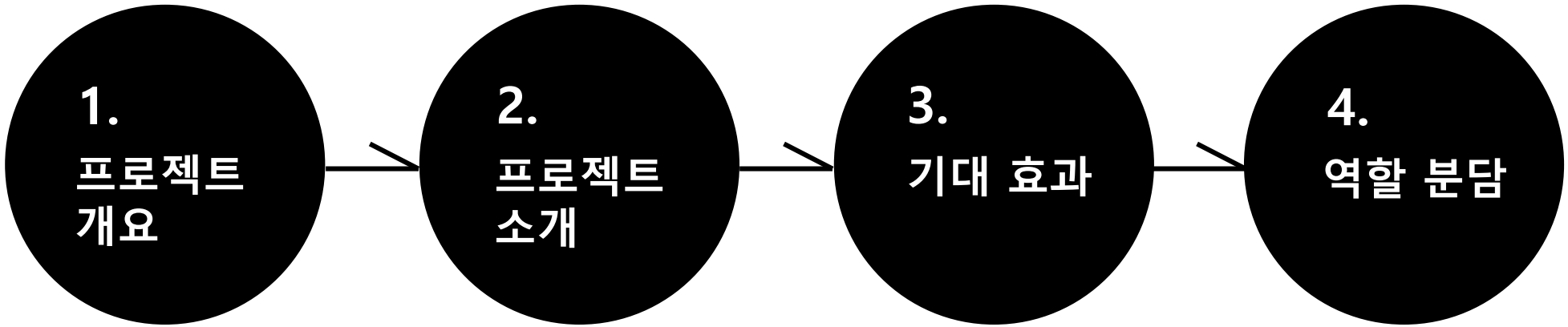
DREAM

Do it !

문다민 김기환 김현석 정혜리 방유한

목차



1. 프로젝트 개요

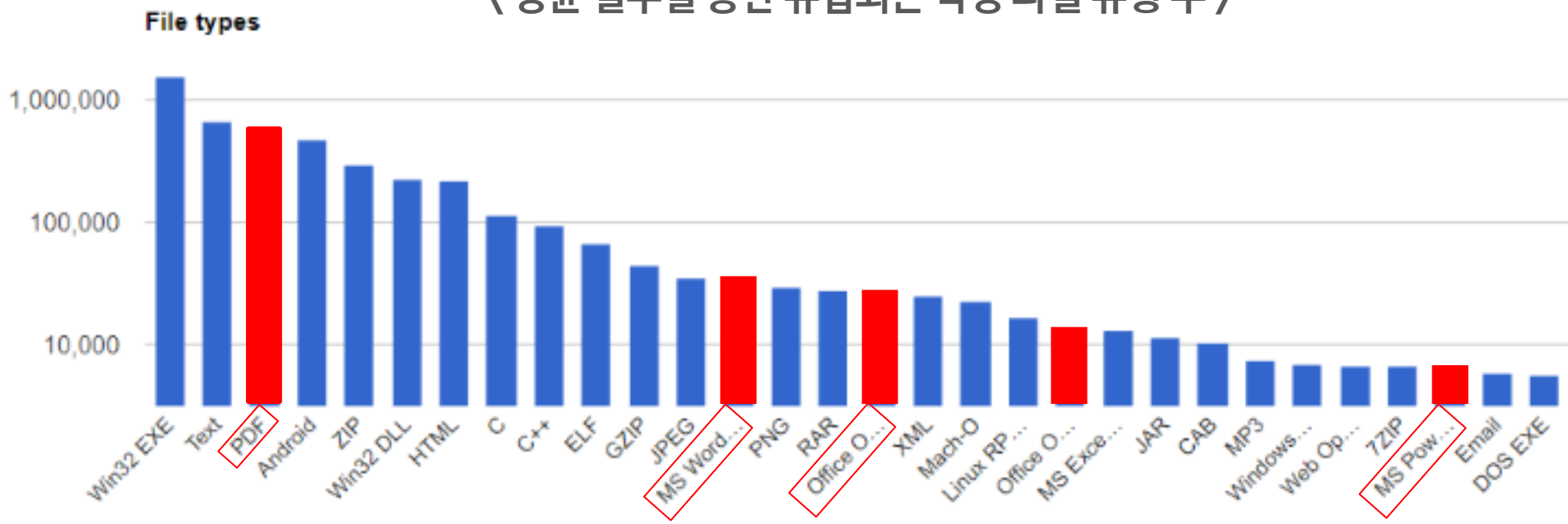
지능적으로 발전하는 악성코드, 이를 이용한 사이버 공격

출처 :
https://www.ahnlab.com/kr/site/main.do

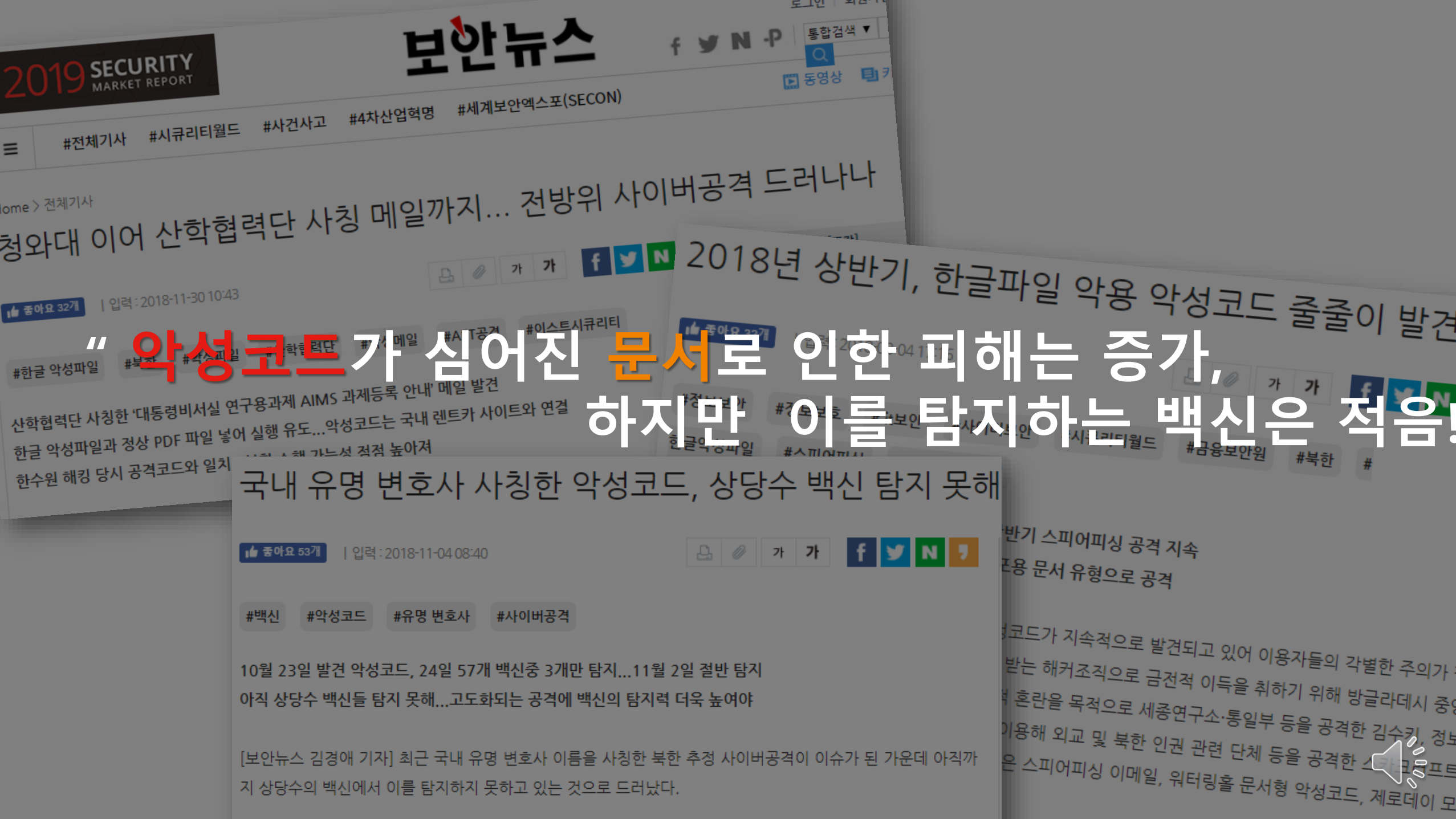
1. 프로젝트 개요

그 중의 상당 수는 우리가 자주 접하는 PDF, MS Office 등 **문서**에 심어진 악성코드

〈 평균 일주일 동안 유입되는 악성 파일 유형 수 〉



출처 : https://www.virustotal.com/ko/statistics/



" 악성코드가 심어진 문서로 인한 피해는 증가,
하지만 이를 탐지하는 백신은 적음!

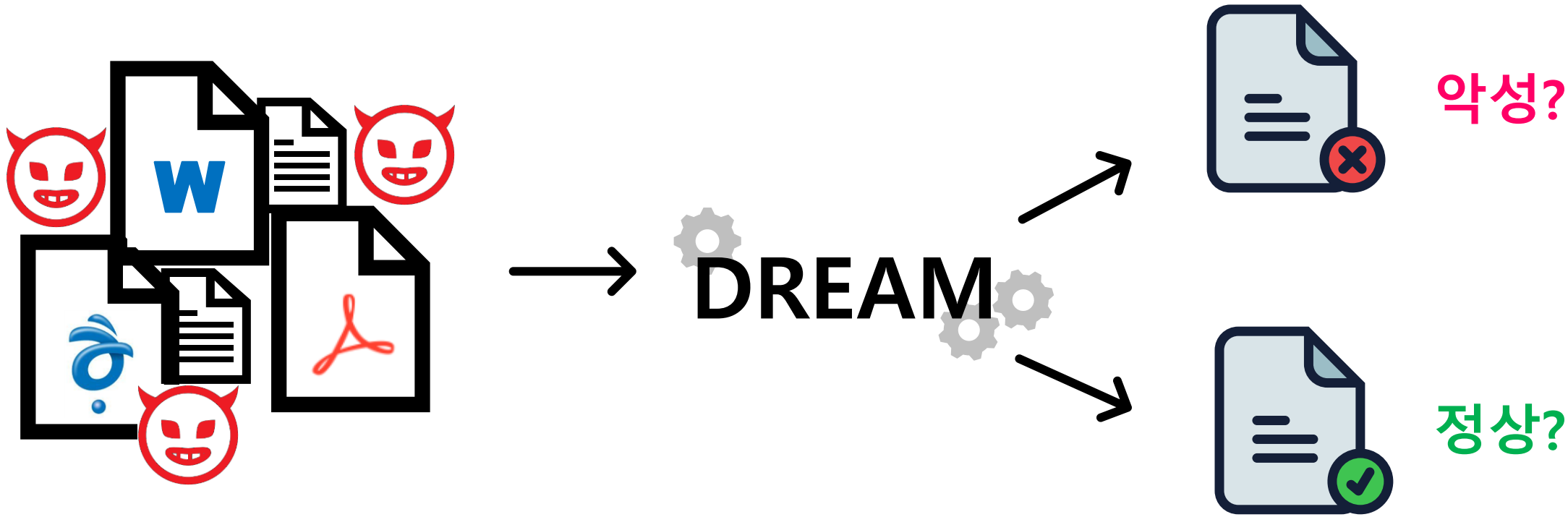
2. 프로젝트 소개

프로젝트 이름

Detecting
in **R**eal – tim**E**
m**A**licious documents
using **M**achine learning

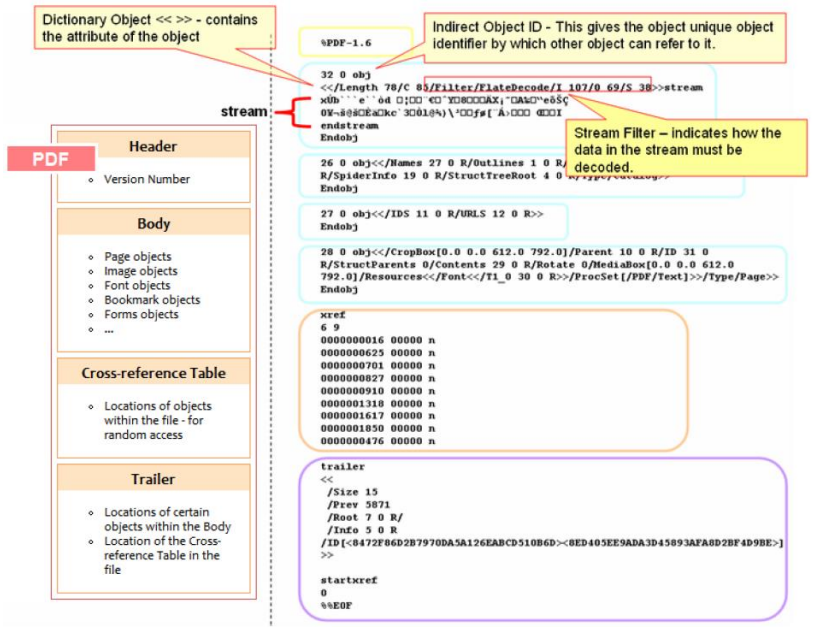
2. 프로젝트 소개

프로젝트 목표

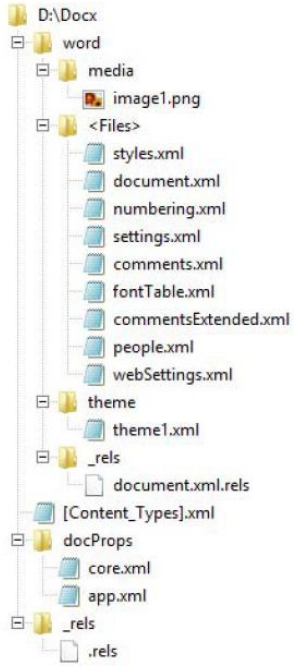


2. 프로젝트 소개

개발 내용



출처 : PDF 구조 및 악성코드 분석. (2013). 누리랩, p.6.



Storage Stream

설명	구별 이름	길이(바이트)	레코드 구조	압축/암호화
파일 인식 정보	FileHeader	고정		
문서 정보	DocInfo	고정	√	√
본문	BodyText Section0 Section1 ...	가변	√	√
문서 요약	\005HwpSummaryInformation	고정		
바이너리 데이터	BinData BinaryData0 BinaryData1 ...	가변		√
미리보기 텍스트	PrvText	고정		
미리보기 이미지	PrvImage	가변		
문서 옵션	DocOptions _LinkDoc DrmLicense ...	가변		
스크립트	Scripts DefaultJScript JScriptVersion ...	가변		
XML 템플릿	XMLTemplate Schema Instance ...	가변		
문서 이력 관리	DocHistory VersionLog0 VersionLog1 ...	가변	√	√

출처 : 한글문서파일구조. (2013). 한글과 컴퓨터

2. 프로젝트 소개

개발 내용

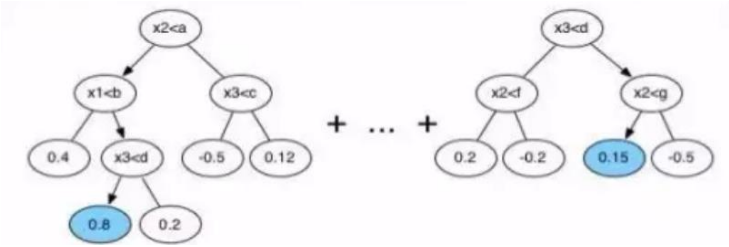

```
obj 1 0
 Type: /Catalog
 Referencing: 2 0 R, 17 0 R, 1151 0 R, 1152 0 R

  <<
    /Type /Catalog
    /Pages 2 0 R
    /Lang (ko-KR)
    /StructTreeRoot 17 0 R
    /MarkInfo
      <<
        /Marked true
      >>
    /Metadata 1151 0 R
    /ViewerPreferences 1152 0 R
  >>


obj 2 0
 Type: /Pages
 Referencing: 3 0 R, 14 0 R

  <<
    /Type /Pages
    /Count 2
    /Kids [ 3 0 R 14 0 R]
  >>
```

F_1	F_2	F_3	F_4	...	F_n
1.7	12	9	124	...	0

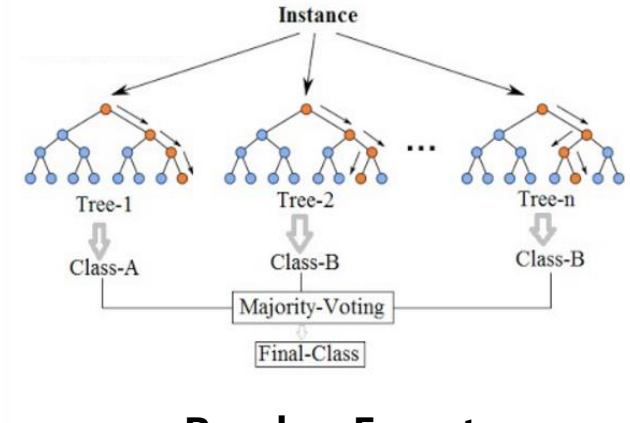
Feature Vector

2. 프로젝트 소개

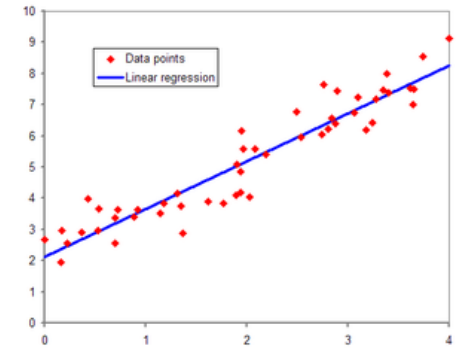
개발 내용



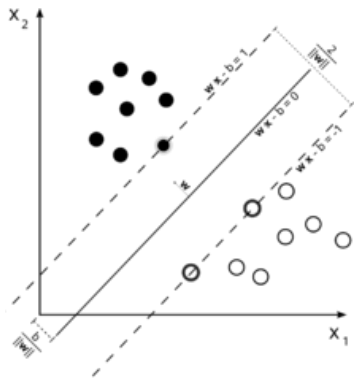
GBDT



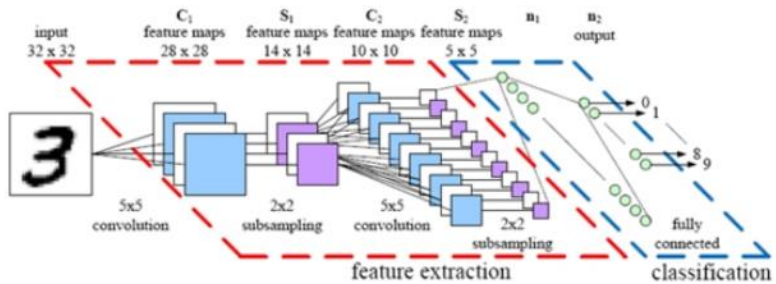
RandomForest



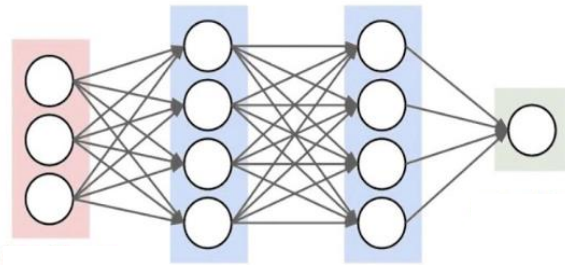
Linear Regression



SVM



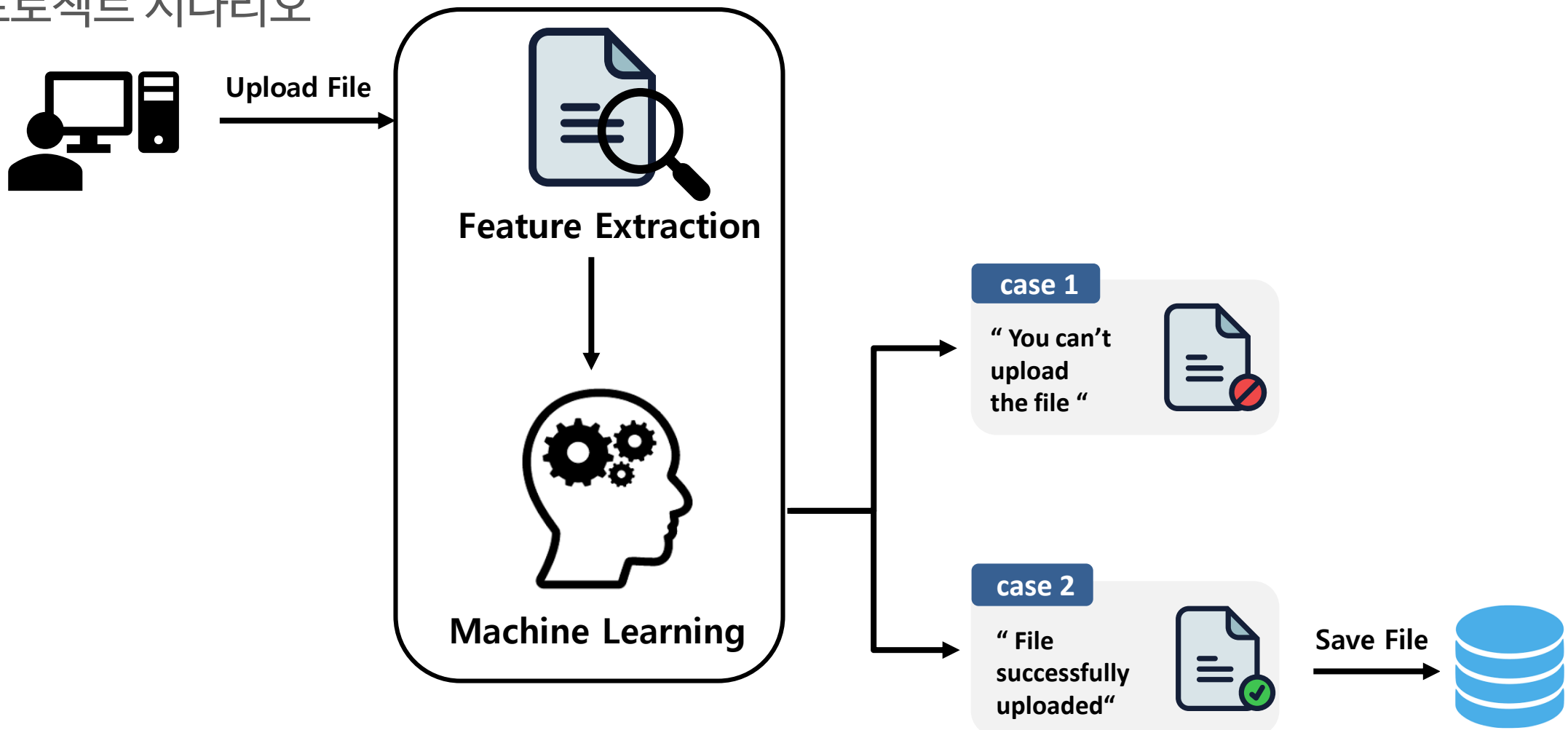
CNN



DNN

2. 프로젝트 소개

프로젝트 시나리오



3. 기대효과

1. 문서형 악성코드 유포 방지

문서형 악성코드 유포 방지로 사회 문제 해소

2. 오픈소스 소프트웨어

다른 시스템 환경에 맞게 코드 수정 및 사용 가능

4. 역할분담

문다민 (팀장)

머신 러닝 모델 설계 및 구축

김현석

데이터 라벨링, 웹 서버 구축

김기환

특징 추출 및 가공, 데이터 수집

정혜리

특징 추출 및 가공, 문서 작업

방유한

유저 인터페이스 구현

감사합니다.

문다민 김기환 김현석 정혜리 방유한

Reference

- GBDT 사진 : https://ask.hellobi.com/blog/mlanddlanddm/7103
- RandomForet 사진 : https://yeo0.github.io/data/2018/11/06/6.-%EB%9E%9C%EB%8D%A4-%ED%8F%AC%EB%A0%88%EC%8A%A4%ED%8A%B8/
- 선형회귀 사진 : https://ko.wikipedia.org/wiki/%EC%84%A0%ED%98%95_%ED%9A%8C%EA%B7%80
- SVM 사진 : https://ko.wikipedia.org/wiki/%EC%84%9C%ED%8F%AC%ED%8A%B8_%EB%B2%A1%ED%84%B0_%EB%A8%B8%EC%8B%A0
- CNN 사진: http://physics2.mju.ac.kr/juhapruwp/?p=1517
- DNN 사진: https://www.slideshare.net/JinwonLee9/ss-70446412

Reference

- 기사 1 – https://www.boannews.com/media/view.asp?idx=75093
- 기사 2 – https://www.boannews.com/media/view.asp?idx=71995&kind=1&search=title&find=%B9%AE%BC%AD%C7%FC
- 기사 3 – https://www.boannews.com/media/view.asp?idx=74302