**AI를 활용한 Spam mail Detection**

프로젝트 수행 보고서

작성일 : 2021.09.04

팀명 : AI를 활용한 SPAM MAIL DETECTION

문서 수정 내역(Revision Chart)

|  |  |  |
| --- | --- | --- |
| 문서 버전 | 완료일 | 주요 수정 내용 |
| v.1 | 09.04 | 프로젝트 기본 개요 작성: 권채연 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

목차

**1. 프로젝트 개요 3**

1.1 프로젝트 개요 3

1.2 추진 배경 및 필요성 4

1.3 서비스 내용 4

1.4 프로젝트 수행 역할 4

1.5 기대효과 5

**2. 기존 스팸 업무 문제점 5**

2.1 기존 스팸 업무 문제 및 현황 5

2.2 기존 업무 구성도 6

2.3 개선 방향 6

**3. AI 스팸 분석 시스템 설계 6**

3.1 개발 환경 6

3.2 시스템 구성 6

1. 프로젝트 개요
   1. 프로젝트 개요

|  |  |
| --- | --- |
| 프로젝트 개요 | |
| 프로젝트명 | AI를 활용한 Spam mail Detection |
| 프로젝트 목표 | 스팸 메일 분석 자동화 프로세스를 구축하고 분석 시간을 단축하여 빠른 대응 및 업무 효율 증진 |
| 타겟 | - 스팸 신고 접수를 받는 스팸부서, 관제부서  - 스팸 솔루션이 없는 중소기업 |

* 1. 추진 배경 및 필요성
* 스팸 신고를 받아 악성 여부를 판단하는 스팸 부서 또는 관제 부서의 업무 과부하 발생
* 분석 업무의 과도한 시간 소모로 인한 타 업무 효율 저하
* 유입 되어진 악성 메일에 대한 빠른 분석 및 인지와 대응 필요
  1. 서비스 내용
* DNS Server Mail server 구축
* 메일 스크래핑을 통한 메일 자동 수집.
* 스팸 메일 분류 서비스에 머신러닝을 접목하여 스팸 필터링
* 분석 report 바탕 머신러닝 스팸 여부 판단
* 스팸 메일 분류 후 시각화(ELK)

**1.4** 프로젝트 수행 역할

|  |  |
| --- | --- |
| 수행 역할 | |
| PM | 내용 : 프로젝트 전체 총괄 및 인력 관리  수행자 : 정병현 |
| PL ,기획, 설계 | 전체 |
| 코어 개발 | 내용 : 머신러닝 학습 시스템 개발, 로그 파싱  수행자 : 정병현 |
| 서브 개발 | 내용 : 메일 스크래핑, 데이터 가공, 머신러닝 보조  수행자 : 권채연 |
| 시스템 구축 | 내용 : Mail server, DNS server, ELK 구축  수행자 : 연람희 |
| QA(테스터) | 내용 : 전체 시스템 테스트 및 버그, 오류 탐지  수행자 : 전체 |

**1.5** 기대효과

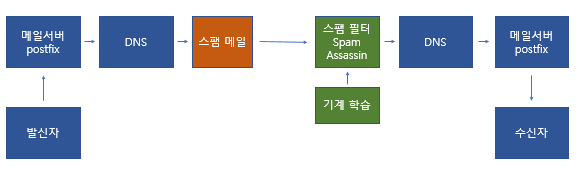
* 신고된 스팸메일에 대한 자동 분석으로 업무 효율 상승
* 머신러닝 분석의 효율성 지표화(분석 시간 단축, 분석 메일 수 증가 등) 및 KPI 반영
* 기존 스팸 차단 시스템의 개선점 도출

1. 기존 스팸 업무 문제점
   1. 기존 스팸 업무 문제 및 현황

* 일반 메일(광고성, 단순 텍스트) 신고 다수
* 타 업무와 동시에 이루어져 분석 소요시간 과다 발생
* 스팸메일 신고 현황

|  |  |  |
| --- | --- | --- |
| 스팸신고 건수 | 실제 악성메일 | 분석 시간 |
| 평균 200건 이상 | 평균 3~4건 이상 | 2~3시간 소요 |

* 1. 기존 스팸메일 유입 및 분석 구성도

****

* 1. 개선 방향
* 신고 메일 수집 자동화
* 분석 자동화

1. AI 스팸 분석 시스템 설계
   1. 개발 환경

|  |  |
| --- | --- |
| **개발 환경** | |
| **메일 스크래핑, 머신러닝** | 개발 언어 : Python 3.x  OS : centos 7 |
| **대시보드** | 분석 툴: ELK  OS : centos 7 |

* 1. 시스템 구성

|  |
| --- |
| **[머신러닝 학습]** |
| **1) 데이터에 대한 전처리**  Kaggle 에서 spam mail data를 다운로드 후 중복과 불 필요한 열 제거 훈련 데이터와 테스트 데이터의 분리   1. Labeling spam일 경우 0으로 spam 이 아닐 경우 1로 표시 2. Spam 유무를 의미하는 레이블 값을 분포를 시각화 3. 단어의 개수를 학습 파라미터로 정하여 인코딩 4. 메일 데이터 길이 분포 시각화 5. 훈련 데이터와 테스트 데이터 분리   **2) CNN으로 스팸 메일 분류하기**  캐글에서 제공하는 스팸메일 데이터를 활용하여 스팸메일 분류  참고: 딥러닝을 이용한 자연어 처리 입문 |
| **[시스템 구성도]** |
|  |

1. 프로젝트 관리 기본 방향
   1. 의사소통 방식

* PM 주도 아래 1주 단위로 목표 계획을 수립하고 개인 역량에 맞게 범위 조절
* 진행 사항에 대해 매일 10분씩 간략 보고
* 이슈 사항에 대해 48시간 이상 해결 못할 시 미결 문제로 오픈
* 개인사정으로 인한 지연 발생 시 즉시 보고
  1. 협업 도구

|  |  |
| --- | --- |
| 협업 툴 | 내용 |
| Git hub | 개발 코드 공유 |
| ZIRA | 일정 관리, 프로젝트 진행관리 |
| Slack | 정보공유 |

1. 테스트 계획
2. 일정 계획

~ 8.31 수행 계획 완성

~9.6 플젝준비