

# Мы учим Облака – Облака учат нас (Машинное Обучение как облачная услуга)

Олег Фатеев  
эксперт по облачной разработке

Cloud & Digital Transformation  
Москва, Digital October  
23 марта 2017

# Где применяется Машинное Обучение

- ❑ Распознавание речи, жестов, образов, рукописного ввода
- ❑ Компьютерное зрение, машинное осязание, движение роботов
- ❑ Распознавание эмоций и эмоциональный интерфейс
- ❑ Анализ сложных последовательностей и структур
- ❑ Биоинформатика, исследование ДНК, молекулярная информатика, химическая информатика
- ❑ Медицинская диагностика, лабораторные анализы
- ❑ Прогнозирование технического состояния
- ❑ Анализ финансовых рынков, технический анализ, биржевая торговля
- ❑ Обнаружение мошенничества, кредитный скоринг
- ❑ Защита от спама, адаптивные сайты, контекстная реклама
- ❑ Предсказание пользовательского поведения, рекомендательные системы
- ❑ Понимание естественного языка, смысловой поиск, машинный перевод
- ❑ Разработка программного обеспечения
- ❑ Логические и азартные игры

# Паттерны Машинного Обучения

- ❑ Machine Learning = Model Self-Learning
- ❑ Обученная модель – это программа, «написанная» данными
- ❑ Машинное Обучение невозможно без Больших Данных
- ❑ Неявное обнаружения скрытых закономерностей
- ❑ Эмпиризм, а не рационализм. Индукция, а не дедукция
- ❑ Используется классический математический аппарат
- ❑ Машинное Обучение не всеильно

# Современные тенденции Машинного Обучения

- Машинное Обучение стало операционным
- Для пользователей не требуется глубоких знаний ни в математике, ни в программировании. Глубокие знания им требуются только в собственной предметной области
- Машинное Обучение идет в Облака
- Машинное Обучение невозможно без открытого программного обеспечения и открытых данных
- Машинное Обучение невозможно без широкого сотрудничества и честного соперничества

# Рекомендательные системы без использования Машинного Обучения



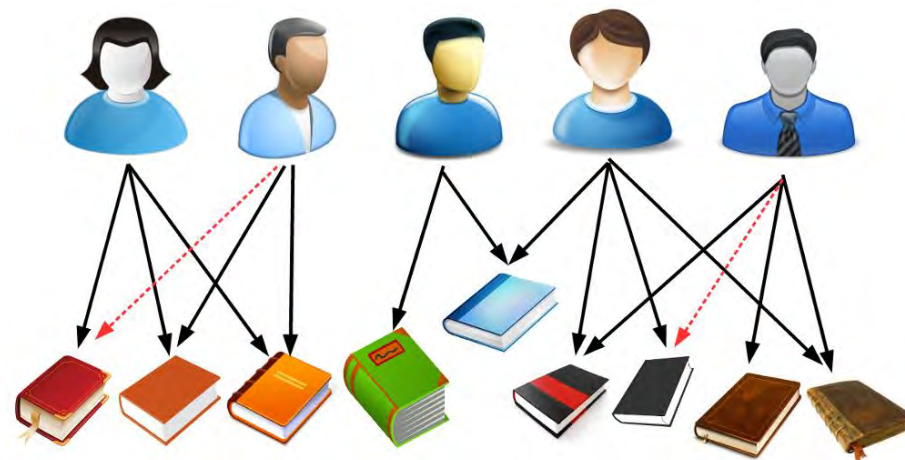


# Netflix Prize – начало эпохи Машинного Обучения в рекомендательных системах

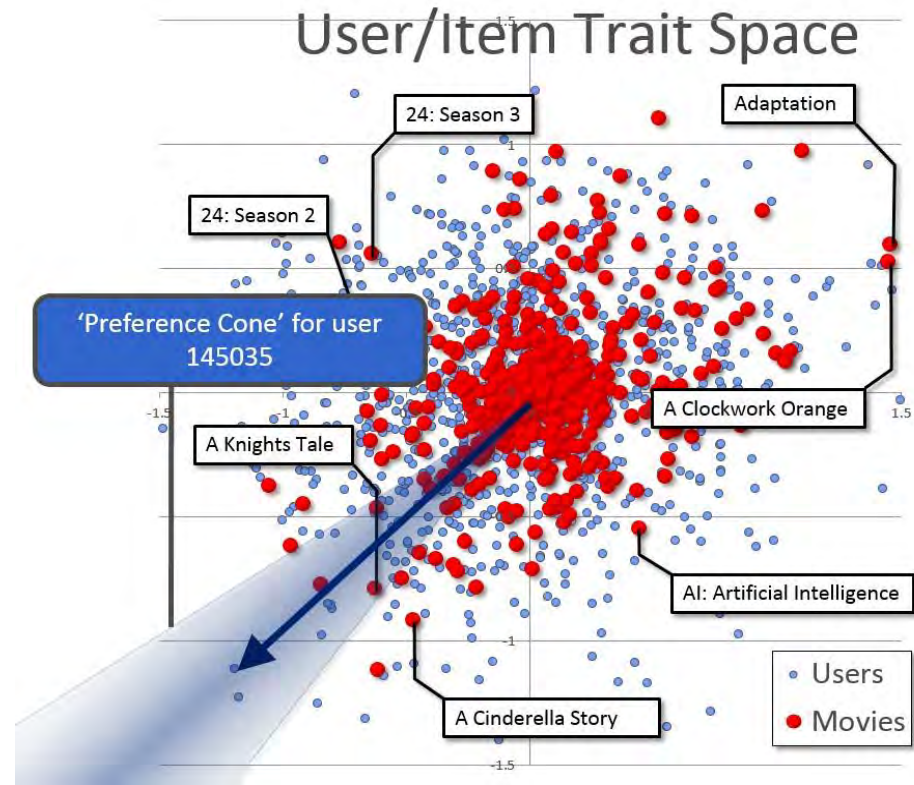
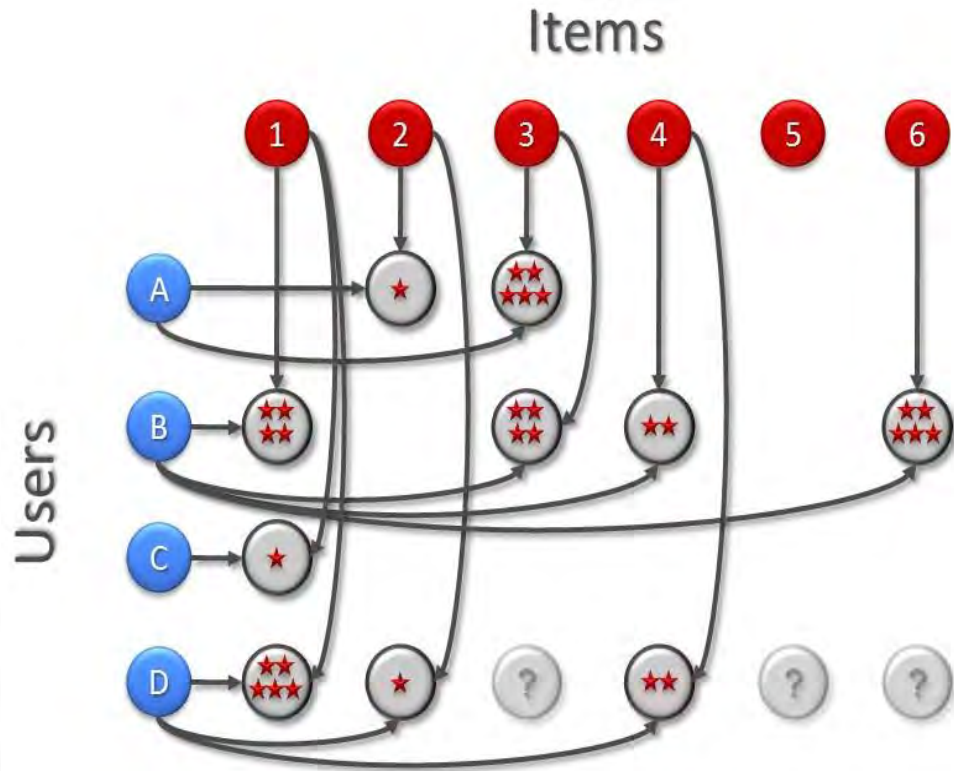


# Как работают рекомендательные системы

- ❑ Позволяют предсказать неизвестные предпочтения пользователей по известным предпочтениям других пользователей. Предсказывает, что пользователю может понравиться, на основании схожести его поведения с другими пользователями
- ❑ Обрабатывает прошлое поведение одних пользователей, их активности или оценки, чтобы строить прогнозы на будущее для других пользователей
- ❑ Позволяет рекомендовать товары/услуги, за которые пользователь/покупатель заплатит с наибольшей вероятностью



# Рекомендательные системы на базе Машинного Обучения

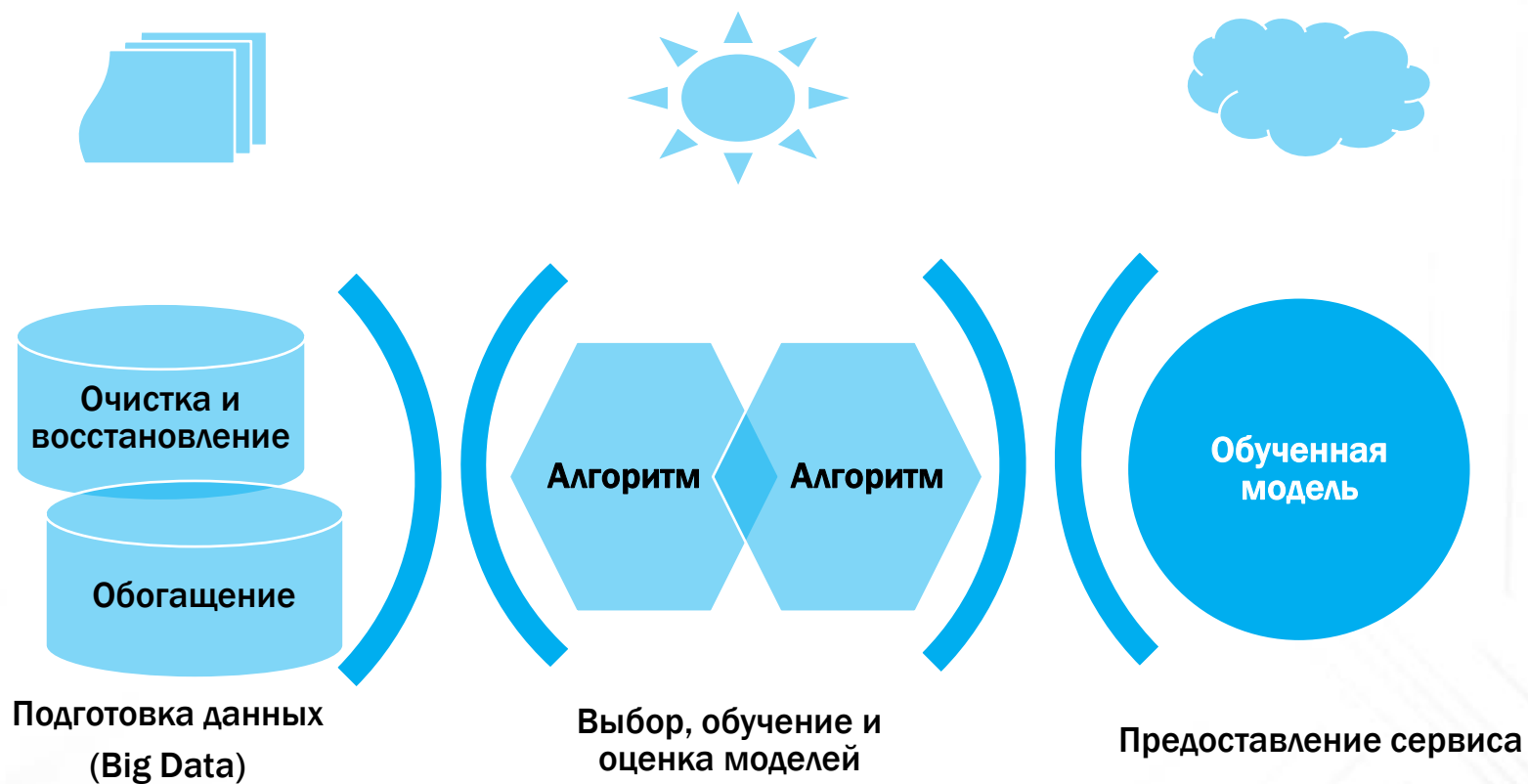




# Почему Машинное Обучение идет в Облака

- ❑ Начиная разработку, вы еще не знаете, что вам будет нужно
- ❑ Неравномерная структура потребления вычислительных ресурсов
- ❑ Необходимость проводить большое количество экспериментов и подключать различные источники данных
- ❑ Желание переложить разработку моделей на провайдеров
- ❑ Желание использовать практический опыт других

# Процесс облачного Машинного Обучения



# Как выбрать облачную платформу Машинного Обучения (список требований)

- Наличие большого набора готовых моделей и модулей предобработки данных
- Возможность бесплатно испытывать модели и наличие тестовых наборов данных для этого
- Возможность подключать различные источники данных
- Интегрированные визуальные средства разработки (для управления workflow, обработки данных, настройки моделей)
- Средства сплита данных, скоринга и оценки моделей
- Возможность дорабатывать модели, делать их тонкую настройку и подключать свои модули предобработки данных
- Возможность сразу развертывать обученную модель как облачный сервис

# Начните использовать Машинное Обучение в Облаках уже сейчас

Крупнейшие провайдеры (все с наличием бесплатных режимов работы)

Microsoft Azure Machine Learning: <https://azure.microsoft.com/ru-ru/services/machine-learning/>

Amazon Machine Learning: <https://aws.amazon.com/ru/machine-learning/>

Google Cloud Machine Learning: <https://cloud.google.com/products/machine-learning/>

TensorFlow – самая известная библиотека с открытым исходным кодом: <https://www.tensorflow.org/>

Kaggle – самая популярная краудсорсинговая платформа и площадка для проведения соревнований: <https://www.kaggle.com/>



# Выводы

- ❑ Машинное Обучение, базирующееся на Больших Данных, обеспечивает реальные конкурентные преимущества компаниям, которые его активно используют!
- ❑ Справиться со сложностью математического аппарата Машинного Обучения и потребностью в вычислительных ресурсах, при этом минимизируя затраты и время внедрения, можно с помощью Облаков!

# Координаты

Олег Фатеев

[E-mail: oleg.fateev@elpis.global](mailto:oleg.fateev@elpis.global)

Big Big Data Group (Big Data + Machine Learning)

<https://www.facebook.com/groups/bigbigdata/>

