In [ ]:
【ハイパーパラメタ・チューニング】
問6
(まず)パラメータチューニングなしでXGBoostで予測をし精度検証をしてみてください。
※ XGBoostの使い方は、SVCやランダムフォレストと同じです
(まずインポートします)->from xgboost import XGBClassifier
(次に、インスタンスを作ります)->xgb = XGBClassifier()
あとは'xgb'で.fit()します

In [2]:
```
import pandas as pd

df = pd.read_csv('train.csv')
df = df.drop(['PassengerId','Age','SibSp','Parch','Ticket',
              'Cabin','Name'], axis=1)
df['Embarked'] = df['Embarked'].fillna('S')
df['Embarked'] = df['Embarked'].map({'S': 0, 'C': 1, 'Q': 2})
df['Sex'] = df['Sex'].apply(lambda x: 0 if x=='male' else 1)
df.head()
```

Out[2]:

	Survived	Pclass	Sex	Fare	Embarked
0	0	3	0	7.2500	0
1	1	1	1	71.2833	1
2	1	3	1	7.9250	0
3	1	1	1	53.1000	0
4	0	3	0	8.0500	0

In [3]:
```
from xgboost import XGBClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score

x = df.drop('Survived', axis=1)
y = df['Survived']
x_train, x_test, y_train, y_test = train_test_split(x, y,
                                                    test_size=0.3,
                                                    random_state=0)


xgb = XGBClassifier()

xgb.fit(x_train, y_train)
```

Out[3]:
```
XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=
1,
              colsample_bynode=1, colsample_bytree=1, gamma=0, gpu
_id=-1,
              importance_type='gain', interaction_constraints='',
              learning_rate=0.300000012, max_delta_step=0, max_dep
th=6,
              min_child_weight=1, missing=nan, monotone_constraint
s='()',
              n_estimators=100, n_jobs=0, num_parallel_tree=1, ran
dom_state=0,
              reg_alpha=0, reg_lambda=1, scale_pos_weight=1, subsa
mple=1,
              tree_method='exact', validate_parameters=1, verbosit
y=None)
```

In [4]:
```
pred = xgb.predict(x_test)
acc = accuracy_score(pred, y_test)
print('accuracy score : {:.5f}' .format(acc))
```

```
accuracy score : 0.81343
```

In [ ]:
```
# 今回の場合。チューニングなしではランダムフォレストの方がいいようですね
```