИИ решает одну задачу чрезвычайно хорошо, но эффективность падает, если чуть изменить условия

ИИ обучается самостоятельно

- человек формулирует проблему, готовит данные, определяет соответствующие наборы данных, удаляет необъективности в обучающих данных
- человек постоянно обновляет программное обеспечение для интеграции новых знаний и данных в следующий цикл обучения

Миф 3: ИИ заменит человека

- ИИ - это ПО для решения частной задачи
- ИИ может выполнять рутинную работу



Yury Sirota

PhD in Mathematics, MSc in Data Science

MSc in Software Engineering, MSc in Finance

Artificial Intelligence and Data Analytics Office
Business Intelligence Competency Center
Chief Data and Analytics Officer, Senior vice president

Tel: +7 (495) 785-12-12 ext. (054) 8448 8, Efremova st.,
Mob.: +7 (926) 720-0-790    Moscow, 119048, Russia
E-mail: SirotaYN@uralsib.ru www.uralsib.ru