

# Улучшение модели

## Разделение на тренировочную и валидационную выборку

```
In from sklearn.model_selection import train_test_split

df_train, df_valid = train_test_split(df, test_size=0.25, random_state=12345)

# test_size – доля валидационной выборки от общего числа объектов, число от 0 до 1
```

## Инициализация модели случайного леса

```
In from sklearn.ensemble import RandomForestClassifier

model = RandomForestClassifier(random_state=12345, n_estimators=3)

# n_estimators – количество деревьев
```

## Инициализация модели логистической регрессии

```
In from sklearn.linear_model import LogisticRegression

model = LogisticRegression(random_state=12345)
```

# Словарь

### Валидационная выборка

набор данных, извлекаемый из исходного датасета, на которой проверяется качество работы алгоритма во время обучения модели

### Гиперпараметры

настройки алгоритмов обучения

### Параметры

настройки модели, определяющие её работу и получаемые из обучающих данных