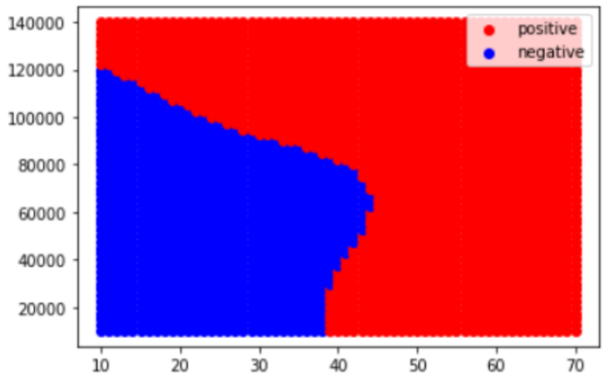
Instruction for HW 5

Due date: Oct 14th, Monday, 11:59 PM

1) Prepare the data:
   a) Import the libraries
   b) Load the dataset
   c) Split the train set and the test set, with random_state = 0, test_size = 0.2
   d) Apply feature scaling (standard scaler)

2) Train the models:
   a) Use train set as the input of GridSearchCV to find the best parameters of: (set the parameter grid as you wish)
      1. Logistic Regression Classifier
      2. KNN (KNeighborsClassifier)
      3. SVC (svm.LinearSVC)
      4. Kernel SVC (svm.SVC)
      5. Decision Tree Classifier
      6. Random Forest Classifier
      7. Naive Bayes Classifier
      8. XGBoost (xgboost.XGBClassifier)
      9. CatBoost (catboost.CatBoostClassifier)
   b) Use each model with the best parameters to predict the test set.

3) Evaluate the model and visualization:
   a) Calculate confusion_matrix and classification_report in sklearn.metrics for training results and test results.
   b) Visualize the result:
      i) Scatter the train result, set the TP, FP, FN, TN in different colors.
      ii) Scatter the test result, set the TP, FP, FN, TN in different colors.
      iii) Visualize the classification boundary of the model, for example:



   c) Compare the performances of those models and write a brief report.

You can put all your code into one script. Please submit both IPYNB and PDF files.