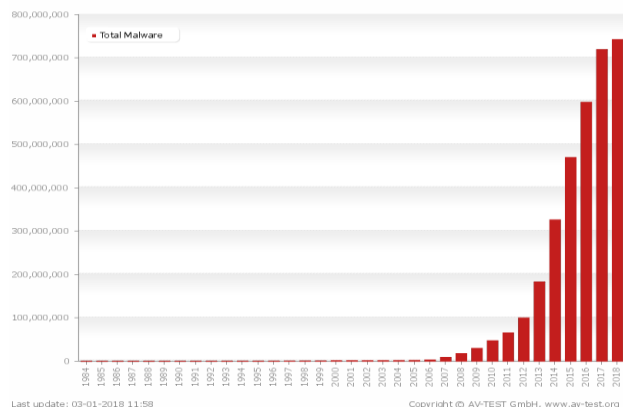


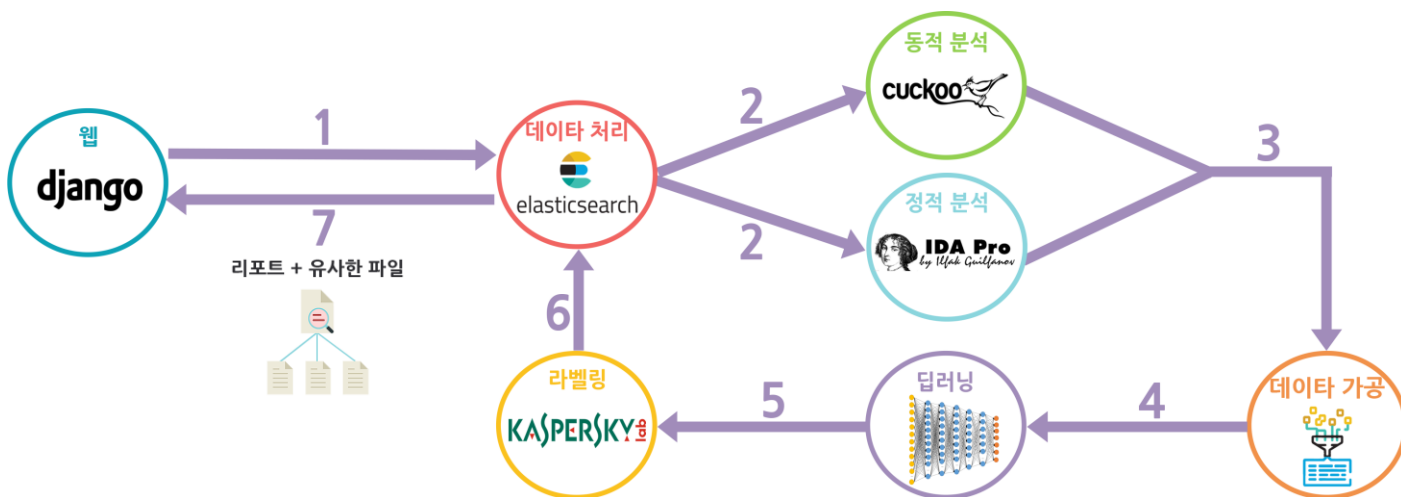
## 1. 프로젝트 소개



보안제품 인증 기관 AV-TEST에 통계에 따르면, 해마다 총 악성코드의 개수가 늘어나고 있습니다. 하지만 **악성코드 전문가의 수는 한정적**입니다. 때문에 **악성코드 전문가**에게 효율적으로 악성코드를 분석하는 **자동화된 분석 시스템**이 필요합니다. 우리는 4차 산업 혁명이 대두되면서 각광받고 있는 인공지능과 빅데이터 기술을 적용하여 이 문제를 해결하고 악성코드 전문가를 돕고자 합니다.

MASK(Malware Analysis System in Kookmin)는 파일을 동적, 정적 분석 기술을 사용하여 분석하고 결과를 보여주는 오픈소스 소프트웨어입니다. 때문에 악성코드 분석에 대한 전문 지식이 있는 **악성코드 분석 전문가**를 사용자로 정의합니다. 우리는 IDA를 이용하여 정적 정보를, Cuckoo Sandbox를 이용하여 동적 정보를 추출한 뒤 Tensorflow 이용하여 분류 모델을 학습하고 분석 결과를 보여줍니다. 추가로 우리 데이터베이스에 있는 데이터의 검색을 위해 Elasticsearch를 도입하였습니다.

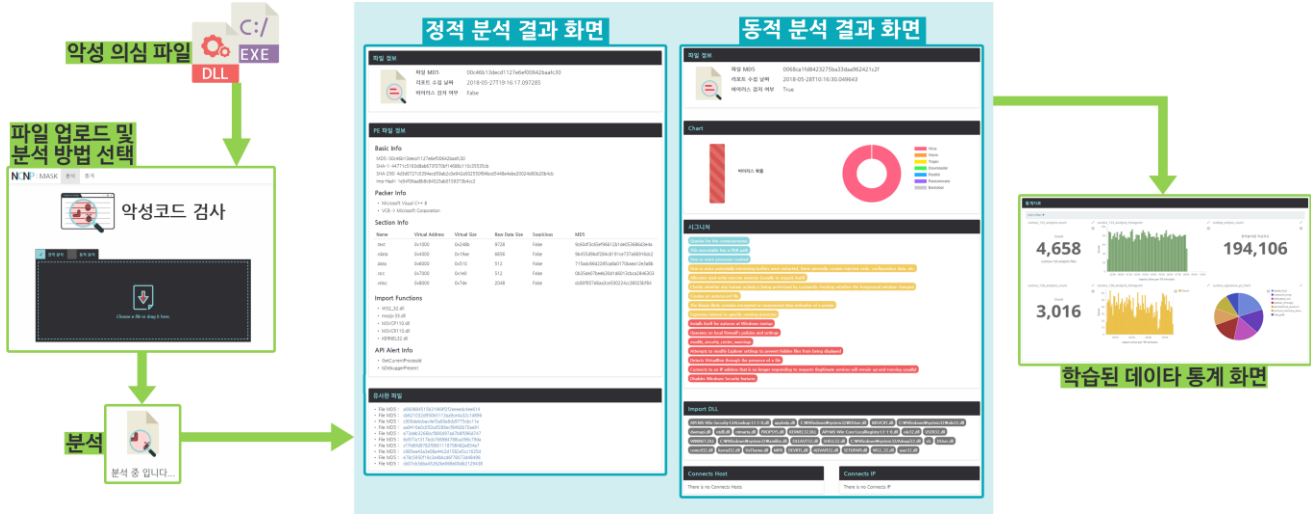
## 2. 시스템 구조도



"동적 분석과 심층 신경망 모델을 이용한 악성코드 분류",

한채연 · 김영재 · 허준녕 · 명준우, KCC 2018(한국정보과학회) 논문 발표 예정

### 3. 시나리오



- I. 사용자(악성코드 전문가)가 MASK 웹에 악성으로 의심되는 파일을 업로드한다.
- II. 사용자는 파일 업로드시 정적/동적 분석을 선택할 수 있다.
- III. 해당 파일에 대한 분석결과가 데이터베이스에 이미 있는 경우 해당 분석결과를 웹을 통하여 사용자에게 보여준다.
- IV. 데이터베이스에 분석결과가 없는 파일이 업로드 되면 정적 분석의 경우 IDA Pro, 동적 분석의 경우 Cuckoo Sandbox를 통해 분석하여 분석 결과를 생성해낸다.
- V. 정적/동적 분석 단계에서 생성된 분석 결과로부터 추출된 특징을 이용하여 학습된 딥러닝(Tensorflow) 모델로부터 탐지 결과를 구한다.
- VI. 4, 5 과정을 거쳐 생성된 분석 및 탐지 결과는 데이터베이스에 저장하고 웹을 통하여 사용자에게 보여준다.

### 4. MASK의 장점



#### I. 오픈소스 소프트웨어로 공개

현재 대다수의 악성코드 분석 서비스는 일정 범위에 한해서만 무료로 제공되며 더 많은 서비스에 대해서는 유료로 서비스가 제공된다. 본 프로젝트를 오픈소스 소프트웨어로 제공함으로써 개인 및 기업에서 더 쉽게 사용할 수 있도록 하여 악성코드 분석 서비스를 더 쉽게 사용할 수 있도록 한다.



#### II. 딥러닝 기반 악성코드 분류

기존 안티바이러스 소프트웨어들은 시그니처 기반 혹은 머신러닝을 이용하여 악성코드를 탐지하였다. 이러한 방법은 기하급수적으로 증가하는 신종/변종 악성코드를 탐지하는데 한계가 있다. 그러나 본 프로젝트에서는 딥러닝을 이용하여 기존 방법에 비해 신종/변종 악성코드를 더욱 정교하게 탐지해낸다.



#### III. 정적/동적 분석 동시 활용

기존 악성코드 탐지에는 정적/동적 분석 중 하나의 방법만을 선택하여 이용해왔다. 그러나 정적/동적 분석은 각각의 한계점을 지니고 있다. 본 프로젝트에서는 정적/동적 분석을 동시에 이용하여 각각의 한계점을 극복해낸다.



#### IV. 분석 파일과 유사한 파일 정보 제공

본 프로젝트는 업로드 된 파일에 대해 분석 후 ssdeep이라는 파일 유사도 측정 툴을 사용하여 데이터베이스에 있는 유사한 파일의 기존 분석 결과를 동시에 보여준다. 유사한 파일에 대한 정보를 제공함으로써 업로드 된 파일 분석에 단서를 제공한다.