

## Project 3

Make a classification dataset as following

```
from sklearn.datasets import make_blobs
X_2, y_2 = make_blobs(n_samples = 100, n_features = 2, centers = 8,
                      cluster_std = 1.3, random_state = 4)
y_2 = y_2 % 2
```

1. Split data into training data and test data by (25 points)

```
X_train, X_test, y_train, y_test = train_test_split(X_2, y_2, random_state=0)
Scatter plot training dataset and test dataset on one figure
```

2. Write a program that fits a SVM model (sklearn.svm.SVC) on the *training data* X\_train for linear kernel, RBF kernel, and Polynomial kernel with degree of 3 (using default parameter setting). For each model, plot the decision boundary and class region like the figure below. (50 points)
3. For linear model, use C = 0.01, 0.1, 1, 10, 100 respectively and calculate the fitting score for test data for each case. What conclusion you might draw? For RBF model, Use C=0.1, 1, 10 and gamma = 0.01, 1, 10, 100 and calculate fitting score for test data. What conclusion you might draw? For polynomial model with degree of 3, use parameters C = 0.1, 1, 10 and degree = 2, 3, 4, 5 and calculate fitting score for test data. What conclusion you might draw? (25 Points)



