

# 题目背景

目前现网中恶意软件猖獗，数量级别不断抬高，但传统检测方法效率太低，故我司急寻安全人才帮助检测。我们收集了大量的恶意软件并进行了数据清洗及预处理工作。并提取出所有二进制文件的gdl文件。

gdl文件由两部分构成：node和edge；

- node中的title是唯一标识，label是函数名称
- edge中的sourcename是源调用函数，targetname是被调用函数，两个值均与node中的title关联

请根据黑白文件夹中给定的数据信息进行二分类

# 提交答案

两列，第一列是文件名（不带后缀），第二列是该文件label（黑为1，白为0），以空格间断，最后一行不要留空行

例子：

```
0b3cc62a77d2e720ebe93fe352e88f6e 0
0dc03cb8b5d79303fff2869826d501ee 1
```

比赛期间随时可以提交答案，提交后10分钟内给出成绩

# 评分标准

混淆矩阵：

预测/真实 ↕	恶意软件 ↕	正常软件 ↕	↕
恶意软件 ↕	TP ↕	FP ↕	↕
正常软件 ↕	FN ↕	TN ↕	↕

召回率：

$$recall = TP / (TP + FN)$$

误报：

$$FPR = FP / (FP + TN)$$

评分标准:

$$\text{评分标准} = \text{召回} - 1.0 * \text{误报}$$