

# Машинное обучение для HR-задач

Борис Вольфсон



**Борис Вольфсон**  
**Директор по развитию**  
**HeadHunter**  
**[boris@hh.ru](mailto:boris@hh.ru)**



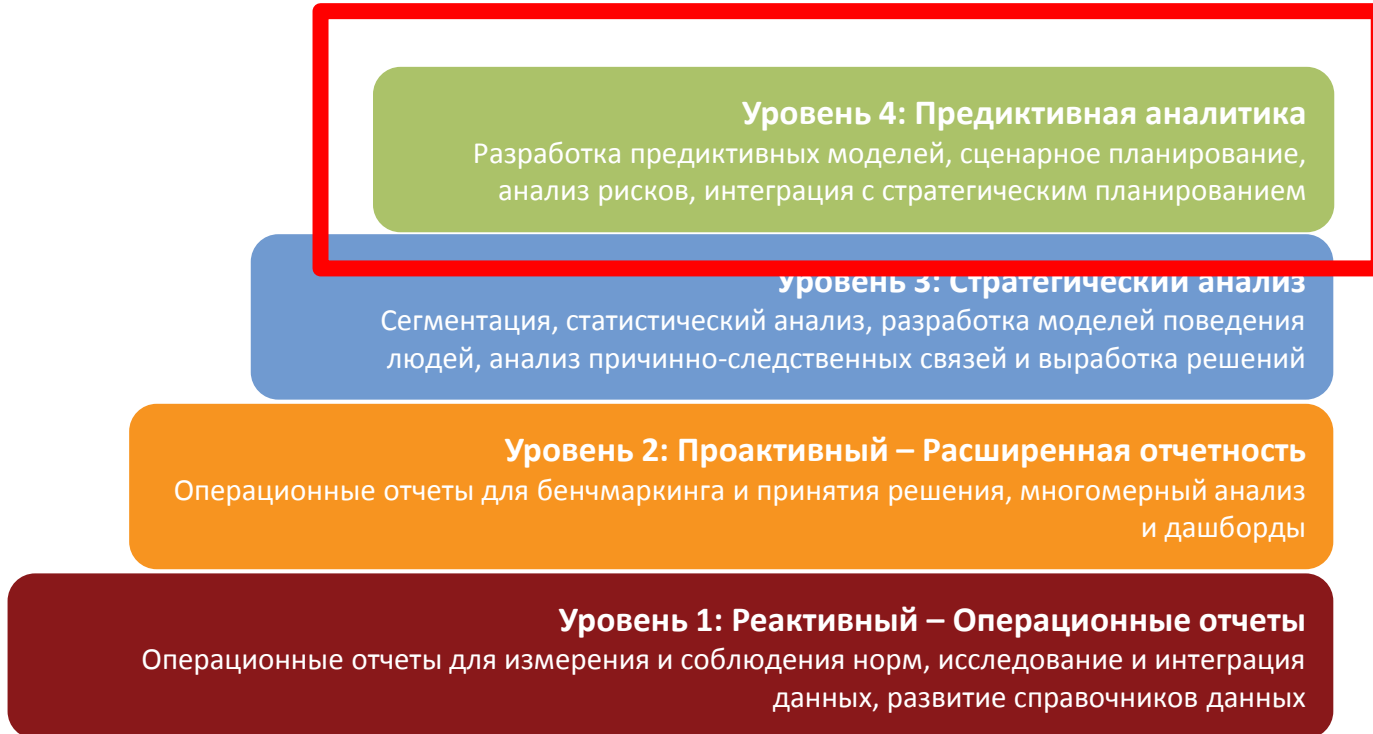



# HR Digital

Подход к организации управления  
человеческими ресурсами основанный на  
построении, автоматизации, измерении,  
**аналитики** и оптимизации процессов

*Этот термин будет со временем расширяться и включать в себя новые области*

# Модель зрелости аналитики талантов





Что представляет аудитория, когда я  
рассказываю про искусственный интеллект  
и машинное обучение

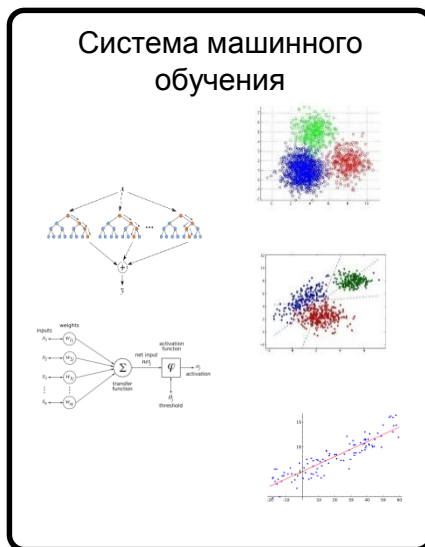


**Как будет на самом деле**

# От Excel к большим данным и машинному обучению

Данные для обучения	
_____	8123
_____	946
_____	2457
_____	2658
_____	642
_____	243
_____	86
_____	946
_____	774

Данные для предсказания	
_____	???
_____	???
_____	???



Данные с предсказанием	
_____	334
_____	563
_____	176

# Примеры использования машинного обучения

Кредитный  
скоринг

Определение  
спама

Распознавание  
изображений

Диагностика  
заболеваний

Управление  
транспортом  
без водителя

HR



# От Excel к большим данным и машинному обучению

Данные для обучения	
_____	8123
_____	946
_____	2457
_____	2658
_____	642
_____	243
_____	86
_____	946
_____	774

Данные с известной целевой метрикой:

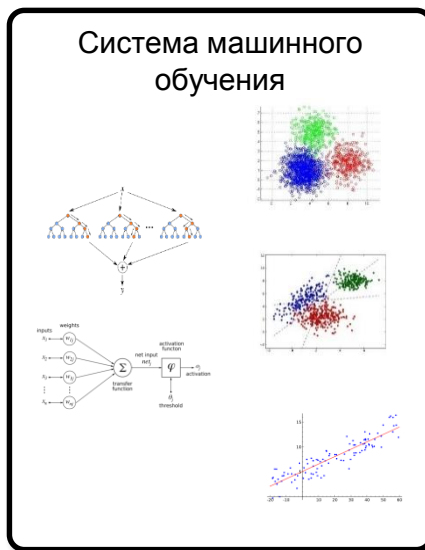
- Резюме с сроками работы
- Вакансии с ЗП

Данные для предсказания	
_____	???
_____	???
_____	???

Данные с неизвестной целевой метрикой:

- Резюме без сроков работы
- Вакансии без ЗП

# От Excel к большим данным и машинному обучению



Алгоритмы, которые строят модель для предсказаний на основе исторических данных

# От Excel к большим данным и машинному обучению

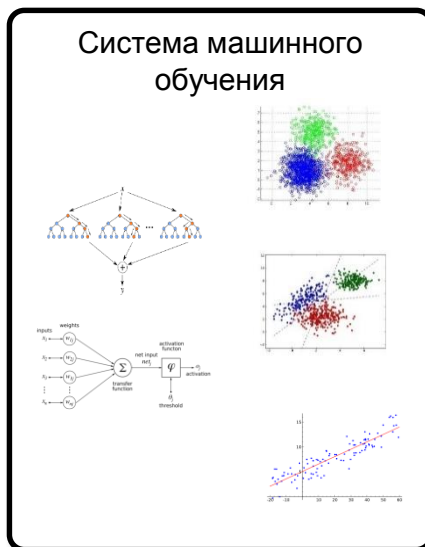
Данные, где для каждого  
элемента сделано  
предсказание моделью

Данные с предсказанием	
_____	334
_____	563
_____	176

# От Excel к большим данным и машинному обучению

Данные для обучения	
_____	8123
_____	946
_____	2457
_____	2658
_____	642
_____	243
_____	86
_____	946
_____	774

Данные для предсказания	
_____	???
_____	???
_____	???



Данные с предсказанием	
_____	334
_____	563
_____	176

## Кейс 1: Ранжирование откликов

### Задача:

- На вакансию поступили отклики, необходимо понять кого приглашать на собеседование или телефонное интервью

### Проблема

- HR-ры тратят значительное количество времени на разбор резюме и определение списка для собеседования

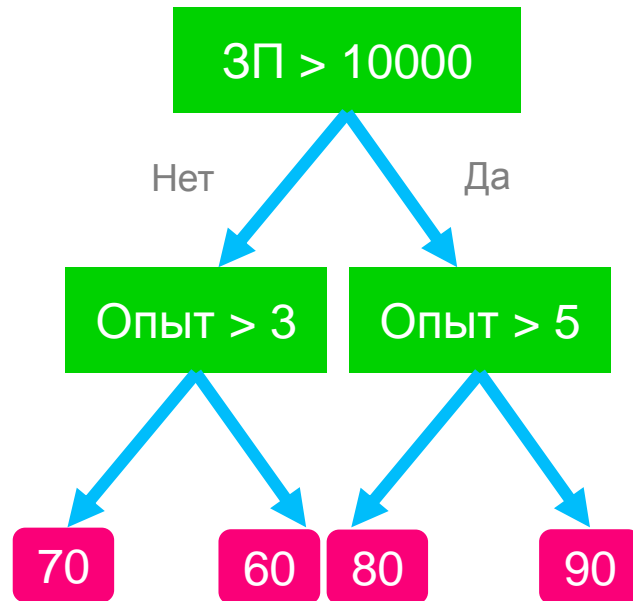
# Какие критерии вы бы использовали для ранжирование кандидатов?

Неформализованный список:

1. Опыт работы
2. Образование
3. Навыки
4. Зарплатные ожидания
5. ...
6. ...
7. ...



Дерево решений

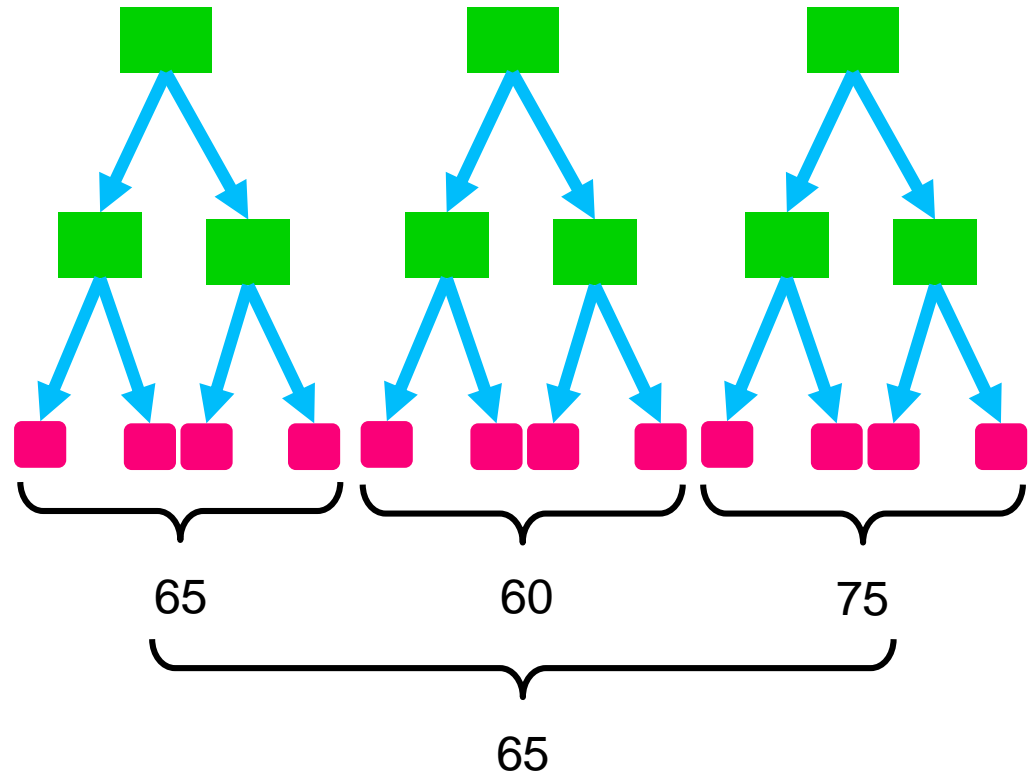


## Как принимает решение человек

1. Строим алгоритм оценки кандидатов: критерии и веса определяем экспертно
2. Прогоняем алгоритм оценки к каждому кандидату
3. Упорядочиваем кандидатов по количеству набранных баллов

# Как принимает решение машина

1. Берем исторические данные с известными результатами, т.е. мы знаем каких кандидатов пригласили на собеседование
2. На основе этих данных система машинного обучения строит лес решающих деревьев, которые вычисляют известный результат с минимальной погрешностью
3. Прогоняем через систему машинного обучения кандидатов





## Кейс 2: Прескриннинг/модерация резюме

### Задача

- Проверять качество заполнения резюме на самом верху воронки подбора

### Проблема

- Значительная часть соискателей плохо и неполно заполняют резюме

### Решение

- Использовать исторические данные по модерации резюме, чтобы автоматически подтверждать самые качественно заполненные

## Схема работы

1. Отмодерированные резюме занесены в систему в качестве исторических данных
2. Все новые резюме проверяются «роботомодератором» и если он их относит к самым качественным, то они автоматически подтверждаются.
3. Неподтвержденные резюме направляются модераторам.

**В итоге качество модерации неотлично от человека!**

**16 лет меняем мир  
рекрутинга и HR к лучшему**

**hh**