Dataset

The passengers on titanic,with features consisted of sex,pclass,age etc.The goal is to predict whether they survived or not.

∞ PassengerId	# Survived	# Pclass	▲ Name	▲ Sex	# Age	# SibSp
1	0	3	Braund, Mr. Owen Harris	male	22	1
2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Thayer)	female	38	1
3	1	3	Heikkinen, Miss. Laina	female	26	0
4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35	1
5	0	3	Allen, Mr. William Henry	male	35	0
6	0	3	Moran, Mr. James	male		0

Download:https://www.kaggle.com/competitions/titanic/

Scaling the features

Sex:male,female

Pclass:1,2,3

Age:divided to intervals,[0,10],[0,20],[0,30]...

Sibsp:0 or >=1

Parch:0 or >=1

Port of Embarkation:C,Q,S

```cpp
int binaryFeature(passenger& p)
{
    int f = 0;

    /* 0-1 bits for sex */
    f |= 1 << p.sex;
    /* 2-4 bits for pclass */
    f |= 1 << (p.pclass + 1);

    /* 5-14 bits for age */
    for (int i = 1; i <= 10; i++) {
        if (p.age <= i * 10) {
            f |= 1 << (i + 4);
        }
    }

    /* 14-bit for sibsp */
    if(p.sibsp) {
        f |= 1 << 15;
    }

    /* 15-bit for parch */
    if(p.parch) {
        f |= 1 << 16;
    }

    /* 16-18 bits for port C,Q or S */
    f |= 1 << (p.port + 17);

    return f;
}
```

Accuracy

The dataset has been split into two groups,training set and test set.

Accuracy = the proper predictions / the number of test cases

```
Correct predictions: 356
Total test cases: 418
Accuracy: 0.851675

Process returned 0 (0x0)   execution time : 0.739 s
Press any key to continue.
```

Thinking

When choose age intervals as [0,10),[10,20),[20,30),...

```
Correct predictions: 346
Total test cases: 418
Accuracy: 0.827751

Process returned 0 (0x0)   execution time : 1.222 s
Press any key to continue.
```

When choose age intervals as [0,100],[10,100],[20,100],...

```
Correct predictions: 351
Total test cases: 418
Accuracy: 0.839713
```