

# Машинное обучение: ошибки при работе с DS проектами

МТС Тета  
Эмели Драль

# Проектная работа

Весь объем работы можно разделить на **три** стадии:

- Предпроектное исследование
- Работа над проектом
- Работа после окончания проекта

Что может пойти не так?

# Ошибки при работе с DS проектами

1. Постановка задач и оценка эффекта
2. Метрики и критерии качества
3. Обучение моделей
4. Тестирование моделей

# Постановка задач и оценка эффекта

# Конфликт постановок

Мы измеряем успех с точки зрения бизнес-цели  
А задачу оптимизации решаем математически

Постановка  
задачи

Поэтому, оптимизируя некоторые показатели качества,  
мы надеемся оптимизировать KPI бизнеса.

# Два взгляда на задачу

Постановка  
задачи

Business goal



Math problem statement



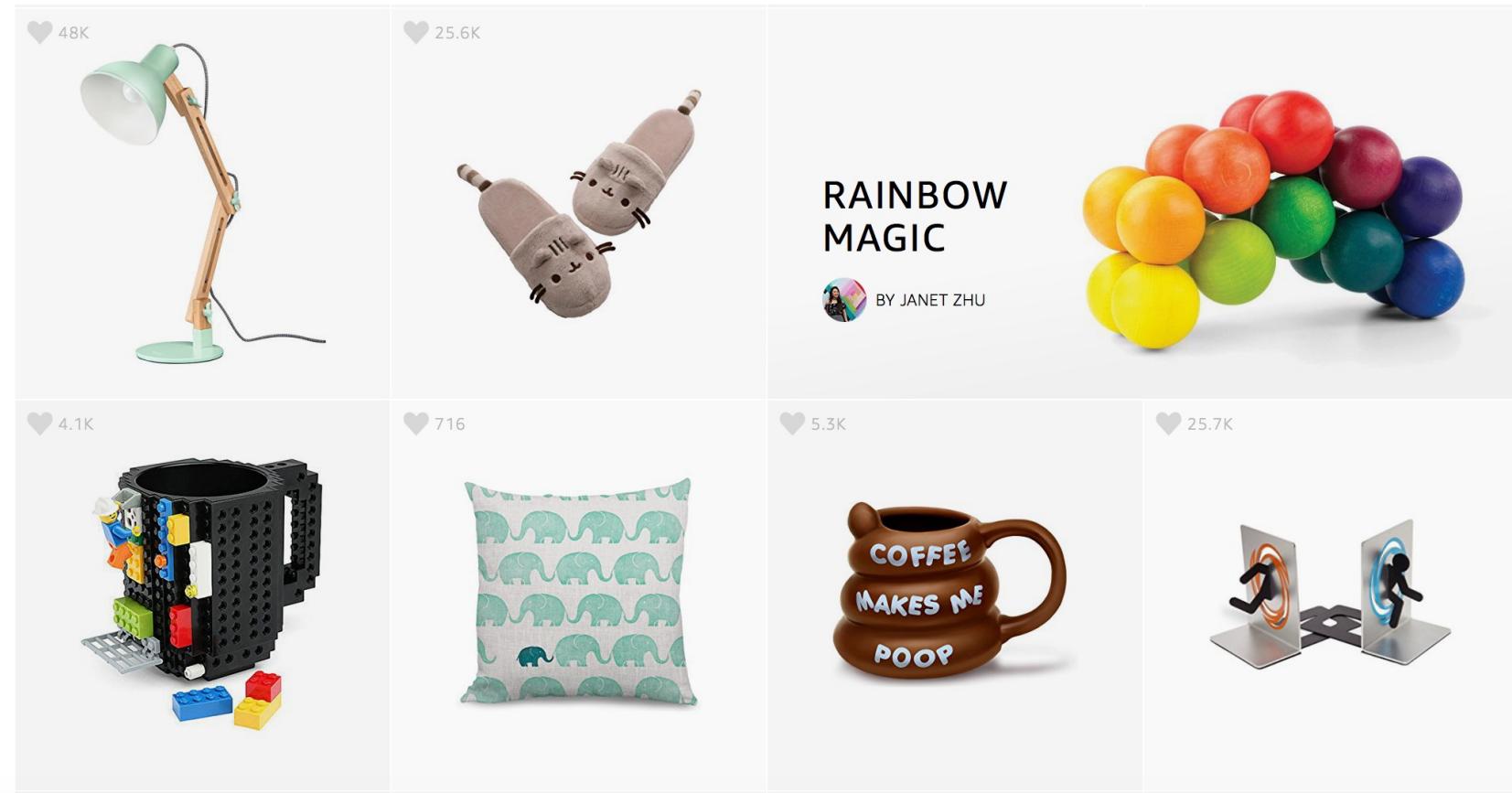
# Постановка задачи

## Постановка задачи

- Соответствует ли математическая постановка бизнес-задаче?

# Постановка задачи

## Постановка задачи



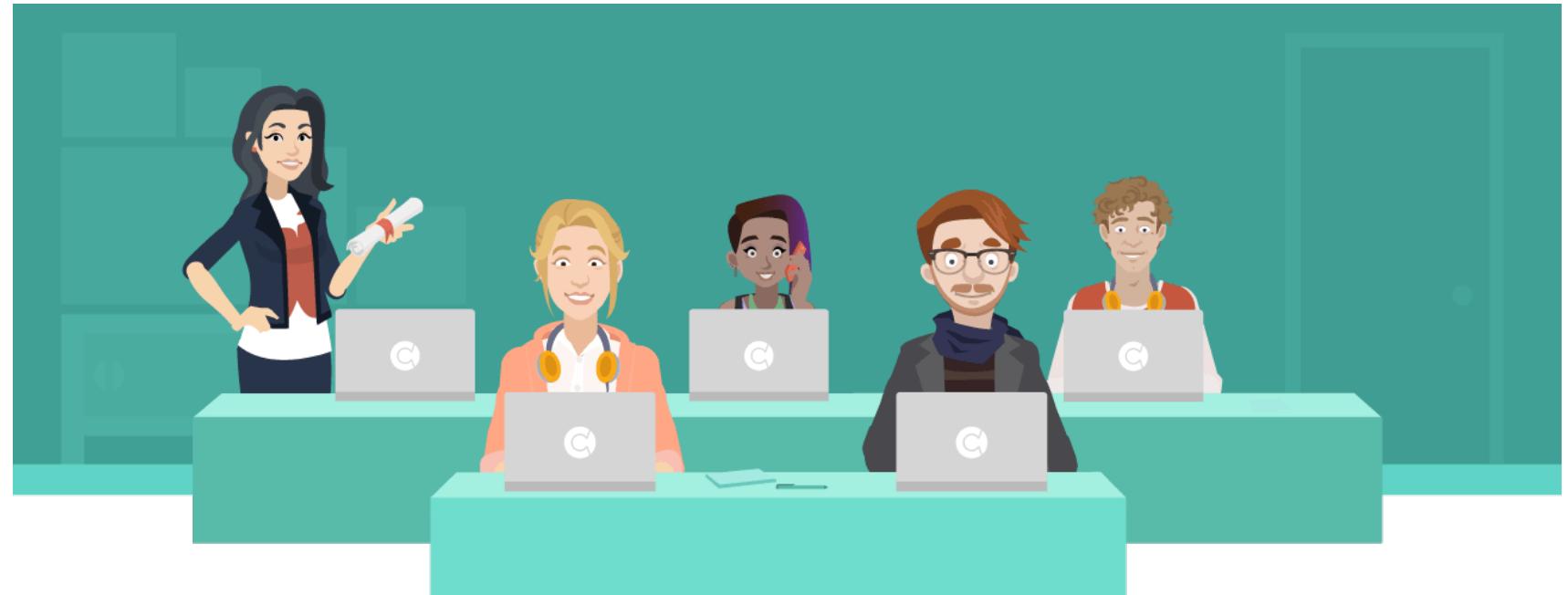
## Оценка эффекта

### Бизнес-процесс

- Нуждается ли бизнес процесс в оптимизации?

# Оценка эффекта

# Бизнес-процесс



## Оценка эффекта

### Экономическая эффективность

- Обоснован ли проект экономически?

## Оценка эффекта

# Последовательная оптимизация модели

- Are you making more money than you are spending?
- Knowing when to stop improving a once profitable model is also an important lesson to learn.

<http://www.cbronline.com/news/big-data/analytics/five-questions-every-business-must-ask-starting-machine-learning-project/>

# Определение метрик и критериев успеха

Метрики и  
критерии  
успеха

# Метрики качества



# Метрики и критерии успеха

## Метрики качества

Метрик качества очень много, одно и то же решение можно оценить сразу несколькими.

Например, метрики для качества прогноза:

- MAE
- MSE
- RMSE
- MAPE
- WAPE
- SMAPE

и пр.

# Метрики и критерии успеха

## Фиксация метрик, критериев успеха

### MAPE

- Средняя ошибка прогноза в процентах
- $MAPE = \sum_{i=1}^n \frac{|a_i - f_i|}{|a_i|}$

### SMAPE

- Симметричная средняя ошибка прогноза в процентах
- $SMAPE = \sum_{i=1}^n \frac{|a_i - f_i|}{|a_i + f_i|/2}$

# Метрики и критерии успеха

## Фиксация метрик, критериев успеха

### MAPE

- Недопрогнозирование:  
actual = 100  
forecast = 90  
mape = 10%
- Перепрогнозирование  
actual = 100  
forecast = 110  
mape = 10%

# Метрики и критерии успеха

## Фиксация метрик, критериев успеха

### MAPE

- Недопрогнозирование:  
actual = 100  
forecast = 90  
mape = 10%
- Перепрогнозирование  
actual = 100  
forecast = 110  
mape = 10%

### SMAPE

- Недопрогнозирование:  
actual = 100  
forecast = 90  
mape = 10,5%
- Перепрогнозирование  
actual = 100  
forecast = 110  
mape = 9,5%

# Метрики и критерии успеха

## Метрики качества

Важно, чтобы метрики успеха:

- Были адекватны математической постановке задачи
- Отвечали потребностям бизнеса
- Были зафиксированы до начала разработки и тестирования
- Не пересматривались в процессе или после тестирования (особенно актуально для пилотов, АБ-тестов)

# Обучение моделей

# Обучение модели

- Будет ли модель, построенная по историческим данным, работать в production?

Обучение  
модели

# Обучение модели

## Внешние данные

Moscow

Wed

Cloudy



9 °C | °F

Precipitation: 20%

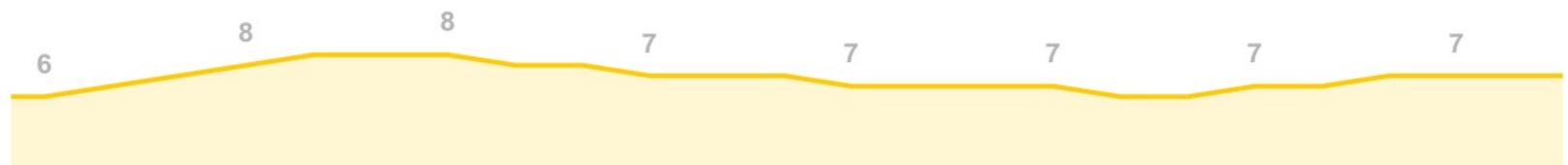
Humidity: 83%

Wind: 6 m/s

Temperature

Precipitation

Wind



11 AM

2 PM

5 PM

8 PM

11 PM

2 AM

5 AM

8 AM

Wed



9° 6°

Thu



9° 5°

Fri



10° 6°

Sat



8° 4°

Sun



7° 6°

Mon



9° 5°

Tue



9° 5°

Wed



8° 4°

# Доступность данных

Обучение модели

Schicht / Datum	Einheiten	D/02.02.03
Sorte		
FLG		
Stoffverhältnis DIP / Etik.		
V - Sieb		
V - Poproller		
Arbeitsbreite		
Stoffauflauf		
Auslaufverhältnis		
Druck		
PD Innendruck		
Druckwaage / Spülung		
Lippenöffnung		
Vorderwand		
Pumpendrehzahl		
Schüttelbock Freq.		
Schüttelbock Hub		
Duoformer D		
Scimmer / Entwässerung		
1. Zone		
2. Zone		
Obersiebentwässerung		
Druck Leiste 1		
Druck Leiste 2		
Druck Leiste 3		
Druck Leiste 4		
Druck Leiste 5		
Druck Leiste 6		
Druck Leiste 7		
Einaufwalze Duoformer		
Vakuumeinstellen		
1.Vakufoil		
2. Vakufoil / Naß		
Doppelvakuf		
Scimmer		
1.Formationszone		
2. Zone (Trockengehalt)		
Trennsauger		
Flachsäuger		
SSW		
PU Haltezone		
PU Preßzone		
Pressenpartie / Linienkraft		
1.Presse		
2.Presse		
3.Presse		
Pressmantelstellung		
Pressmantelstellung		
C		
B		
A		
Che		
D		
E		
F		
G		
H		
I		
J		
K		
L		
M		
N		
O		
P		
Q		
R		
S		
T		
U		
V		
W		
X		
Y		
Z		
Korrektur		

# Тестирование моделей

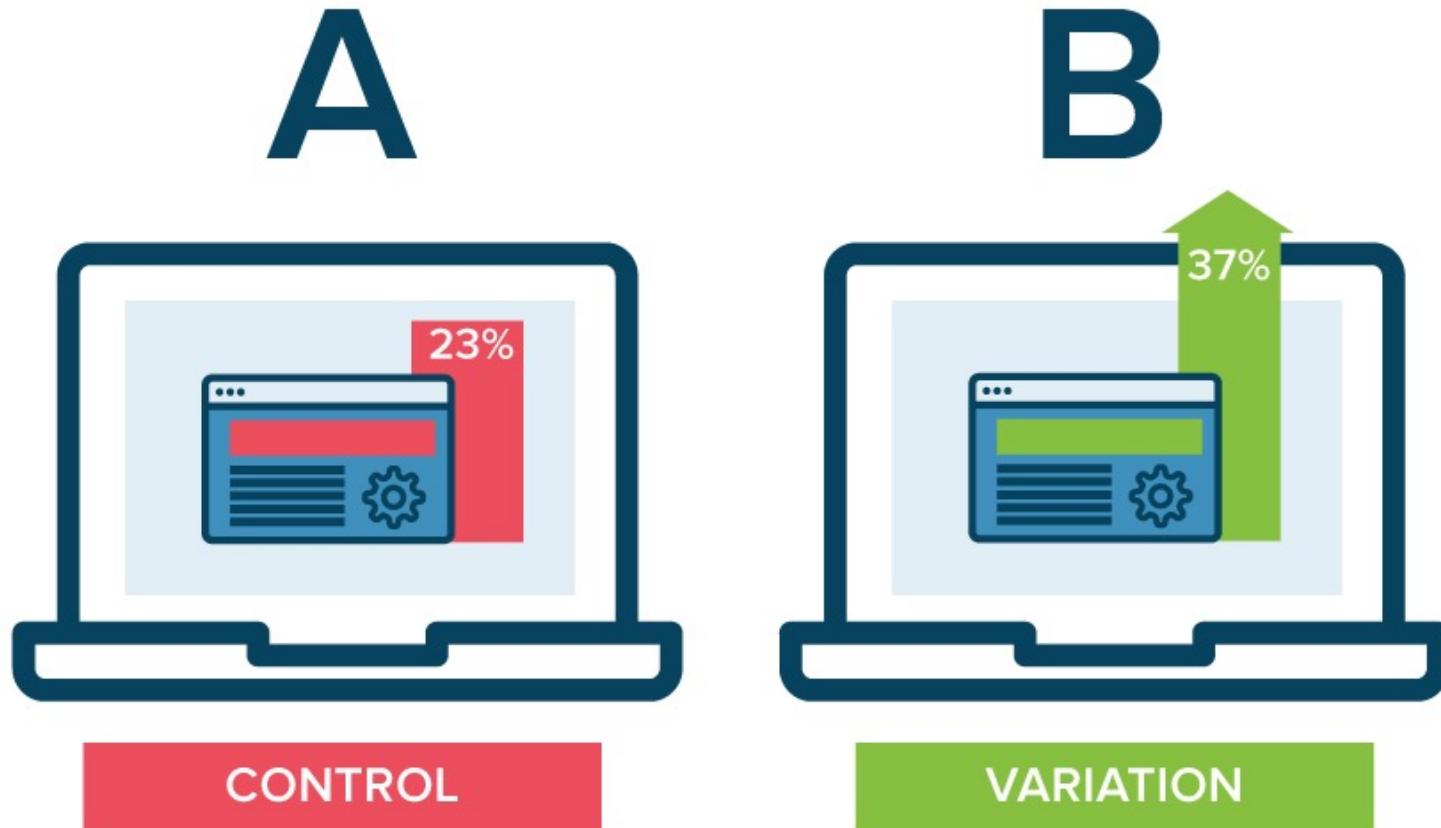
# Дизайн эксперимента

## Дизайн эксперимента

- Позволяет ли эксперимент достоверно оценить качество модели в production?

## Дизайн эксперимента

# Дизайн эксперимента



# Разработка дизайна эксперимента

Дизайн  
эксперимента



# Дизайн эксперимента

## Значимость результатов

- Будут ли результаты, полученные в процессе эксперимента, воспроизводиться в дальнейшем?

# Дизайн эксперимента

## Статистическая и практическая значимость

- Статистическая значимость: проверяется, свидетельствуют ли данные против основной гипотезы в пользу альтернативы?
- Практическая значимость: имеет ли изменение практическую (прикладную) ценность?

# Осьминог Пауль

Проверка  
гипотез



# Проверка гипотез

## Осьминог Пауль

- Гипотеза – Пауль просто осьминог
- Альтернатива – Пауль обладает предсказательной супер-способностью
- Успех – верно предсказанный исход матча
- Статистика – количество верно предсказанных матчей

## Проверка гипотез

Верно спрогнозировал исход 11  
матчей из 13!

Соперник	Турнир	Этап	Прогноз Пауля	Итог	
Польша	Евро 2008	Групповая стадия	Германия	2:0	верно
Хорватия	Евро 2008	Групповая стадия	Германия	1:2	неверно
Австрия	Евро 2008	Групповая стадия	Германия	1:0	верно
Португалия	Евро 2008	1/4 финала	Германия	3:2	верно
Турция	Евро 2008	Полуфинал	Германия	3:2	верно
Испания	Евро 2008	Финал	Германия	0:1	неверно
Австралия	ЧМ 2010	Групповая стадия	Германия	4:0	верно
Сербия	ЧМ 2010	Групповая стадия	Сербия	0:1	верно
Гана	ЧМ 2010	Групповая стадия	Германия	1:0	верно
Англия	ЧМ 2010	1/8 финала	Германия	4:1	верно
Аргентина	ЧМ 2010	1/4 финала	Германия	4:0	верно
Испания	ЧМ 2010	Полуфинал	Испания	0:1	верно
Уругвай	ЧМ 2010	Матч за 3-е место	Германия	3:2	верно

# Проверка гипотез

## Осьминог Пауль

- Результат – отклоняем  $H_0$  с уверенностью 95%
- Вывод – Пауль не просто осьминог

# Проверка гипотез

## Осьминог Пауль

- Результат – отклоняем  $H_0$  с уверенностью 95%
- Вывод – Пауль не просто осьминог
- Но на самом деле Пауль просто осьминог =(

# Проверка гипотез

## Проверка гипотез

- Позволяет проверять предположения по данным
- Не является универсальным инструментом, область применения ограничена
- Позволяет уменьшить неопределенность, но не дает гарантий
- Результаты требуют проверки на здравый смысл

Верно ли я интерпретирую результаты?

Проверка  
гипотез



# Чек-лист по работе с DS проектами

## 1. Постановка задач и оценка эффекта

- Нуждается ли сервис в оптимизации (на основе данных)?
- Соответствует ли математическая постановка продуктовой/бизнес задаче?

## 2. Метрики и критерии качества

- Как метрика качества модели связана с экономическим эффектом? Каковы ограничения на качество?
- Учитывает ли метрика качества продуктовые требования и др. ограничения, накладываемые на модель?

## 3. Обучение моделей

- Будут ли данные, используемые для обучения модели, доступны в production?
- Соответствует ли сложность модели её качеству?

## 4. Тестирование моделей

- Соответствует ли дизайн тестирования сценарию применения модели?
- Верные ли выводы сделаны по результатам тестирования?

# Прикладное машинное обучение

1. Предпроектное исследование: от постановки задачи до оценки потенциального эффекта
2. Проектная работа: оптимизация и валидация модели, демо-стенд, разработка сервиса
3. Поддержка и сопровождение сервиса. Чек-лист data-саентиста.

# Спасибо за курс!

Telegram: <https://t.me/emeliml>

Twitter: <https://twitter.com/EmeliDral>

LinkedIn: <https://www.linkedin.com/in/emelidral/>