

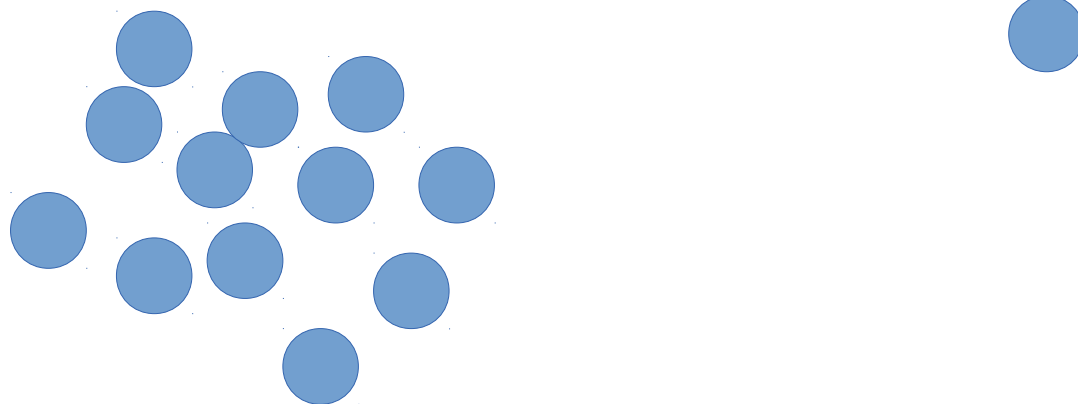
# Scikit-learn ゼミ

## 1.2.3. Density estimation, novelty detection

新納浩幸

# 外れ値検出

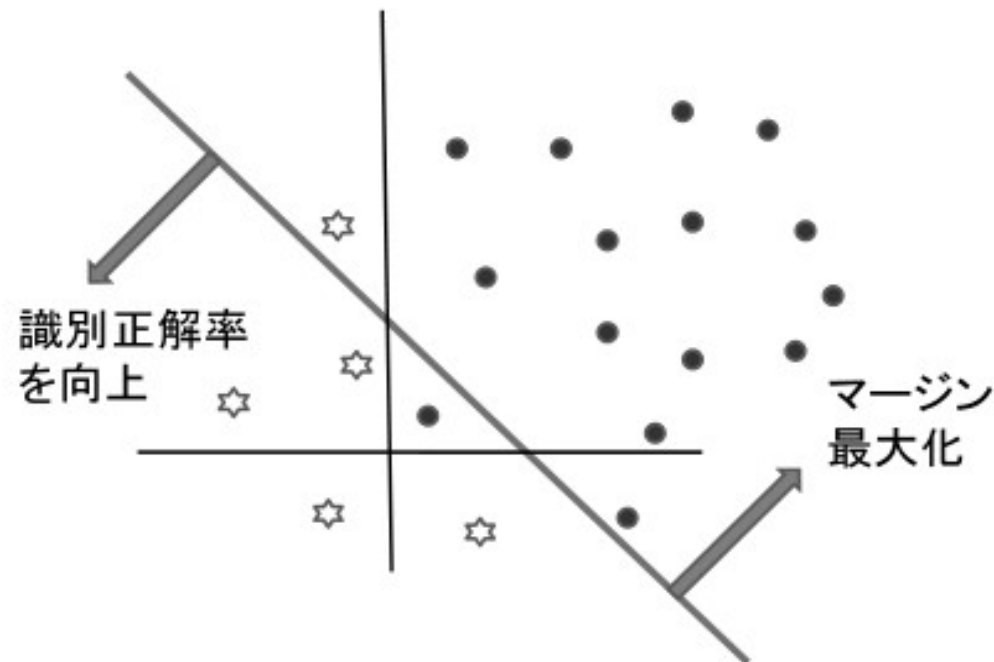
データ集合の中から特異なデータを見つける



# One Class SVM

SVM を利用した外れ値検出手法

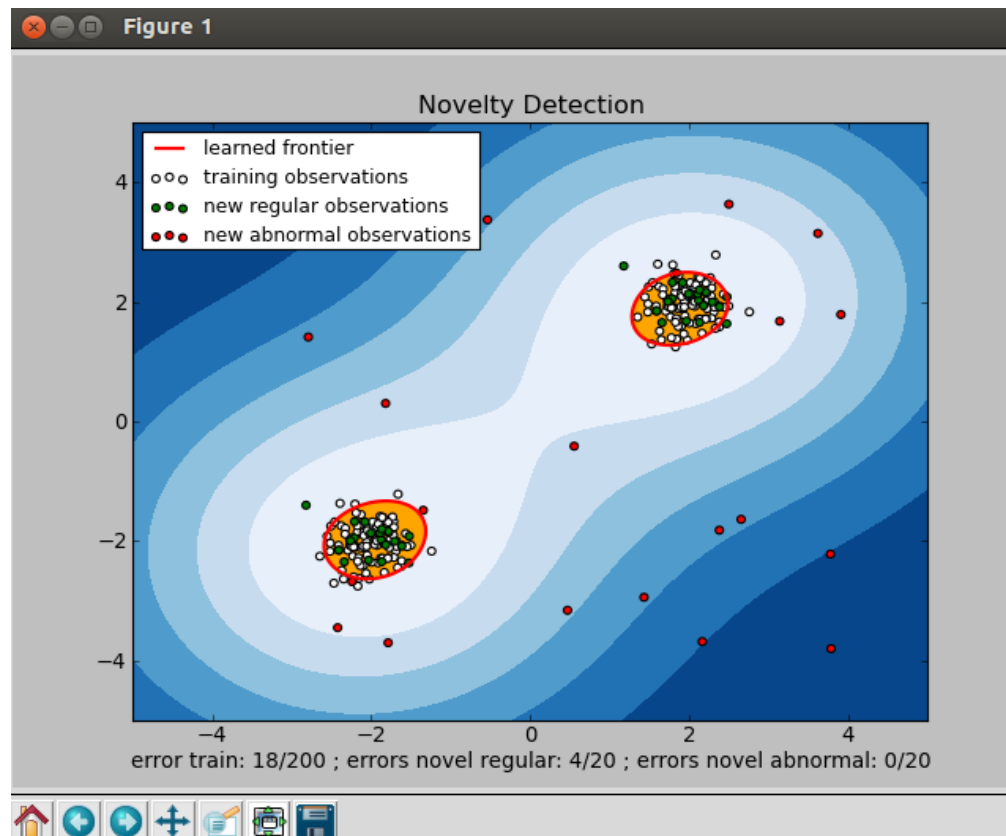
全てのデータを +1 のクラス、原点のみを -1 のクラス  
ソフトSVM で識別面を学習し、全データを識別、  
-1 のクラスが外れ値



# OneClassSVM

sklearn で提供されている One Class SVM の関数

[http://scikit-learn.org/stable/auto\\_examples/svm/plot\\_oneclass.html#example-svm-plot-oneclass-py](http://scikit-learn.org/stable/auto_examples/svm/plot_oneclass.html#example-svm-plot-oneclass-py)



# 簡易な自作の例

```
import random

import numpy as np
from sklearn import svm

data = []
for i in range(100):
    x = random.gauss(0, 1)
    y = random.gauss(0, 1)
    data = data + [[x, y]]

for i in range(10):
    x = random.uniform(-3, 3)
    y = random.gauss(-3, 3)
    data = data + [[x, y]]

clf = svm.OneClassSVM(nu=0.1, kernel="rbf", gamma=0.1)
clf.fit(data)
ans = clf.predict(data)

for a in ans:
    print a
```