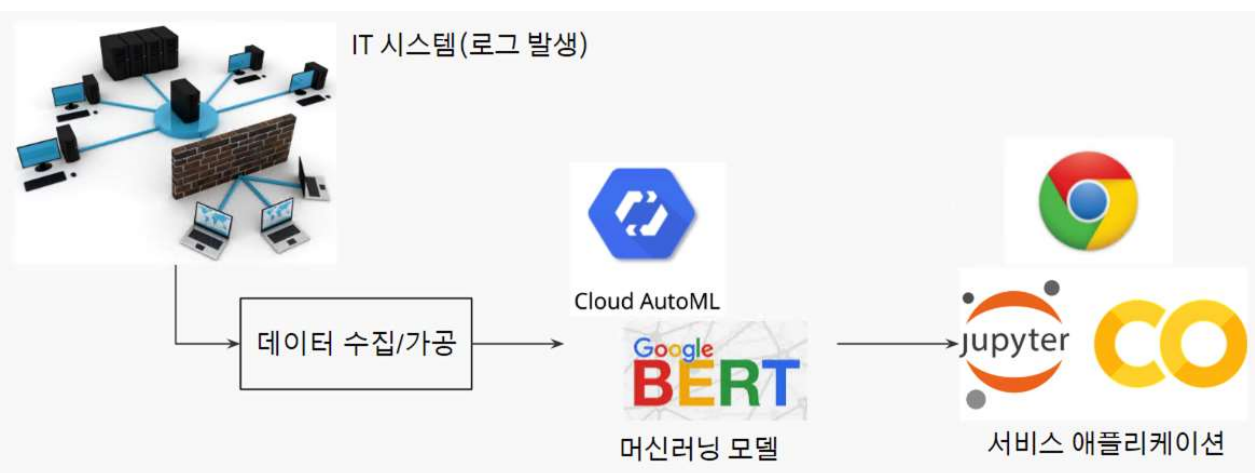


<SW분야 산학협력프로젝트(제안서): 샘플>

과 제 명	자연어 처리(NLP) 기반 IT 시스템 로그 분석		
참여기업	구글	담당자(직위)	정명훈 이사
	휴대전화	010-7287-0851	이메일
팀원요건	<ul style="list-style-type: none"> ○ Python 또는 Java 프로그래밍 가능 ○ Tensorflow 또는 유사 ML 프레임워크 유경험자 ○ 데이터 분석 경험자 선호 		
추진배경	<ul style="list-style-type: none"> ○ 산업계의 AI/ML 수요 비약적으로 증가하지만 역량 있는 기술인력 부족 ○ AutoML(ML 모델 자동 생성) 등 다양한 머신러닝 방법 활용 ○ 머신러닝을 위한 end-to-end 분석 시스템 구축 (데이터 수집, 가공, 분석, 서비스) 		
프로젝트 목표 및 내용	<p><그림을 포함하여 1~2페이지로 작성></p> <ul style="list-style-type: none"> ○ 대상 선정 – 분석 대상 IT 시스템과 로그의 종류 선택 ○ 데이터 수집 – 로그 데이터를 가져오기 위한 파이프라인 구축. 모델 개발 과정에서는 수집된 샘플 로그 활용 가능 ○ 데이터 가공 – Raw format의 로그 데이터를 자연어 분석에 적합한 형태로 가공하는 작업 ○ 모델 개발 – 구글클라우드(GCP)의 AutoML NLP와 커스텀 모델인 BERT를 이용하여 자연어 분석 기반 로그분석 머신러닝 모델 개발 ○ 서비스 애플리케이션 – 실제 IT 시스템으로부터 로그 수집하여 모델을 통해 분석하고 해당 결과를 출력하는 웹애플리케이션 제작 ○ 개발 환경은 GCP 사용(예상 비용 약 30만원) <div style="text-align: center;">  <p>IT 시스템(로그 발생)</p> <p>데이터 수집/가공</p> <p>Cloud AutoML</p> <p>Google BERT</p> <p>머신러닝 모델</p> <p>jupyter CO</p> <p>서비스 애플리케이션</p> </div>		
기대효과	<ul style="list-style-type: none"> ○ 데이터 분석의 기초적인 방법 습득 ○ 데이터 분석에 필요한 end-to-end 분석 시스템(데이터 수집, 가공, 분석, 서비스) 이해 		