Machine Learning:
the main goal,
some examples and facts

Mikhailova Darya,
a third-year student of MIPT

ABBYY®

ОИВТ

Generalizing is the main idea of ML

MAGIC

Available data →

(machine learning algorithm)
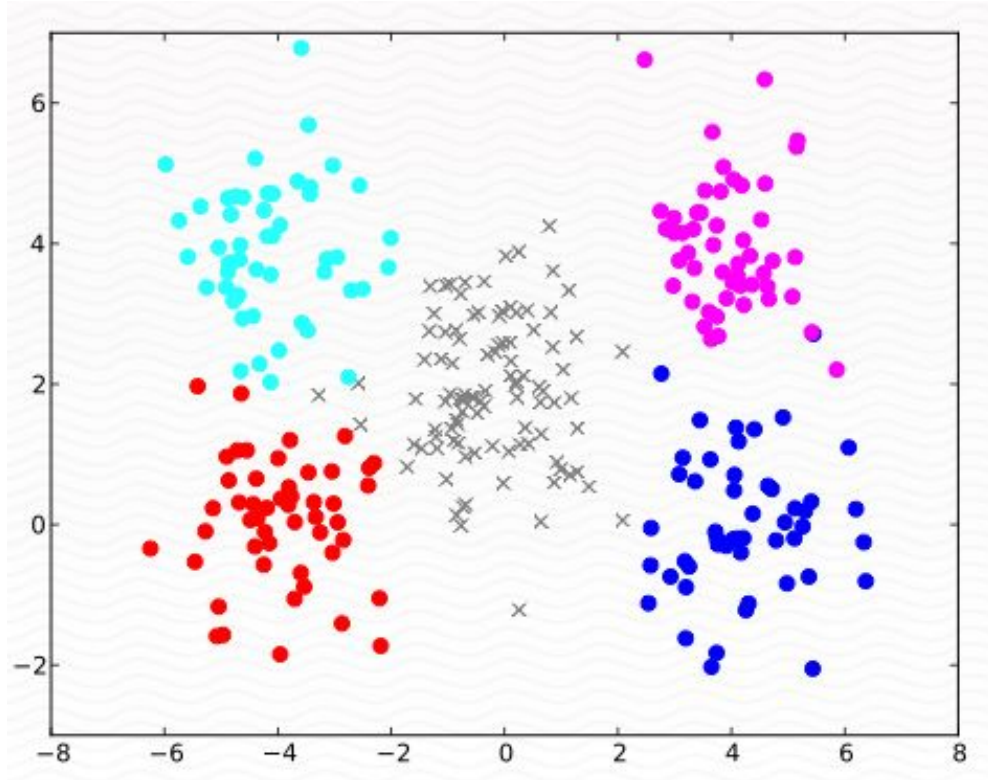
Prediction for new data

Automatically learning is advisable!

Machine learning is everywhere!!!

Classification is the most mature type of ML
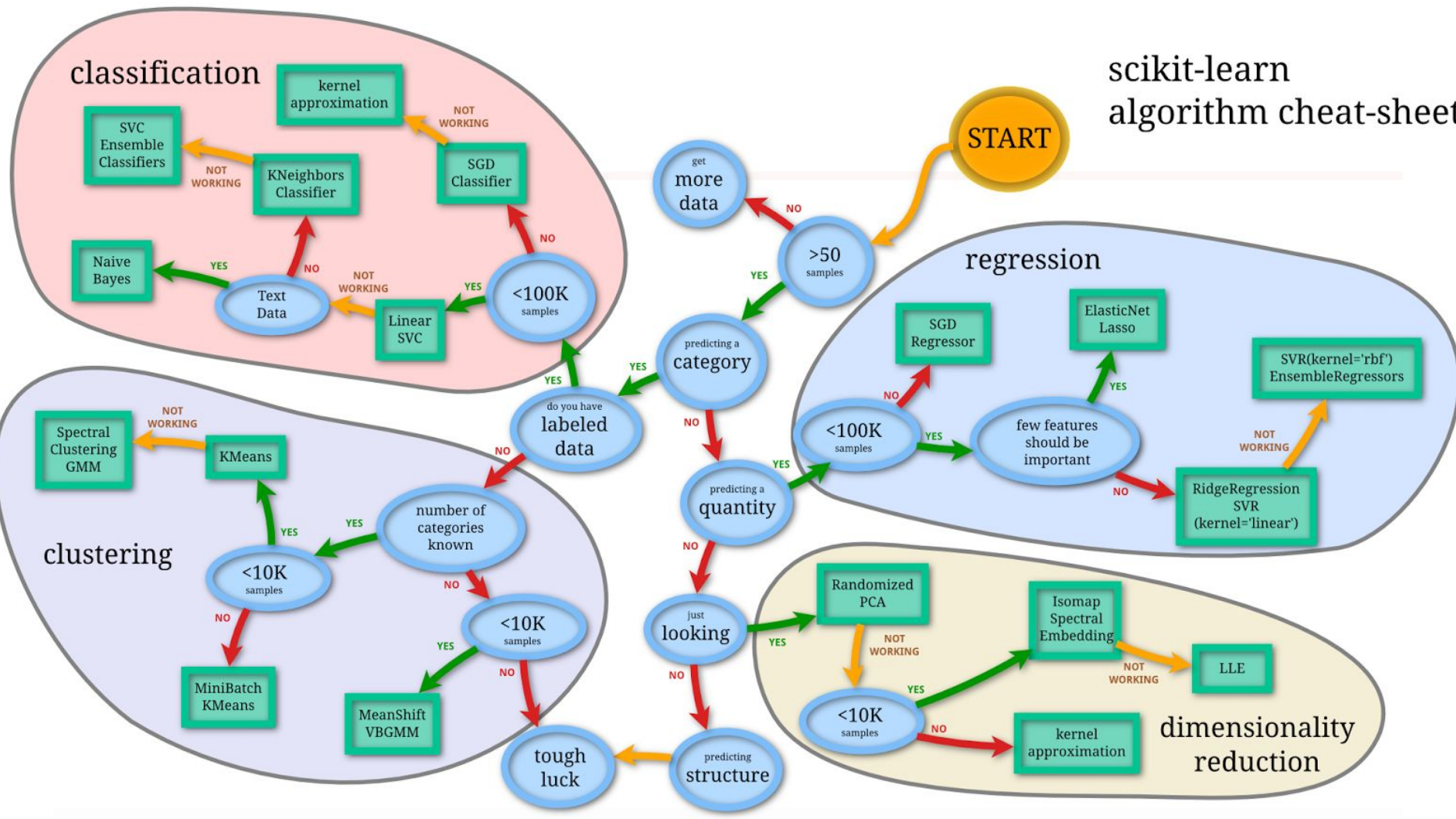
Classification is the most mature type of ML

$X_j = 1$
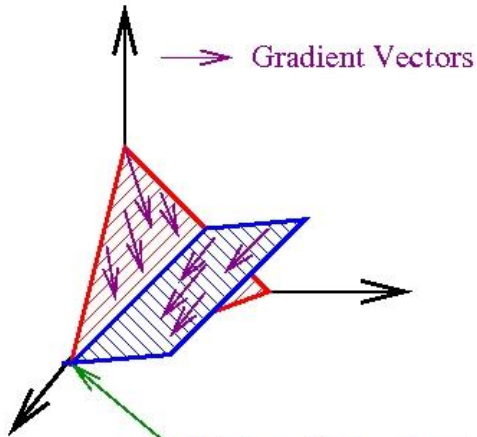*if the j-th word in the dictionary appears in the email*
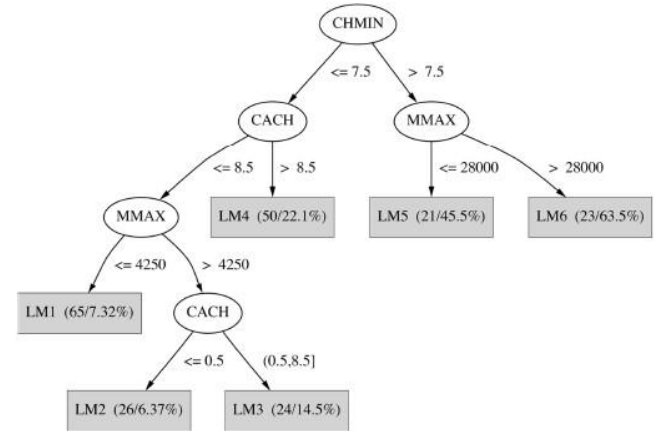
$X_j = 0$
otherwise

scikit-learn algorithm cheat-sheet

classification

kernel approximation

SVC Ensemble Classifiers

KNeighbors Classifier — NOT WORKING

SGD Classifier — NOT WORKING

Naive Bayes — YES

Text Data — NO

Linear SVC — NOT WORKING

<100K samples — YES

START

>50 samples — NO — get more data

predicting a category — YES

do you have labeled data — YES

regression

SGD Regressor

ElasticNet Lasso

SVR(kernel='rbf') EnsembleRegressors

<100K samples — NO / YES

few features should be important — YES / NO

RidgeRegression SVR (kernel='linear') — NOT WORKING

predicting a quantity — YES

clustering

Spectral Clustering GMM — NOT WORKING

KMeans — YES — YES

number of categories known — NO

<10K samples — YES / NO

<10K samples — YES / NO

MiniBatch KMeans

MeanShift VBGMM

just looking — YES — NO

tough luck

predicting structure

dimensionality reduction

Randomized PCA — NOT WORKING

Isomap Spectral Embedding — NOT WORKING — LLE

<10K samples — YES / NO

kernel approximation

The main steps of selecting:

- Representation

- Evaluation

- Optimization

> Gradient Vectors

Minima. Note: almost no gradient vectors point here

Examples of this three components

Representation	Evaluation	Optimization
Instances	Accuracy/Error rate	Combinatorial optimization
K-nearest neighbor	Precision and recall	Greedy search
Support vector machines	Squared error	Beam search
Hyperplanes	Likelihood	Branch-and-bound
Naive Bayes	Posterior probability	Continuous optimization
Logistic regression	Information gain	Unconstrained
Decision trees	K-L divergence	Gradient descent
Sets of rules	Cost/Utility	Conjugate gradient
Propositional rules	Margin	Quasi-Newton methods
Logic programs		Constrained
Neural networks		Linear programming
Graphical models		Quadratic programming
Bayesian networks		
Conditional random fields		

Today we have known:

✓ What machine learning is

✓ Where machine learning is used

✓ What the criteria for selection algorithms exist.

Thank you for your attention!