

## 산학캡스톤 프로젝트 제안서

신청 기업명	DB FIS			사업장 위치	서울시 강남구 삼성로 96	
담당자	소 속	빅데이터파트	성 명	윌리엄	직 위	프로
	연락처	010-2285-2234	e-mail	williamwong@dbfis.co.kr		
관련분야	<input checked="" type="checkbox"/> 빅데이터 처리 <input type="checkbox"/> 애플리케이션SW <input type="checkbox"/> 인공지능 <input type="checkbox"/> 컴퓨터 보안 <input type="checkbox"/> 임베디드SW <input type="checkbox"/> 멀티미디어콘텐츠(AR/VR, 게임 등) <input checked="" type="checkbox"/> 기타 ( 자연어처리, 텍스트 마이닝 )					
프로젝트 명	반도체 원자재 시황 RISK 예측 모델을 위한 RISK 뉴스 라벨링 자동화					
프로젝트 개요	반도체 원자재 시황 RISK 예측 모델 개발을 위해서 필요한 시황 뉴스 데이터의 라벨링 자동화를 텍스트 마이닝 및 머신러닝 기반으로 개발					
추진 배경	<input type="checkbox"/> 배경/ 필요성 - 산업현장 문제점 기술					
	<p>최근 반도체 원자재 공급망 교란이 지속되면서 기업 생존에 큰 위협이 되는 가운데, 원자재 구매 RISK를 조기에 감지 및 선제적 대응의 필요성이 증가</p> <p>이를 위해서 비정형 뉴스 데이터를 분석하여 원자재 공급망에 영향을 미치는 시황 RISK의 조기 식별 및 예측하려는 시도를 하고 있음</p> <p>그러나 시황 RISK 식별 및 예측의 정확도를 높이기 위해서는 학습용 뉴스 빅데이터의 정확한 라벨링 작업(방대한 뉴스 데이터에서 반도체 관련 시황 RISK 뉴스만 선별해서 추출하는 작업)이 선행되어야 하는데, 이때 상당한 시간투입과 비용이 발생</p> <p>따라서, 텍스트 마이닝 및 머신러닝 기법을 사용하여 자동으로 RISK 뉴스 데이터를 정확히 라벨링 하고자 함</p>					
개발 목표	<input type="checkbox"/> 개발 목표 - 개발하고자 하는 시스템의 목표					
	1) 반도체 전문 뉴스 웹사이트에서 관련 뉴스 수집 및 클라우드 적재 2) 텍스트 마이닝 기법을 통해 시황 RISK 키워드 추출 및 딕셔너리 정의 3) 시황 RISK 뉴스 데이터 라벨링 자동화 모델 개발 4) 시황 RISK 뉴스 데이터 라벨링 자동화 및 정확도 검증을 위한 서비스					
	<input type="checkbox"/> 기업체 지원 가능 사항 (산업체 멘토, 개발 도구, 개발용 샘플 등)					

	<ul style="list-style-type: none"> <li>- 산업체 멘토 (텍스트 마이닝 분석가)</li> <li>- 수행 환경 (사내 오피스 공간 &amp; 원격 디지털 공간 제공)</li> <li>- 분석 환경 (클라우드 기반의 데이터 분석 인프라 및 도구 제공)</li> </ul>
<b>결과물</b>	<input type="checkbox"/> 최종 기대 결과물 (졸업작품으로 전시할 소프트웨어 시현물)
	1) RISK 딕셔너리 2) 반도체 원자재 시황 RISK 뉴스 라벨링 자동화 모델 3) 라벨링 검증 및 자동화 서비스
<b>인턴십 연계</b>	<input type="checkbox"/> 여름방학 인턴십 <input type="checkbox"/> 겨울방학 인턴십 <input type="checkbox"/> 인턴십 안함

※ 학생들이 본 제안서를 읽고 본인에게 적합한 과제를 선택할 수 있도록 작성 요망

※ 추진배경 및 개발목표 추가 기술이 필요한 경우 별지 작성 및 첨부 요망