

Начало

чтение датасета в переменную data

разделение данных на тестовую и обучающую выборки в отношении 20-80
X_test, X_train, Y_test, Y_train

применить алгоритм векторизации к
X_train и X_test, получив vec_train и
vec_test

обучить классификатор clf на
vec_train и Y_train

посчитать f-мера и accuracy
классификатора clf на vec_test и
y_test

конец