

# Машинное обучение

Немного про карьерные перспективы

# Примеры ML

- Беспилотные авто
- Кредитный скоринг
- Рекомендательные системы
- Антифрод
- Прогнозирование
- Прочая автоматизация

World's first victim of automation:



# Два пути

Наука  
(исследования)

Индустрия  
(практическое  
применение)



\$\$\$

- Стажер  
до ~60к, 40к в среднем
- Junior  
до ~100к, 70к в среднем
- Middle  
до ~250к, 150к в среднем
- Senior/Lead  
250к в среднем





# Технологии и навыки

- Python (+ библиотеки) / R
- SQL
- Иногда: C++



- Математика
- Линейная алгебра
- Теория вероятности
- Статистика
- Машинное обучение
- Разработка ПО
- Алгоритмы и структуры данных
- Понимание бизнеса

И, ВОЗМОЖНО, САМОЕ ВАЖНОЕ:

# SOFT SKILLS

# Job titles

- Data Scientist
- Machine Learning Engineer
- Deep Learning Engineer



what society thinks I do



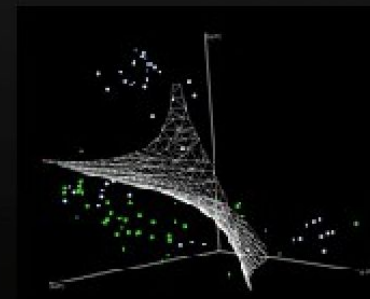
what my friends think I do



what my parents think I do

$$\begin{aligned} L_j &= \frac{1}{2} \|w\|^2 - \sum_{i=1}^n a_i y_i (x_i \cdot w + b) + \sum_{i=1}^n a_i \\ a_i &\geq 0, \forall i \\ w &= \sum_{i=1}^n a_i x_i, \sum_{i=1}^n a_i = 1 \\ \nabla_j L(\theta) &= \frac{1}{n} \sum_{i=1}^n \nabla L(x_i, y_i; \theta_i) + \nabla v_j(\theta) \\ \theta_{i,j-1} - \theta_i &= \eta \nabla L(x_i, y_i; \theta_i) - \eta v_j \cdot \nabla v_j(\theta_i) \\ \mathbb{E}_{(x_i, y_i)} [L(x_i, y_i; \theta_i)] &= \frac{1}{n} \sum_{i=1}^n L(x_i, y_i; \theta_i) \end{aligned}$$

what other programmers think I do



what I think I do

```
>>> from scipy import svm
```

what I really do

# Подробнее про тайтлы

Data scientist

- Больше бизнесовой стороны
- Анализ и визуализация данных, построение выводов
- Построение моделей

Floor is software development best practices

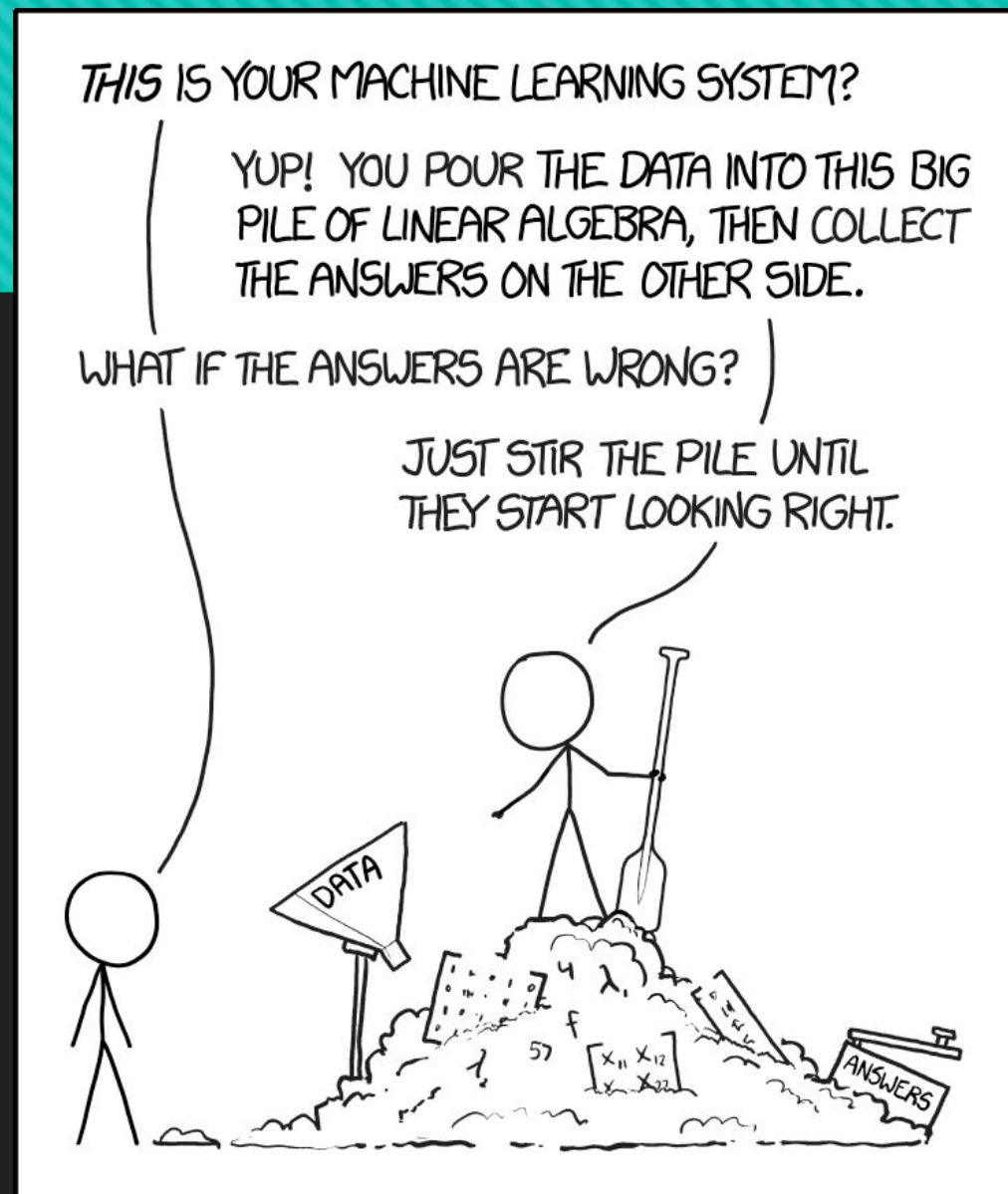




# Подробнее про тайтлы

ML Engineer

- Больше Software Engineering
- Построение моделей
- Выведение моделей в продакшн



# Образование

## Высшее образование

- Магистратура
- Долгое ожидание
- Сложно поступить
- Эффективно

## Самообразование

- Онлайн-курсы/книги
- Относительно быстро (~год)
- Можно получить первую работу раньше

# Kaggle.com

- Соревнования за призовые (до \$100-200к)
- Отлично подойдет для практики
- Медали – строчки в резюме

# OpenDataScience Slack

- 37к+ участников
- Множество интересных обсуждений
- Тонны полезной информации
- **ЭКСКЛЮЗИВНЫЕ ВАКАНСИИ**





# Курсы

- [mlcourse.ai](https://mlcourse.ai)
- “Machine learning”  
by Stanford University  
([coursera.org](https://coursera.org))