정보보호 R & D 데이터 챌린지 2018

AI기반 안드로이드 악성앱 탐지

너의 3대는
김현석 외 4명

CONTENTS

1. Introduction
2. Suggestion
3. Experiment
4. Result

/01

Introduction

Introduction

안드로이드 모바일 사용자 정보를 훔쳐가는 악성앱 발견

좋아요 10개 | 입력 : 2018-08-20 14:27

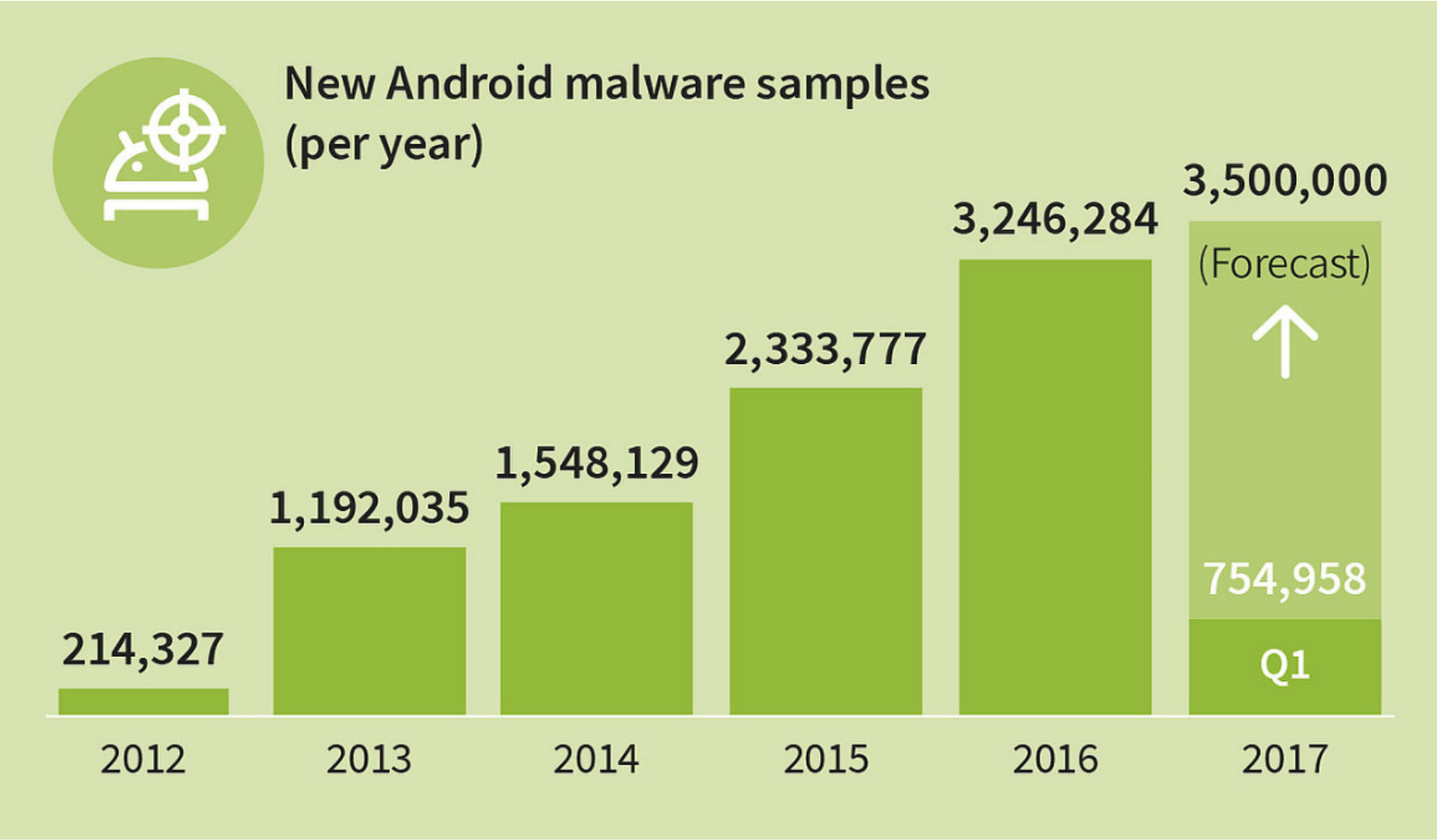
#안드로이드 #악성앱 #정보보호 #정보보안 #사이버보안 #IT보안 #이스트시큐리티

안드로이드 모바일 사용자 정보 훔쳐가는 악성앱 'Misc.Android.InfoStealer' 발견
설정앱 설치 시 기기 정보, 사용자가 현재 어떤 앱 실행하는지 등 실시간으로 정보 전송

[보안뉴스 김경애 기자] 안드로이드 모바일 사용자 정보를 훔쳐가는 악성앱 'Misc.Android.InfoStealer'이 발견됐다. 이 악성앱은 게임의 이용을 돕는 게임패드와 연결되는 설정 앱으로 과도한 권한을 요구하며 권한을 이용해 사용자 몰래 기기 정보를 수집하고 전송한다.

출처 : 보안뉴스, www.boannews.com [Internet}, Available : https://www.boannews.com/media/view.asp?idx=72315, 2018.10.11

Introduction



출처 : G data software, www.gdatasoftware.com [Internet], Available : https://www.gdatasoftware.com/news/2017/04/29715-350-new-android-malware-apps-every-hour, 2018.10.17

Introduction

/02

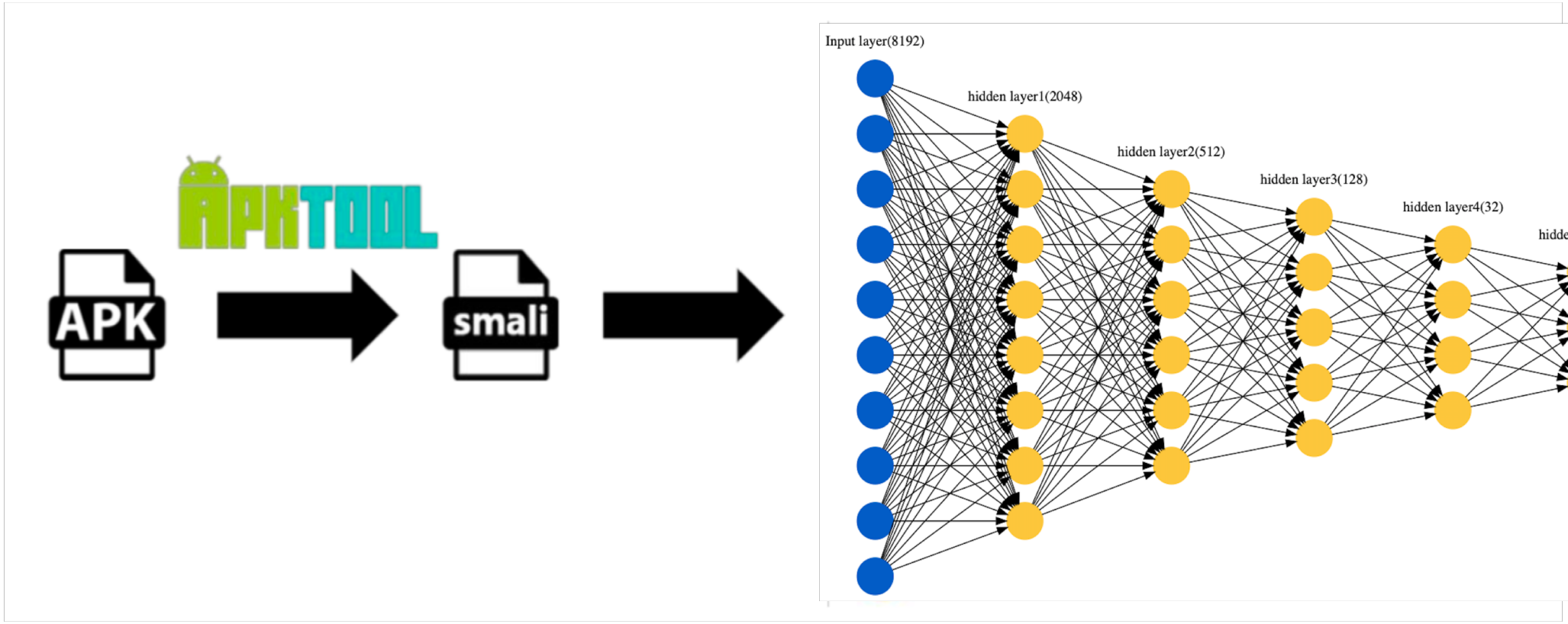
Suggestion

- Preprocess
- Model

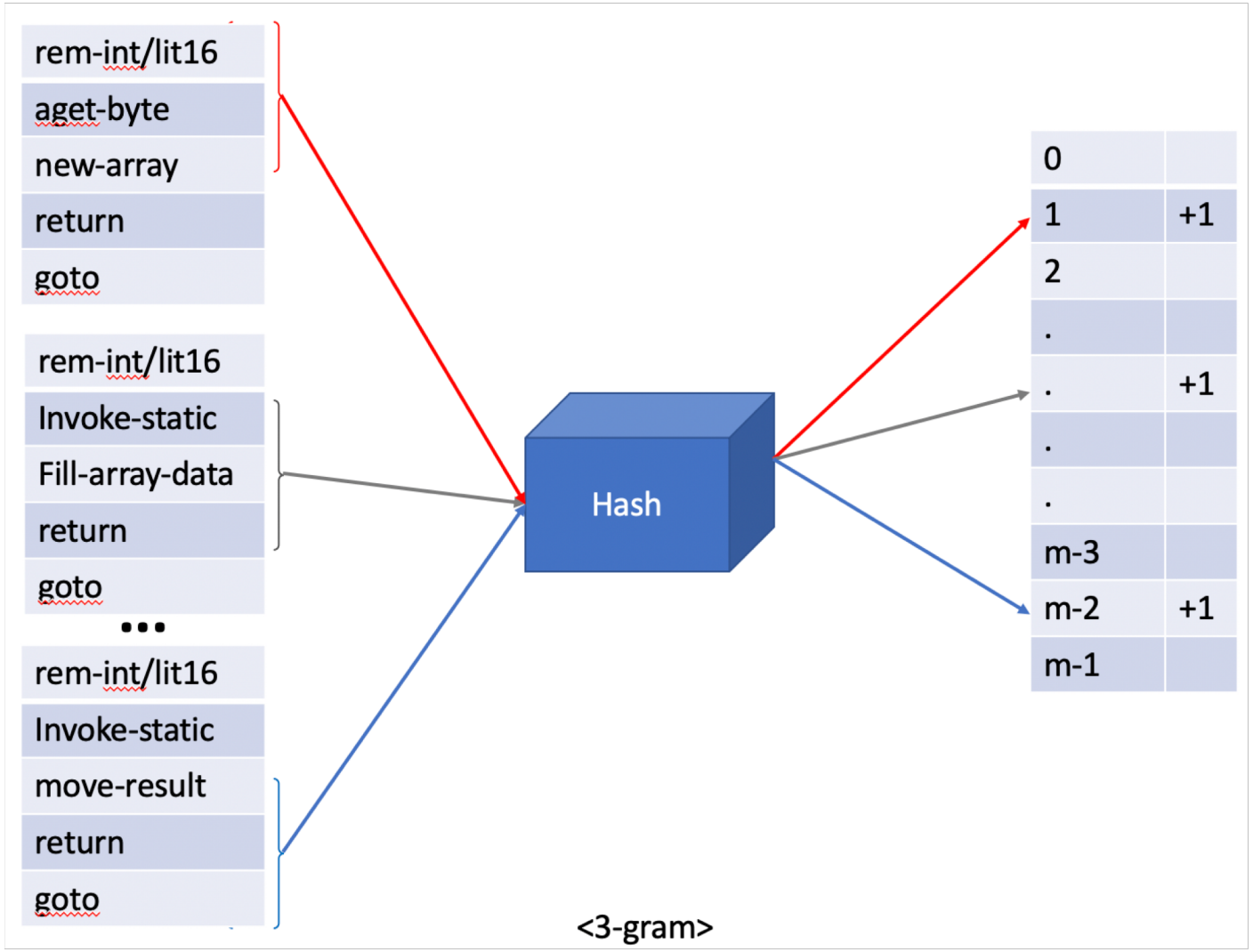
Preprocess

- Bytecode
- Method
- Multiple hash

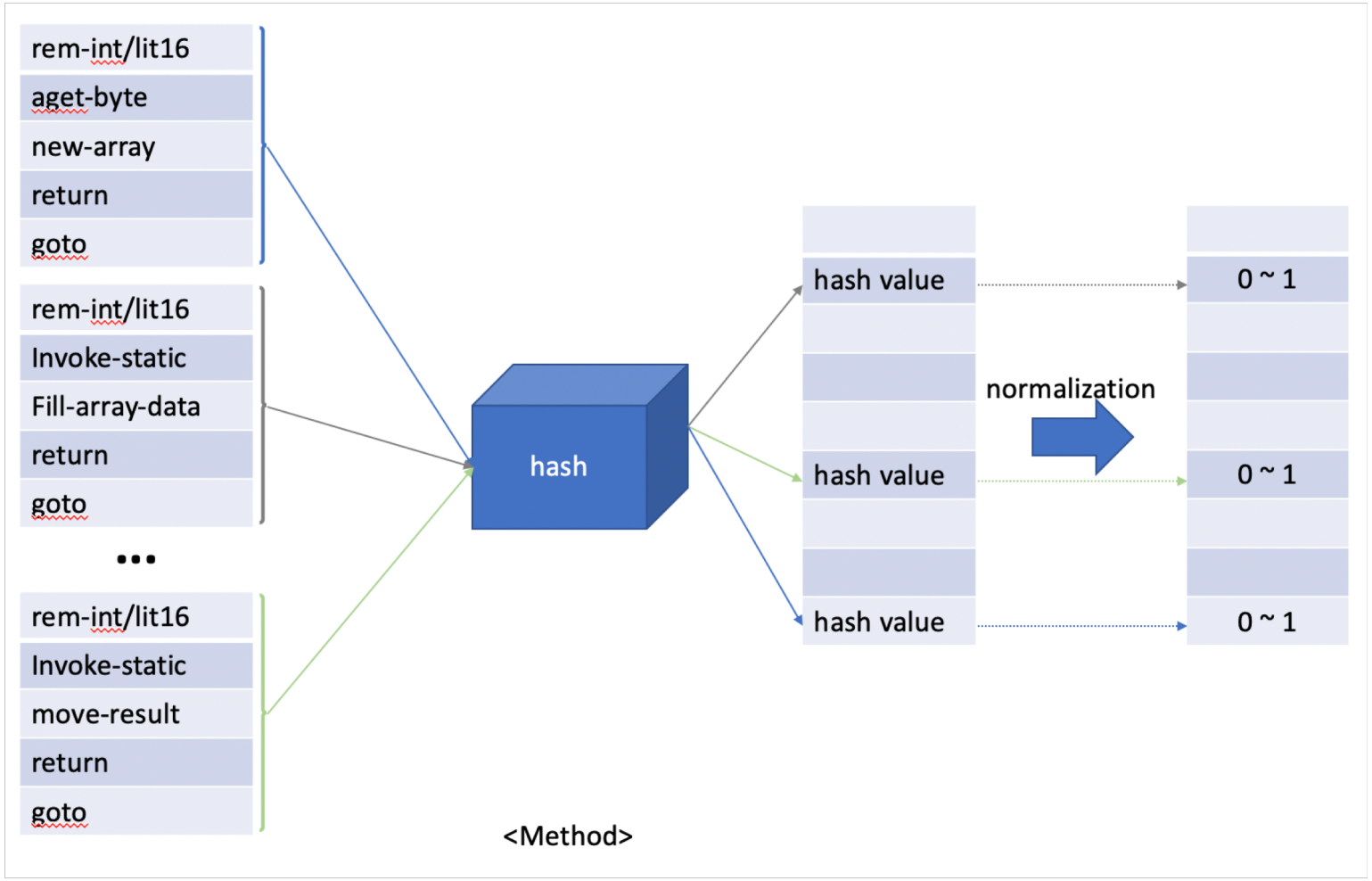
Preprocess


APKTOOL
APK
smali
Input layer(8192)
hidden layer1(2048)
hidden layer2(512)
hidden layer3(128)
hidden layer4(32)
hidden layer5(8)
output layer

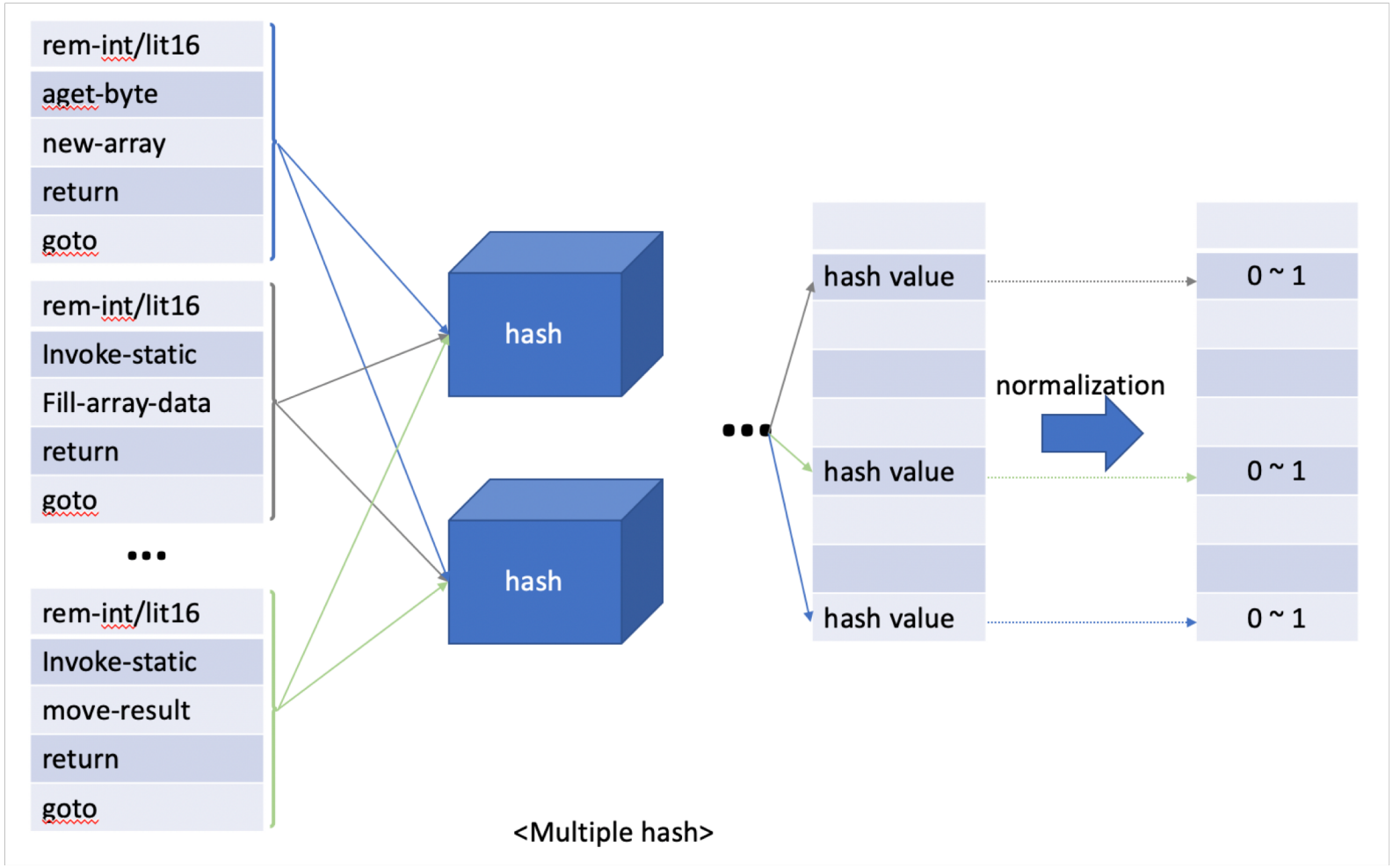
Bytecode


rem-int/lit16
aget-byte
new-array
return
goto
rem-int/lit16
Invoke-static
Fill-array-data
return
goto
...
rem-int/lit16
Invoke-static
move-result
return
goto
Hash
0
1 +1
2
.
. +1
.
.
m-3
m-2 +1
m-1
<3-gram>

Method


rem-int/lit16
aget-byte
new-array
return
goto
rem-int/lit16
Invoke-static
Fill-array-data
return
goto
...
rem-int/lit16
Invoke-static
move-result
return
goto
hash
hash value
hash value
hash value
normalization
0 ~ 1
0 ~ 1
0 ~ 1
<Method>

Multiple Hash



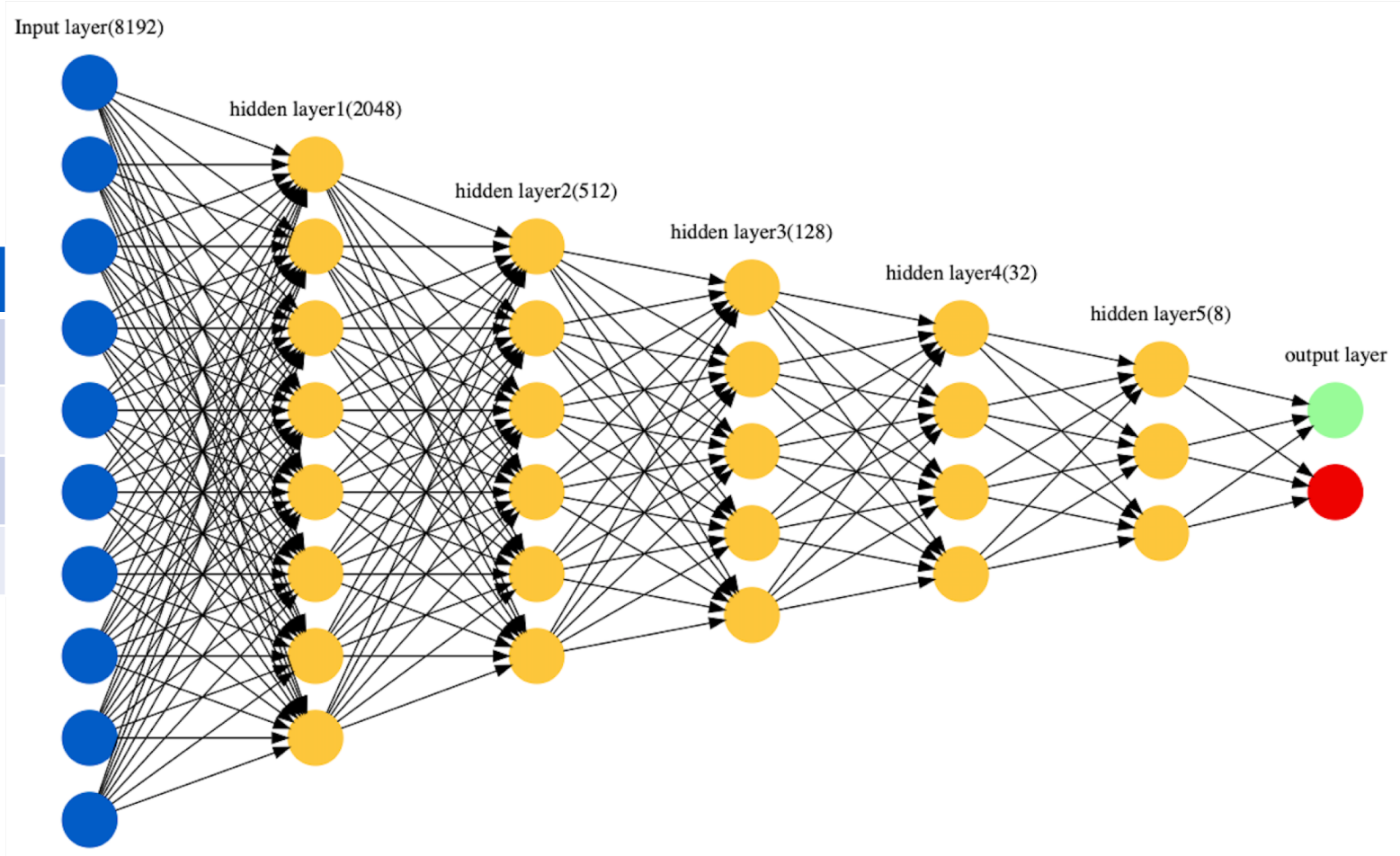
<Multiple hash>

Model

- ANN
- CNN

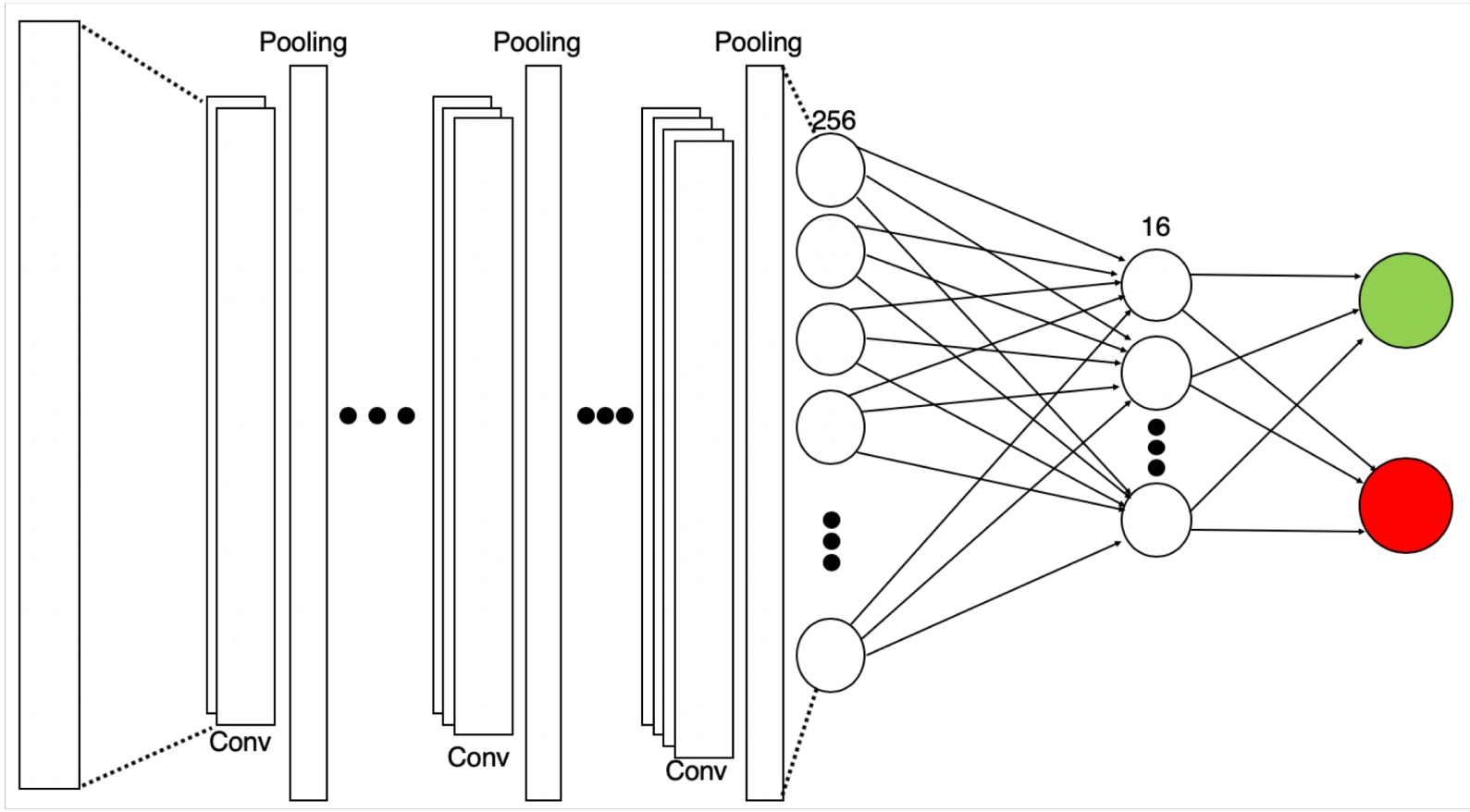
Model_ANN

Hyperparameter	
Activation function	ReLU
Loss function	Cross-entropy
Optimizer	Adam
Regularizer	Batch norm



Model_CNN

Hyperparameter	
Activation function	ReLU
Loss function	Cross-entropy
Pooling	Max pooling
Optimizer	Adam
Regularizer	Batch normalization

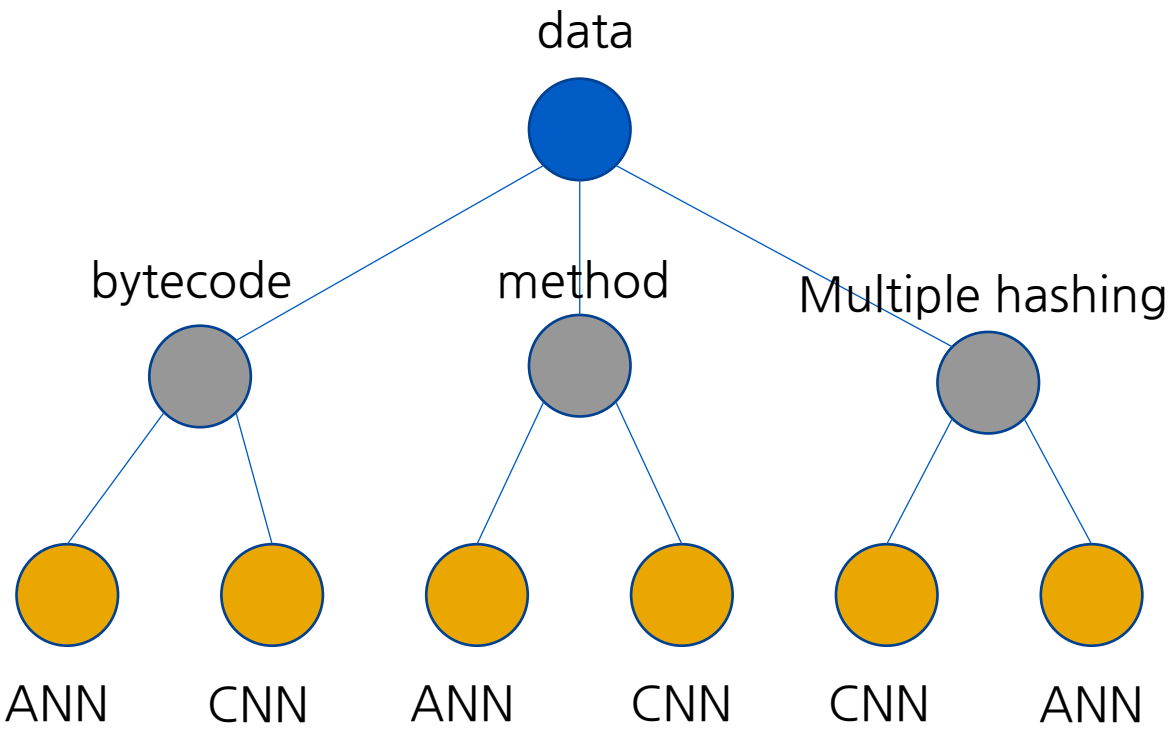


/03

Experiment

Experiment

Environment	
GPU	GTX1080 Ti 11G
CPU	Intel(R) Core(TM)i7-8700K
RAM	DDR4 16G
Language	Python 3.6
Framework	Tensorflow



Experiment

	Malware	Benign
Train	대략 35,000	대략 110,000
Test	2,000	4,000

	ANN	CNN
3-gram/Feature Hashing	98.00	97.93
method/Feature Hashing	98.26	97.95
method/Multiple Hashing	98.67	97.51

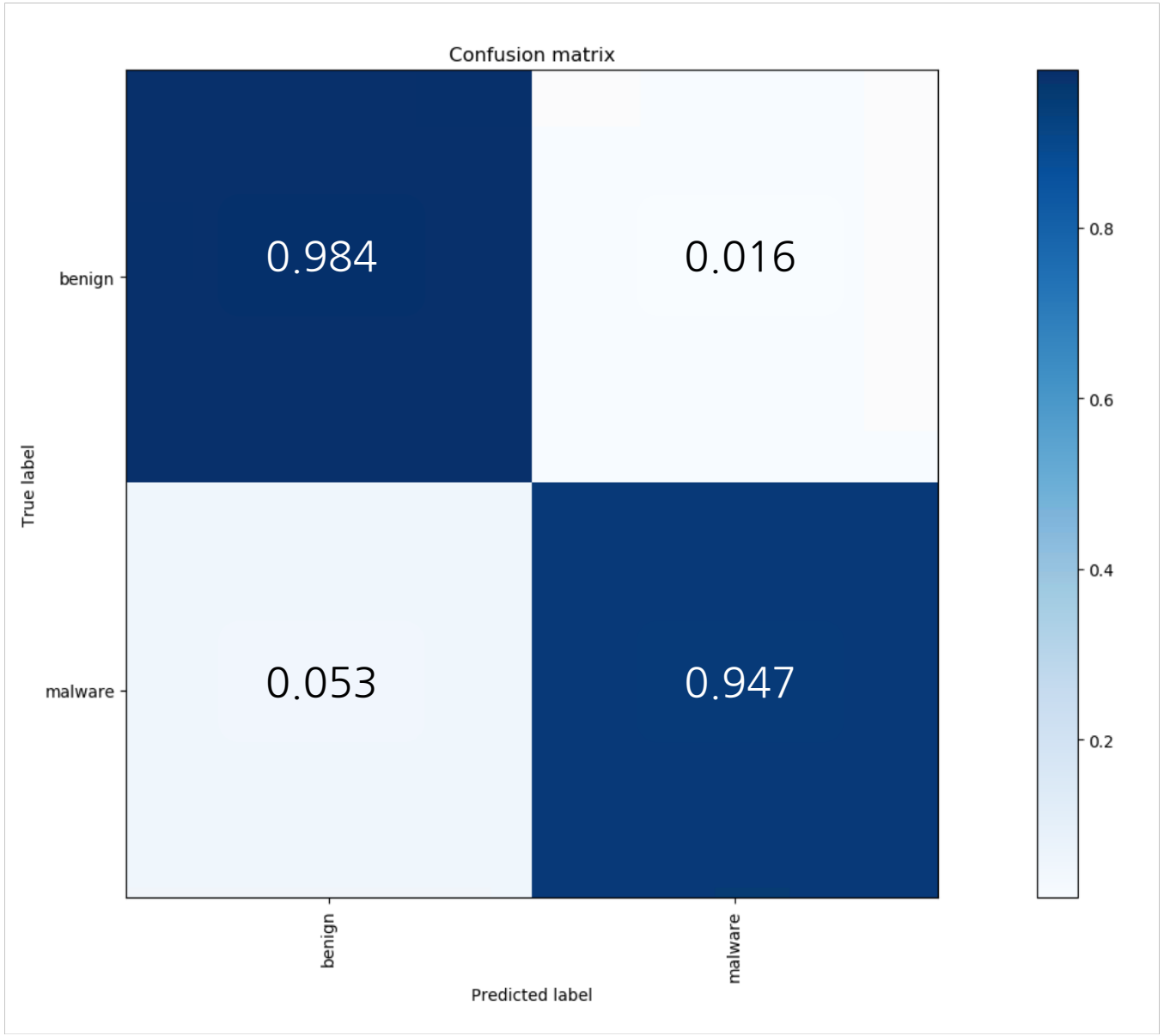
/04

Result

Result of ANN

1st ACC : 97.05

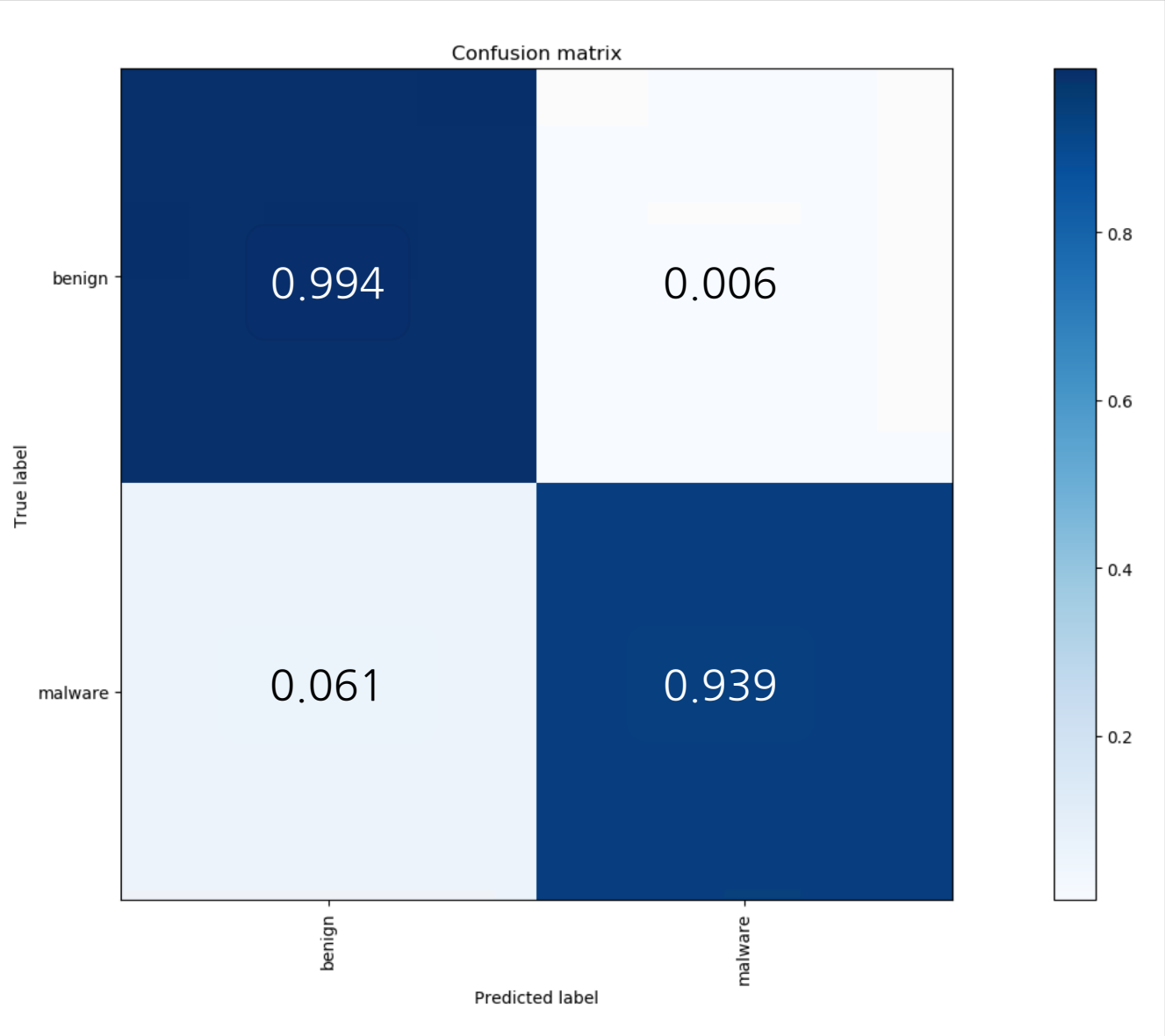
2st ACC : 96.7



Result of CNN

1st ACC : 97.4

2st ACC : 97.65



Thanks
And Your Slogan Here

Speaker name and title

www.islide.cc