인프라 장애 처리 지능화를 위한 데이터 저장 기법

Infra그룹 김동훈

목차

- 서론
- 관련 연구
- 데이터 전처리 플랫폼 아키텍처
- 결론 및 추후 연구과제

서론

- 대규모의 인프라 자원과 서비스를 관리하는 기업의 경우 가장 큰 문제는 장애
- 장애는 서비스 품질에 큰 영향을 주는 요소
- 다양한 플랫폼과 서비스에서 발생하는 장애 메시지를 실시간 수집하고 효율적인 데이터 전처리 하기 위한 데이터베이스 설계와 저장/처리 기법 제안

관련 연구

- 빅데이터 정의 및 특성
  - 가트너 정의: 큰 용량, 빠른 속도 그리고 높은 다양성을 갖는 정보 자산으로써 이를 통해 의사 결정 및 통찰 발견, 프로세스 최적화를 향상시키기 위해서는 새로운 형태의 처리 방식이 필요
  - 3가지 핵심 요소
    - 볼륨(Volume)
    - 속도(Velocity)
    - 다양성(Variety)
- 데이터 전처리
  - 분석 목적에 맞는 데이터를 수집하고 분석이 가능한 데이터로 축소, 제거, 수정 등과 같은 단계를 거쳐 최상의 분석 결과를 도출하기 위한 과정

관련 연구

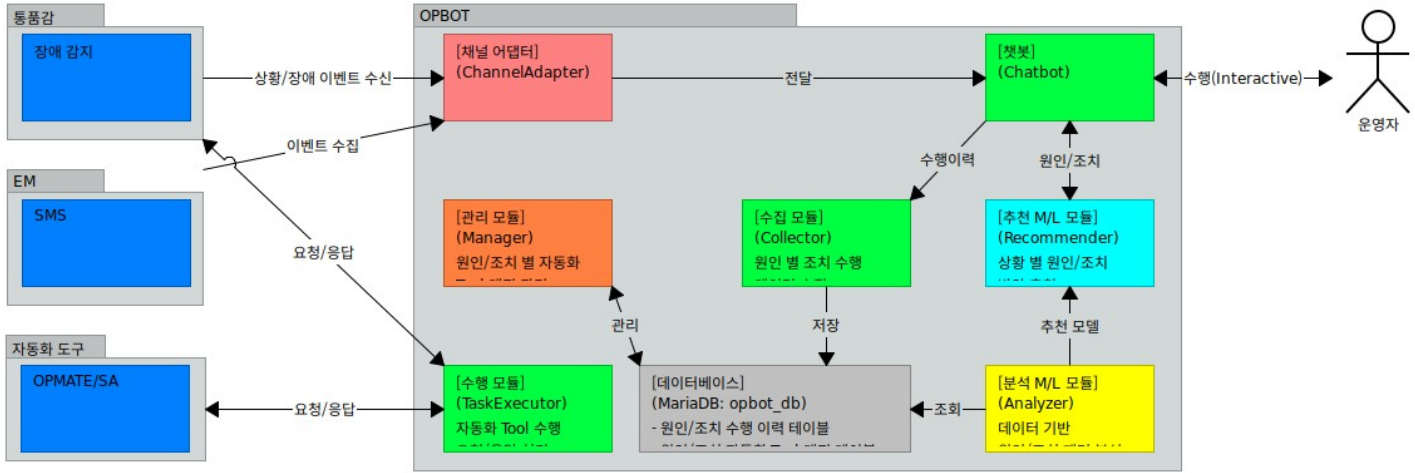
- 자연어 처리
  - 단어 추출/토크나이저/품사판별
- 노이즈 데이터 제거
  - 노이즈 데이터란 손상되거나 왜곡된 혹은 중복된 데이터를 의미
  - 노이즈 데이터는 머신러닝 모델의 부정확하고 잘못된 결과를 야기할 수 있기 때문에 학습 이전에 제거하여 학습 정확도 및 학습 속도 향상

관련 연구

- 정답지
  - 학습, 정확도 검증

데이터 전처리 플랫폼

- 운영체제
  - Ubuntu 18.04.3 LTS
- 데이터베이스
  - MariaDB 10.4
- 개발언어 및 관련 라이브러리
  - Python 3.7.3
  - Flask 1.1.1
  - Celery 4.3.0
  - RESTPlus 0.13.0
- 특징
  - 실시간 데이터 수집
  - 스케줄링을 이용한 데이터 전처리
  - 다양한 인프라 자원, 신규 서비스에 대한 장애 지원



데이터 전처리

- 데이터 구성
  - MS Excel 포맷
  - 장애구분, 문자발송일시, 장애요약, 발송메시지로 구성

	A	B	C	D
1	장애구분	문자발송일시	장애요약	발송메시지
2				[SKT상황]
3				
4				● 내용 : [SMS/MMS GW] Swing포탈 게시글 작성 시 문자 오발송 수신대상 안내 MMS 발송
5				● 영향 : Swing 사용자 34,971건(임직원 제외) 대상 분산 발송 건으로 모니터링 대응
6				● 원인 : Swing포탈 게시글 작성 시 문자 오발송 후속 대응
7				● 조치
8				- 14:24 34,971건 대상 MMS 발송 진행
9				- 14:43 Swing 사용자 34,971건(임직원 제외) MMS GW 발송 완료, 점검 진행
10				- 15:25 점검 완료, 상황종료.
11				
12	상황	202004241526	[SMS/MMS GW]	※ IT종합상황실 상황관리자 이상일 수석
13				[SKT상황]
14				
15				● 내용 : [SMS/MMS GW] Swing포탈 게시글 작성 시 문자 오발송 수신대상 안내 MMS 발송 중
16				● 영향 : Swing 사용자 34,971건(임직원 제외) 대상 분산 발송 건으로 모니터링 대응
17				● 원인 : Swing포탈 게시글 작성 시 문자 오발송 후속 대응
18				● 조치
19				- 14:24 34,971건 대상 MMS 발송 진행
20				- 14:43 Swing 사용자 34,971건(임직원 제외) MMS GW 발송 완료, 점검 진행 중
21				
22	상황	202004241448	[SMS/MMS GW]	※ IT종합상황실 상황관리자 이상일 수석
23				[SKT상황]
24				
25				● 내용 : [SMS/MMS GW] Swing포탈 게시글 작성 시 문자 오발송 수신대상 안내 MMS 발송 중
26				● 영향 : Swing 사용자 34,971건(임직원 제외) 대상 분산 발송 건으로 모니터링 대응
27				● 원인 : Swing포탈 게시글 작성 시 문자 오발송 후속 대응
28				● 조치
29				- 14:24 34,971건 대상 MMS 발송 진행 중
30				
31	상황	202004241426	[SMS/MMS GW]	※ IT종합상황실 상황관리자 이상일 수석
32				[SKT상황]
33				
34				● 내용 : [Swing-포탈] 게시판 글 작성 시 SMS 오발송으로 게시 대상 속성 변경 및 RTA 진행
35				● 영향 : Swing-포탈 게시판 글 작성 시 게시 대상 미지정 Case SMS 오발송(Swing 사용자 대상)
36				● 원인 : 전일 배포 영향
37				● 조치
38				- 09:48 게시 대상 필수 지정을 위한 속성 변경 및 RTA 준비
39				- 09:54 속성 변경(127건) 완료(SMS 오발송 미발생), RTA 준비 지속

시트1

데이터 전처리

- 자연어 처리(soynlp: https://github.com/lovit/soynlp)
  - 단어 추출: 품사 판별/단어 추출
  - 단어 정확도 측정: 단어 확률 계산

```
[ ]  1 list(noun_extractor._compounds_components.items())[:10]
```

```
[('법인정보조회화면', ('법인', '정보', '조회', '화면')),
 ('유선서비스해지화면', ('유선', '서비스', '해지', '화면')),
 ('번호이동인증화면', ('번호이동', '인증', '화면')),
 ('통합접촉이력화면', ('통합접촉이력', '화면')),
 ('고객센터이미지수신', ('고객센터', '이미지수신')),
 ('VDI파일전송시스템', ('VDI', '파일', '전송', '시스템')),
 ('KAIT진위확인처리', ('KAIT', '진위확인', '처리')),
 ('U+번호이동인증처리', ('U+', '번호이동', '인증처리')),
 ('KT번호이동인증처리', ('KT', '번호이동', '인증처리')),
 ('KT번호이동인증요청', ('KT', '번호이동', '인증', '요청'))]
```

```
[ ]  1 # WordExtractor : 단어일 확률을 계산하여 단어 추출
     2 print("유선          : " , word_scores["유선"].cohesion_forward)
     3 print("유선네        : " , word_scores["유선네"].cohesion_forward)
     4 print("유선네트      : ", word_scores["유선네트"].cohesion_forward)
     5 print("유선네트워크 : " , word_scores["유선네트워크"].cohesion_forward)
```

```
유선          :  0.37052456286427976
유선네        :  0.10796736793842097
유선네트      :  0.22674010465665753
유선네트워크 :  0.41050363359653064
```

데이터 전처리

- 자연어 처리(soynlp: https://github.com/lovit/soynlp)
  - word2vec

```
[ ]  1 print(word2vec.most_similar('점검'))
     2 print(word2vec.most_similar('사용자'))
     3 print(word2vec.most_similar('배포'))
     4 print(word2vec.most_similar('할인'))
     5 print(word2vec.most_similar('Email'))
     6 print(word2vec.most_similar('MSA'))
     7 print(word2vec.most_similar('MMS'))
```

```
[('테스트', 0.6649494171142578), ('모니터링', 0.588163435459137), ('기동', 0.5836498737335205), ('정상', 0.574669361114502), ('절체', 0.564501166343689), ('
[('상담원(', 0.9063305854797363), ('장애의심]', 0.9046738147735596), ('1고객센터', 0.891172647476l963), ('사용자(15명)', 0.886872291564941 4), ('가사번사용자(
[('개발기', 0.7823694348335266), ('현업(기획)', 0.7367343902587891), ('상품기준', 0.7137402296066284), ('상용기', 0.7073067426681519), ('RTA', 0.681172370910
[('요금제', 0.8965687155723572), ('/해지', 0.881174683570861 8), ('예약', 0.8718944191932678), ('선해지', 0.8689900636672974), ('해지', 0.8627392649650574), (
[('Marketing', 0.947655200958252), ('MSA', 0.9441289901733398), ('기업홈페이지', 0.9345674514770508), ('PoC]', 0.9333113431930542), ('SBPP(서버)', 0.92919999
[('PoC]', 0.9788060784339905), ('[인터넷]', 0.9683061838150024), ('기업홈페이지', 0.9615873098373413), ('Email', 0.9441288709640503), ('관리시스템,CSP,', 0.9
[('SMS', 0.8864167928695679), ('미지정', 0.8701639175415039), ('T가족나눔데이터', 0.8573136329650879), ('T지원금', 0.8374500274658203), ('할인', 0.836061775
```

데이터 전처리

- 자연어 처리(soynlp: https://github.com/lovit/soynlp)
  - 노이즈 데이터 제거: 한글 띄어쓰기, 불용어 제거(시간, 특수문자)

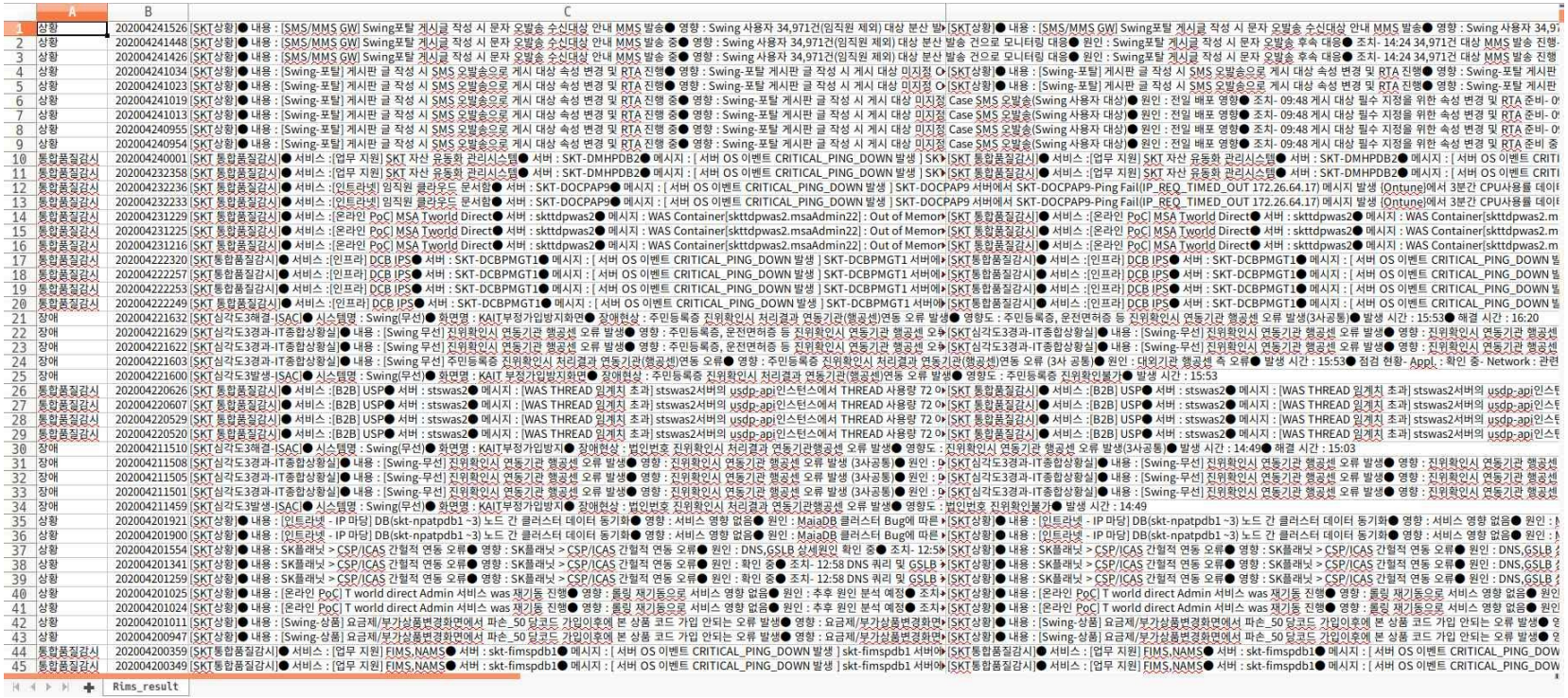
```
[ ]  1 tokenized_text = []
     2 for msg in msg_list:
     3     tokenized_text.append(" ".join(tokenizer.tokenize(msg)))

[ ]  1 tokenized_text[:3]

['[SKT 상황] ● 내용 : [S MS/MMS GW ] Swing 포탈 게시글 작성 시 문자 오발송 수신 대상 안내 MMS 발송 ● 영향 : Swing 사용 자 number 건(임직원 제외) 대상 분산
 '[SKT 상황] ● 내용 : [S MS/MMS GW ] Swing 포탈 게시글 작성 시 문자 오발송 수신 대상 안내 MMS 발송 중 ● 영향 : Swing 사용 자 number 건(임직원 제외) 대상 분
 '[SKT 상황] ● 내용 : [S MS/MMS GW ] Swing 포탈 게시글 작성 시 문자 오발송 수신 대상 안내 MMS 발송 중 ● 영향 : Swing 사용 자 number 건(임직원 제외) 대상 분
```

데이터 전처리

- 정답지: 정확도 검증

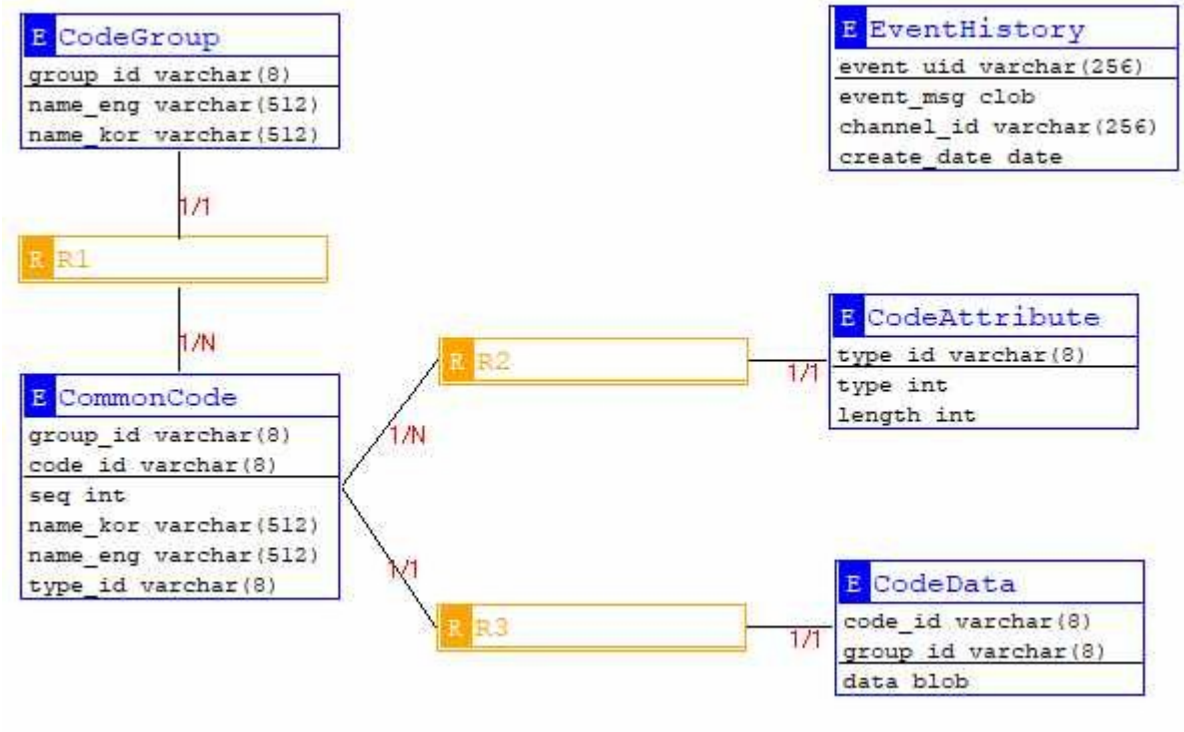


	A	B	C
1	상황	202004241526	[SKT상황]● 내용 : [SMS/MMS GW] Swing포탈 게시글 작성 시 문자 오발송 수신대상 안내 MMS 발송● 영향 : Swing 사용자 34,971건(임직원 제외) 대상 분산 발
2	상황	202004241448	[SKT상황]● 내용 : [SMS/MMS GW] Swing포탈 게시글 작성 시 문자 오발송 수신대상 안내 MMS 발송 중● 영향 : Swing 사용자 34,971건(임직원 제외) 대상 분산 발송 건으로 모니터링 대응● 원인 : Swing포탈 게시글 작성 시 문자 오발송 후속 대응● 조치- 14:24 34,971건 대상 MMS 발송 진행-
3	상황	202004241426	[SKT상황]● 내용 : [SMS/MMS GW] Swing포탈 게시글 작성 시 문자 오발송 수신대상 안내 MMS 발송 중● 영향 : Swing 사용자 34,971건(임직원 제외) 대상 분산 발송 건으로 모니터링 대응● 원인 : Swing포탈 게시글 작성 시 문자 오발송 후속 대응● 조치- 14:24 34,971건 대상 MMS 발송 진행
4	상황	202004241034	[SKT상황]● 내용 : [Swing-포탈] 게시판 글 작성 시 SMS 오발송으로 게시 대상 속성 변경 및 RTA 진행● 영향 : Swing-포탈 게시판 글 작성 시 게시 대상 미지정 O
5	상황	202004241023	[SKT상황]● 내용 : [Swing-포탈] 게시판 글 작성 시 SMS 오발송으로 게시 대상 속성 변경 및 RTA 진행● 영향 : Swing-포탈 게시판 글 작성 시 게시 대상 미지정 O
6	상황	202004241019	[SKT상황]● 내용 : [Swing-포탈] 게시판 글 작성 시 SMS 오발송으로 게시 대상 속성 변경 및 RTA 진행 중● 영향 : Swing-포탈 게시판 글 작성 시 게시 대상 미지정 Case SMS 오발송(Swing 사용자 대상)● 원인 : 전일 배포 영향● 조치- 09:48 게시 대상 필수 지정을 위한 속성 변경 및 RTA 준비- 0
7	상황	202004241013	[SKT상황]● 내용 : [Swing-포탈] 게시판 글 작성 시 SMS 오발송으로 게시 대상 속성 변경 및 RTA 진행 중● 영향 : Swing-포탈 게시판 글 작성 시 게시 대상 미지정 Case SMS 오발송(Swing 사용자 대상)● 원인 : 전일 배포 영향● 조치- 09:48 게시 대상 필수 지정을 위한 속성 변경 및 RTA 준비- 0
8	상황	202004240955	[SKT상황]● 내용 : [Swing-포탈] 게시판 글 작성 시 SMS 오발송으로 게시 대상 속성 변경 및 RTA 진행 중● 영향 : Swing-포탈 게시판 글 작성 시 게시 대상 미지정 Case SMS 오발송(Swing 사용자 대상)● 원인 : 전일 배포 영향● 조치- 09:48 게시 대상 필수 지정을 위한 속성 변경 및 RTA 준비- 0
9	상황	202004240954	[SKT상황]● 내용 : [Swing-포탈] 게시판 글 작성 시 SMS 오발송으로 게시 대상 속성 변경 및 RTA 진행 중● 영향 : Swing-포탈 게시판 글 작성 시 게시 대상 미지정 Case SMS 오발송(Swing 사용자 대상)● 원인 : 전일 배포 영향● 조치- 09:48 게시 대상 필수 지정을 위한 속성 변경 및 RTA 준비 중
10	통합품질감시	202004240001	[SKT 통합품질감시]● 서비스 :[업무 지원] SKT 자산 유동화 관리시스템● 서버 : SKT-DMHPDB2● 메시지 : [서버 OS 이벤트 CRITICAL_PING_DOWN 발생] SK
11	통합품질감시	202004232358	[SKT 통합품질감시]● 서비스 :[업무 지원] SKT 자산 유동화 관리시스템● 서버 : SKT-DMHPDB2● 메시지 : [서버 OS 이벤트 CRITICAL_PING_DOWN 발생] SK
12	통합품질감시	202004232336	[SKT 통합품질감시]● 서비스 :[인트라넷] 임직원 클라우드 문서함● 서버 : SKT-DOCPAP9● 메시지 : [서버 OS 이벤트 CRITICAL_PING_DOWN 발생] SKT-DOCPAP9 서버에서 SKT-DOCPAP9-Ping Fail(IP_REQ_TIMED_OUT 172.26.64.17) 메시지 발생 [Ontune]에서 3분간 CPU사용률 데이터
13	통합품질감시	202004232333	[SKT 통합품질감시]● 서비스 :[인트라넷] 임직원 클라우드 문서함● 서버 : SKT-DOCPAP9● 메시지 : [서버 OS 이벤트 CRITICAL_PING_DOWN 발생] SKT-DOCPAP9 서버에서 SKT-DOCPAP9-Ping Fail(IP_REQ_TIMED_OUT 172.26.64.17) 메시지 발생 [Ontune]에서 3분간 CPU사용률 데이터
14	통합품질감시	202004231229	[SKT 통합품질감시]● 서비스 :[온라인 PoC] MSA Tworld Direct● 서버 : skttdpwas2● 메시지 : WAS Container[skttdpwas2.msaAdmin22] : Out of Memor
15	통합품질감시	202004231225	[SKT 통합품질감시]● 서비스 :[온라인 PoC] MSA Tworld Direct● 서버 : skttdpwas2● 메시지 : WAS Container[skttdpwas2.msaAdmin22] : Out of Memor
16	통합품질감시	202004231216	[SKT 통합품질감시]● 서비스 :[온라인 PoC] MSA Tworld Direct● 서버 : skttdpwas2● 메시지 : WAS Container[skttdpwas2.msaAdmin22] : Out of Memor
17	통합품질감시	202004222320	[SKT통합품질감시]● 서비스 :[인프라] DCB IPS● 서버 : SKT-DCBPMGT1● 메시지 : [서버 OS 이벤트 CRITICAL_PING_DOWN 발생] SKT-DCBPMGT1 서버에
18	통합품질감시	202004222257	[SKT통합품질감시]● 서비스 :[인프라] DCB IPS● 서버 : SKT-DCBPMGT1● 메시지 : [서버 OS 이벤트 CRITICAL_PING_DOWN 발생] SKT-DCBPMGT1 서버에
19	통합품질감시	202004222253	[SKT통합품질감시]● 서비스 :[인프라] DCB IPS● 서버 : SKT-DCBPMGT1● 메시지 : [서버 OS 이벤트 CRITICAL_PING_DOWN 발생] SKT-DCBPMGT1 서버에
20	통합품질감시	202004222249	[SKT 통합품질감시]● 서비스 :[인프라] DCB IPS● 서버 : SKT-DCBPMGT1● 메시지 : [서버 OS 이벤트 CRITICAL_PING_DOWN 발생] SKT-DCBPMGT1 서버에
21	장애	202004221632	[SKT심각도3해결-ISAC]● 시스템명 : Swing(무선)● 화면명 : KAIT부정가입방지화면● 장애현상 : 주민등록증 진위확인시 처리결과 연동기관(행공센)연동 오류 발생
22	장애	202004221629	[SKT심각도3경과-IT종합상황실]● 내용 : [Swing 무선] 진위확인시 연동기관 행공센 오류 발생● 영향 : 주민등록증, 운전면허증 등 진위확인시 연동기관 행공센 오
23	장애	202004221622	[SKT심각도3경과-IT종합상황실]● 내용 : [Swing 무선] 진위확인시 연동기관 행공센 오류 발생● 영향 : 주민등록증, 운전면허증 등 진위확인시 연동기관 행공센 오
24	장애	202004221603	[SKT심각도3경과-IT종합상황실]● 내용 : [Swing 무선] 주민등록증 진위확인시 처리결과 연동기관(행공센)연동 오류● 영향 : 주민등록증 진위확인시 처리결과 연동기관(행공센)연동 오류 (3사 공통)● 원인 : 대외기관 행공센 측 오류● 발생 시간 : 15:53● 점검 현황- Appl. : 확인 중- Network : 관련
25	장애	202004221600	[SKT심각도3발생-ISAC]● 시스템명 : Swing(무선)● 화면명 : KAIT 부정가입방지화면● 장애현상 : 주민등록증 진위확인시 처리결과 연동기관(행공센)연동 오류 발생● 영향도 : 주민등록증 진위확인불가● 발생 시간 : 15:53
26	통합품질감시	202004220626	[SKT 통합품질감시]● 서비스 :[B2B] USP● 서버 : stswas2● 메시지 : [WAS THREAD 원계치 초과] stswas2서버의 usdp-api인스턴스에서 THREAD 사용량 72 0
27	통합품질감시	202004220607	[SKT 통합품질감시]● 서비스 :[B2B] USP● 서버 : stswas2● 메시지 : [WAS THREAD 원계치 초과] stswas2서버의 usdp-api인스턴스에서 THREAD 사용량 72 0
28	통합품질감시	202004220529	[SKT 통합품질감시]● 서비스 :[B2B] USP● 서버 : stswas2● 메시지 : [WAS THREAD 원계치 초과] stswas2서버의 usdp-api인스턴스에서 THREAD 사용량 72 0
29	통합품질감시	202004220520	[SKT 통합품질감시]● 서비스 :[B2B] USP● 서버 : stswas2● 메시지 : [WAS THREAD 원계치 초과] stswas2서버의 usdp-api인스턴스에서 THREAD 사용량 72 0
30	장애	202004211510	[SKT심각도3해결-ISAC]● 시스템명 : Swing(무선)● 화면명 : KAIT부정가입방지● 장애현상 : 법인번호 진위확인시 처리결과 연동기관행공센 오류 발생● 영향도 : 진위확인시 연동기관 행공센 오류 발생(3사공통)● 발생 시간 : 14:49● 해결 시간 : 15:03
31	장애	202004211508	[SKT심각도3경과-IT종합상황실]● 내용 : [Swing-무선] 진위확인시 연동기관 행공센 오류 발생● 영향 : 진위확인시 연동기관 행공센 오류 발생 (3사공통)● 원인 : D
32	장애	202004211505	[SKT심각도3경과-IT종합상황실]● 내용 : [Swing-무선] 진위확인시 연동기관 행공센 오류 발생● 영향 : 진위확인시 연동기관 행공센 오류 발생 (3사공통)● 원인 : D
33	장애	202004211501	[SKT심각도3경과-IT종합상황실]● 내용 : [Swing-무선] 진위확인시 연동기관 행공센 오류 발생● 영향 : 진위확인시 연동기관 행공센 오류 발생 (3사공통)● 원인 : D
34	장애	202004211459	[SKT심각도3발생-ISAC]● 시스템명 : Swing(무선)● 화면명 : KAIT부정가입방지● 장애현상 : 법인번호 진위확인시 처리결과 연동기관행공센 오류 발생● 영향도 : 법인번호 진위확인불가● 발생 시간 : 14:49
35	상황	202004201921	[SKT상황]● 내용 : [인트라넷 - IP 마당] DB(skt-npatpdb1 ~3) 노드 간 클러스터 데이터 동기화● 영향 : 서비스 영향 없음● 원인 : MaiaDB 클러스터 Bug에 따른
36	상황	202004201900	[SKT상황]● 내용 : [인트라넷 - IP 마당] DB(skt-npatpdb1 ~3) 노드 간 클러스터 데이터 동기화● 영향 : 서비스 영향 없음● 원인 : MaiaDB 클러스터 Bug에 따른
37	상황	202004201554	[SKT상황]● 내용 : SK플래닛 > CSP/ICAS 간헐적 연동 오류● 영향 : SK플래닛 > CSP/ICAS 간헐적 연동 오류● 원인 : DNS,GSLB 상세원인 확인 중● 조치- 12:5
38	상황	202004201341	[SKT상황]● 내용 : SK플래닛 > CSP/ICAS 간헐적 연동 오류● 영향 : SK플래닛 > CSP/ICAS 간헐적 연동 오류● 원인 : 확인 중● 조치- 12:58 DNS 쿼리 및 GSLB
39	상황	202004201259	[SKT상황]● 내용 : SK플래닛 > CSP/ICAS 간헐적 연동 오류● 영향 : SK플래닛 > CSP/ICAS 간헐적 연동 오류● 원인 : 확인 중● 조치- 12:58 DNS 쿼리 및 GSLB
40	상황	202004201025	[SKT상황]● 내용 : [온라인 PoC] T world direct Admin 서비스 was 재기동 진행● 영향 : 롤링 재기동으로 서비스 영향 없음● 원인 : 추후 원인 분석 예정● 조치
41	상황	202004201024	[SKT상황]● 내용 : [온라인 PoC] T world direct Admin 서비스 was 재기동 진행● 영향 : 롤링 재기동으로 서비스 영향 없음● 원인 : 추후 원인 분석 예정● 조치
42	상황	202004201011	[SKT상황]● 내용 : [Swing-상품] 요금제/부가상품변경화면에서 파손_50 닷코드 가입이후에 본 상품 코드 가입 안되는 오류 발생● 영향 : 요금제/부가상품변경화면
43	상황	202004200947	[SKT상황]● 내용 : [Swing-상품] 요금제/부가상품변경화면에서 파손_50 닷코드 가입이후에 본 상품 코드 가입 안되는 오류 발생● 영향 : 요금제/부가상품변경화면
44	통합품질감시	202004200359	[SKT통합품질감시]● 서비스 : [업무 지원] FIMS,NAMS● 서버 : skt-fimspdb1● 메시지 : [서버 OS 이벤트 CRITICAL_PING_DOWN 발생] skt-fimspdb1 서버에
45	통합품질감시	202004200349	[SKT통합품질감시]● 서비스 : [업무 지원] FIMS,NAMS● 서버 : skt-fimspdb1● 메시지 : [서버 OS 이벤트 CRITICAL_PING_DOWN 발생] skt-fimspdb1 서버에

Rims_result

유형별 데이터 저장 및 처리 기법

- 데이터베이스 설계
  - 장애 메시지
    - EventHistory
  - 설정 정보
    - CodeGroup
    - CommonCode
    - CodeAttribute
    - Code Data



유형별 데이터 저장 및 처리 기법

- 그룹코드(테이블명: CodeGroup)
  - 그룹 아이디(컬럼명: group_id): “GC_0001”
  - 한글명(컬럼명: name_kor): “노이즈 데이터”
  - 영문명(컬럼명: name_eng): “noise data”

- 공통코드(테이블명: CommonCode)
  - 그룹 아이디(컬럼명: group_id): “GC_0001”
  - 코드 아이디(컬럼명: code_id): “CD_0001”
  - 순번(컬럼명: seq): 0
  - 한글명(컬럼명: name_kor): “불용어”
  - 영문명(컬럼명: name_eng): “stopword”
  - 처리 타입 아이디(컬럼명: type_id): “T_0001”

- 코드속성(테이블명: CodeAttribute)
  - 처리 타입 아이디(컬럼명: type_id): “T_0001”
  - 처리 타입(컬럼명: type): 6
    - 0: Integer
    - 1: String
    - 2: Text
    - 3: DateTime
    - 4: Float
    - 5: Boolean
    - 6: PickleType
    - 7: LargeBinary
  - 길이(컬럼명: length): 0

유형별 데이터 저장 및 처리 기법

- 파이썬 객체 관계형 매퍼(ORM) 사용
  - SQLAlchemy(https://www.sqlalchemy.org/)
  - 유형별 데이터 처리
    - varchar: String 객체
    - clob: Text 객체
    - blob: Pickle type/LargeBinary 객체
    - date: DateTime 객체

-코드데이터(테이블명: CodeData)

- 코드 아이디(컬럼명: code_id): “CD_0001”
- 그룹 아이디(컬럼명: group_id): “GC_0001”
- 데이터(컬럼명: data)

```
{
‘employer’: {‘from’: [r'₩D{2,4}?(선임|수석|매니저|메니저)'], ‘to’: ‘담당자’},
‘time’: {‘from’: [r'₩d{2,3}[:₩.₩s]*₩d{2,2}(?!₩d)', r'₩(₩d{3}₩)₩s*₩d{4}[-₩.₩s]??₩d{4}'], ‘to’: ‘time’},
‘number’: {‘from’: [r'₩d{1,3}[,₩.]₩d{1,3}'], ‘to’: ‘number’},
“date”: {‘from’: [r'₩d{1,4}[년]₩d{1,2}[월]₩d{1,2}[일]', r'₩d{1,4}[년]₩s₩d{1,2}[월]₩s₩d{1,2}[일]', r'₩d{1,2}[월]₩s₩d{1,2}[일]', r'₩d{1,2}[월]₩d{1,2}[일]'], ‘to’: ‘date’}
}
```

결론 및 추후 연구과제

- 인프라 장애 이벤트 발생 시 원인 분석 및 조치 지능화 플랫폼의 아키텍처 제시
- 장애 메시지를 이용하여 머신러닝 모델을 만들기 위한 데이터 전처리 방안과 전처리 기법 제시
- 전처리 설정 정보와 장애 이벤트 정보를 저장하기 위한 데이터베이스 설계와 처리 기법 제시
- 인프라 장애 처리 지능화 플랫폼을 구현하고 성능 평가, 효율적인 사용자 인터페이스 및 API의 개발

End Of Document.