目录

[1. 引言 2](#_Toc81314672)

[1.1 编写目的 2](#_Toc81314673)

[1.2 开发目的及意义 2](#_Toc81314674)

[1.3 预期读者和阅读建议 2](#_Toc81314675)

[2. 系统概述 3](#_Toc81314676)

[2.1 运行环境 3](#_Toc81314677)

[2.2 设计和执行约束 3](#_Toc81314678)

[3. 功能需求 3](#_Toc81314679)

[3.1 系统功能列表 3](#_Toc81314680)

[3.2 功能需求 4](#_Toc81314681)

[4. 非功能需求 4](#_Toc81314682)

[5. 性能需求 4](#_Toc81314683)

[6.用户需求 5](#_Toc81314684)

# 引言

## 编写目的

此份软件产品需求分析报告是为识别恶意代码系统编写的，开发这个软件旨在建立一个高效率、高准确率、通用性好的恶意代码系统，通过借助计算机进行恶意代码检测，在很大程度上提高了恶意代码的检测成功率。通过这份软件产品需求分析报告详尽说明了该软件产品的需求规格，包括对产品进行的准确的定义。

## 开发目的及意义

本设计为一个恶意代码检测系统，即通过对恶意代码图像化处理并运用卷积神经网络（CNN）来对恶意代码进行检测。本系统能够对恶意代码进行检测，有助于网络环境的维护以及系统开发的支持。

## 预期读者和阅读建议

本软件产品需求分析报告所针对的各种不同的预期读者，可能包括：

● 用户；

● 开发人员；  
● 测试人员  
● 学生  
● 老师  
● 文档编写入员

# 系统概述

## 运行环境

本软件的运行环境，包括：

● 硬件平台：SUN V880或以上，双CPU，2G内存，200G以上磁盘阵列

● 操作系统：linux系统  
● 使用框架：viper，tensorflow，pytorch  
● 运行环境：python 3

## 设计和执行约束

在开发过程中若出现一些由于开发者技术不足导致无法完成的任务，我们将对这些功能做出简化或必要的删改。

# 功能需求

## 系统功能列表

|  |  |
| --- | --- |
| 功能点 | 功能描述 |
| 导入数据 | 引入viper框架中存储的二进制数据 |
| 数据处理 | 对引入的二进制数据进行降噪处理 |
| 数据图像化 | 讲二进制数据以8个字节为单位，转换成灰度图，使其成为卷积神经网络识别的数据集 |
| 导入训练集 | 导入需要的训练集 |
| 导入测试集 | 导入需要的测试集 |
| CNN识别 | 通过建立的卷积神经网络模型对恶意代码进行识别 |
| 提供接口访问 | 提供Restful API访问接口 |
| 汉化处理 | 对viper中重要的功能部分进行汉化处理，方便更好的使用 |
| 模块添加 | 对训练数据集的模型参数进行替换 |

## 功能需求

● 实现基于CNN和图像的恶意代码检测模块（基于已有模块）

● 将检测模块集成到开放恶意代码管理框架中

● 提供Restful API访问接口

● 使用高斯数据库对训练使用的数据集进行统一管理，支持在viper-web中下载和上传数据集。

● 在CNN识别模块可以选择多个不同模型进行识别

# 非功能需求

● 准确率大于75%

● 使用Python3

● 对训练模块的训练结果的参数进行统一管理

# 5. 性能需求

● 程序运行无明显延迟

● 10000条恶意代码检测时间应小于5分钟

# 6.用户需求

●数据库使用

将模型训练的数据集以及后续单独测试的数据进行分开，将训练数据集依然存放在高斯数据库中，将用户进行测试的数据集存入viper自带的数据库中。

●模块使用

暂时先取消自己编写模块，对已有模块进行重写然后集成到viper中去，主要针对数据处理模块以及训练模块。

●数据集的替换以及增添

基于老师已给的数据集进行进一步的增添，可以识别更多的恶意代码家族类型。

●汉化处理

对viper中重要的功能部分进行汉化处理，方便更好的使用。

●模块添加

可以对训练数据集的模型参数进行替换。

●训练参数更改

用户可以更改训练模块的训练参数。用户将训练参数写成配置文件(.json)，模块可以通过调用该文件中的参数进行训练。

●模型参数更改

用户可以在viper-web界面上直接将自己训练的参数应用到识别模块。