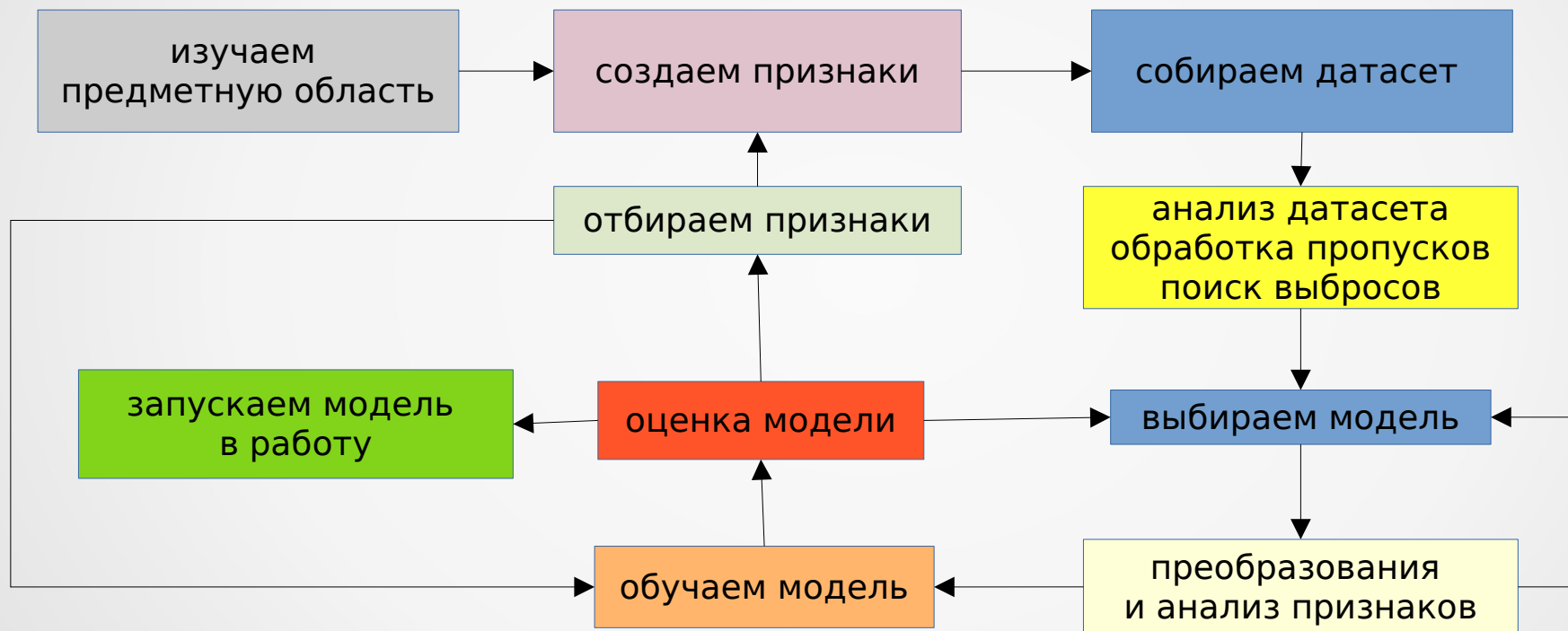




# **О методах машинного обучения**

Евгений Борисов

# О методах машинного обучения



# схема применения методов ML

**создаем признаки** ( *feature extraction / feature engineering* )

отображение данных, специфических для предметной области,  
в точки пространства признаков

## Типы признаков

- бинарные (да/нет)
- категориальные
- количественные ( $\mathbb{R}$ )
- порядковые

## примеры признаков

для текстов

- TF-IDF
- Word2Vec

для изображений:

- Haar-like features,
- HOG (Histogram of Oriented Gradients)

собираем признаки формируем учебный датасет

# О методах машинного обучения

## Основные типы задач ML

Классификация - разделение на части

Кластеризация - формирование групп

Регрессия - восстановление зависимости

# О методах машинного обучения

## Методы решения задач машинного обучения

### способы организации данных

"с учителем" (supervised)

- размеченные данные

"без учителя" (unsupervised)

- не размеченные данные

"частичное обучение" semi-supervised

- частично размеченные данные

"с подкреплением" reinforcement

- датасет в явном виде отсутствует

# О методах машинного обучения

## Методы решения задач машинного обучения

### способы организации данных

"с учителем" (supervised)

- размеченные данные

"без учителя" (unsupervised)

- не размеченные данные

"частичное обучение" semi-supervised

- частично размеченные данные

"с подкреплением" reinforcement

- датасет в явном виде отсутствует

### модели

Метрические: k-Neighbors

Статистические: Naive Bayes

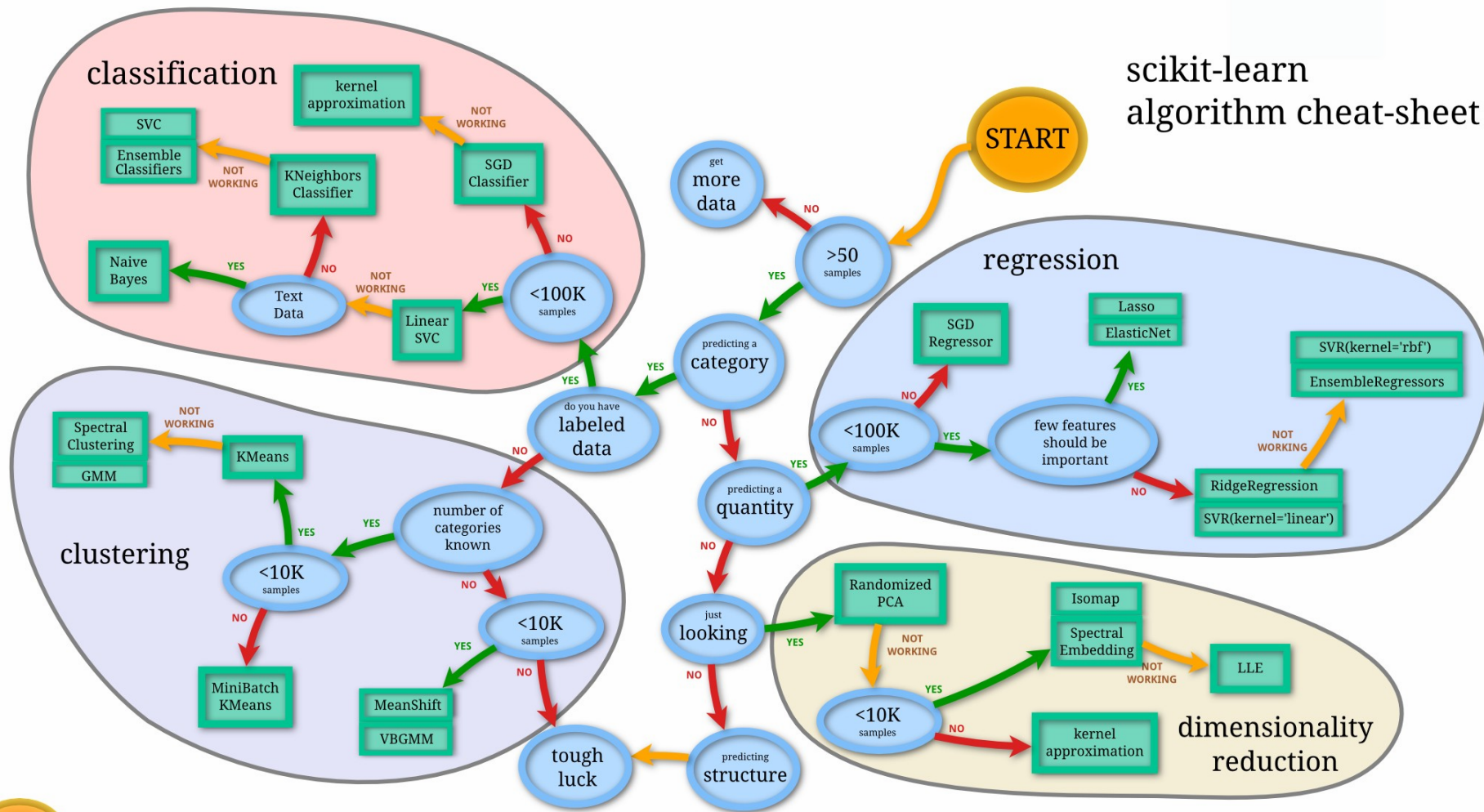
Логические: Decision Tree

Линейные: SVM, MLP

Композиции: AdaBoost

# О методах машинного обучения

scikit-learn  
algorithm cheat-sheet



# Литература

git clone [https://github.com/mechanoid5/ml\\_lectorium.git](https://github.com/mechanoid5/ml_lectorium.git)