三种恶意代码：病毒、蠕虫、木马

特征提取方法：N-grams、可输出字符串、api

特征选择方法：信息增益（IG）、文档频率（DF）、频繁项集（FP-G’rowth）

分类算法：决策树、支持向量机、朴素贝叶斯、k邻近

评价方法：

1. 真正类TP（恶意代码正确判定）、真负类TN（正常代码正确判定）、假正类FP（正确代码错误判定）、假负类FN（恶意代码错误判定）

正确率=

误报率=

总体准确率=