

# 2021 네이버 겨울 인턴십(연합전공 인공지능분야) 과제

## \* 과제목록

1. 컴퓨터 비전을 활용한 피싱사이트 탐지
2. 지능적 문서 분류(Text Categorization)
3. 실시간 교통 정보를 활용한 패턴 교통 정보 모델링
4. 실시간 버스 정보를 활용한 정류장간 소요시간 패턴 모델링
5. ML/DL Data pipeline 구축을 통한 Model 학습 자동화 플랫폼 구축

\* 각 과제 스펙에 지원 해야하는 URL 이 명시되어 있습니다.

- 기간내에 해당 지원 URL 에 지원을 하셔야 전형 진행이 가능합니다.

(지원 가능 기간 : 11/11 ~ 11/15)

- 하나의 지원 URL 에 여러 과제들(타 학과 전형 과제 포함)이 포함되어 있으므로,  
“지원사항” 항목에서 자신의 지원 주제를 선택하시기 바랍니다.

\* 1 인당 하나의 과제만 지원이 가능합니다.

\* 전형방식(코딩테스트 유/무)은 과제별로 상이합니다.

과제 번호	1	모집 부서	회원&인증플랫폼
제목	컴퓨터 비전을 활용한 피싱사이트 탐지	모집인원	1명
지원 URL	<a href="#">링크</a>		
과제 상세 설명	피싱으로 의심되는 사이트 중 실제 네이버로그인 피싱사이트로 운영되는 사이트를 컴퓨터 비전을 활용하여 탐지한다.		
필수 요구 사항	2021년 3월 기준 학부 4학년(7차 학기) 이상 머신러닝 등 인공지능 기초 과목 이수		
우대 사항	python 개발 가능한 자, 컴퓨터 비전 과목 이수자, 프론트엔드개발 관련 기술 보유자		
근무지역	경기도 성남시 분당구 불정로 6 네이버 그린팩토리		
부서의 인재상	연구를 진행함에 있어 진취적인 자세를 보여주고 현재에 얽매이지 않는 참신한 아이디어가 많은 인재		
선발 방식	서류 + 면접		
<b>주차별 계획</b>			
1주차	서버에 수집되는 데이터 마이닝		
2주차	데이터 정제		
3주차	수집된 의심 사이트 visualizing		
4주차	피싱사이트 판단 모델 모델링		
5주차	피싱사이트 판단 모델 모델링		
6주차	피싱사이트 판단 모델 모델링		
7주차	제작된 모델을 활용하여 서비스에 적용		
8주차	프로젝트 개선점 보완 및 마무리		

과제 번호	2	모집 부서	UGC Data Dev
제목	지능적 문서 분류(Text Categorization)	모집인원	1 명
지원 URL	<a href="#">링크</a>		
과제 상세 설명	제공된 한글 문서 분류에 적합한 딥러닝 모델 선택 기존 통계기반 문서 분류와 성능 비교 카테고리 확장 자동화 방법 고안		
필수 요구 사항	2021 년 3 월 기준 학부 4 학년(7 차 학기) 이상 머신러닝에 대한 기본적인 이해. Pytorch 가능.		
우대 사항	1. NLP 관련 지식 및 머신러닝 모델(BERT, TF-IDF 등..) 경험 2. 웹(프론트)를 통한 데이터 시각화 경험		
근무지역	경기도 성남시 분당구 불정로 6 네이버 그린팩토리		
부서의 인재상	논리적인 사고력, 꼼꼼함		
선발 방식	서류 + 면접		
<b>주차별 계획</b>			
1 주차	과제 적합한 모델 조사		
2 주차	과제 적합한 모델 조사		
3 주차	모델 적용 및 성능 분석		
4 주차	모델 적용 및 성능 분석		
5 주차	모델 적용 및 성능 분석		
6 주차	모델 적용 및 성능 분석		
7 주차	시각화 또는 결과 정리		
8 주차	시각화 또는 결과 정리		

과제 번호	3	모집 부서	Maps 개발 / Traffic Intelligence
제목	실시간 교통 정보를 활용한 패턴 교통 정보 모델링	모집인원	2 명
지원 URL	<a href="#">링크</a>		
과제 상세 설명	<p>사용자들에게 정확한 교통정보를 제공하기 위해서 실시간 교통 정보는 필요합니다. 하지만 내비게이션은 미래 시간에 대한 정보가 필요하기 때문에 교통 정보의 통계적 해석을 통한 정확한 정보 예측을 필요합니다.</p> <p>다양한 형태의 모델링을 활용하여 패턴 교통 정보 생성을 해보고 이를 통해서 실제로 정보 예측을 향상시킬 수 있는 방안을 제시합니다.</p>		
필수 요구 사항	2021 년 3 월 기준 학부 4 학년(7 차 학기) 이상 알고리즘 강의 이수		
우대 사항	Spring Framework 사용 우대. Hadoop 기반 빅데이터 경험 우대		
근무지역	경기 성남시 분당구 분당로 55 분당퍼스트타워		
부서의 인재상	문제를 푸는 것이 재미있고, 늘 성장하고자 하는 개발자		
선발 방식	서류 + 면접 + 코딩 테스트		
<b>주차별 계획</b>			
1 주차	모델 설정 및 데이터 준비		
2 주차	모델 설정 및 데이터 준비		
3 주차	모델 코딩		
4 주차	모델 코딩		
5 주차	모델 코딩		
6 주차	모델 평가 및 개선		
7 주차	모델 평가 및 개선		
8 주차	리팩토링		

과제 번호	4	모집 부서	Maps 개발 / MSP(Map Service Platform)
제목	실시간 버스 정보를 활용한 정류장간 소요시간 패턴 모델링	모집인원	2 명
지원 URL	<a href="#">링크</a>		
과제 상세 설명	버스들은 신호대기 등의 이유로 정류장간 소요시간이 불규칙합니다. 그러나 교통상황 등을 고려하더라도 소요시간의 패턴은 있을 것으로 기대합니다. 통계적 정보를 활용하여 버스의 정류장별 소요시간을 어느 정도 예측할 수 있다면 사용자들에게 좋은 정보를 줄 것으로 기대합니다.		
필수 요구 사항	2021 년 3 월 기준 학부 4 학년(7 차 학기) 이상 알고리즘 강의 이수		
우대 사항	Spring Framework 사용 우대		
근무지역	경기 성남시 분당구 분당로 55 분당퍼스트타워		
부서의 인재상	문제를 푸는 것이 재미있고, 늘 성장하고자 하는 개발자		
선발 방식	서류 + 면접 + 코딩 테스트		
<b>주차별 계획</b>			
1 주차	데이터 분석 및 준비. 모델 결정		
2 주차	데이터 분석 및 준비. 모델 결정		
3 주차	모델 코딩		
4 주차	모델 코딩		
5 주차	모델 코딩		
6 주차	평가 및 서비스 데모 버전 개발		
7 주차	평가 및 서비스 데모 버전 개발		
8 주차	리팩토링		

과제 번호	5	모집 부서	미디어 Tech / 이용자피드백플랫폼
제목	ML/DL Data pipeline 구축을 통한 Model 학습 자동화 플랫폼 구축	모집인원	2명
지원 URL	<a href="#">링크</a>		
과제 상세 설명	데이터의 수집, 가공, 서플등 ML/DL Model 개발시 필요한 big data 처리를 pipeline 구축을 통해 네이버 뉴스외 다수의 서비스에 적용된 클린봇 ver 2.0 의 지속적인 강화 학습을 위한 Data Lake 플랫폼 구축		
필수 요구 사항	2021년 3월 기준 학부 4학년(7차 학기) 이상 java, spring Framework 사용 가능자		
우대 사항	학부 인공지능 수업 이수, kafka 경험자 우대, Python 개발경험, 대용량 데이터 설계에 대한 기본 지식과 모델링 경험		
근무지역	경기도 성남시 분당구 불정로 6네이버 그린팩토리		
부서의 인재상	본인과 동료들의 성장, 그리고 플랫폼의 성장 더 나아가 아름다운 세상을 함께 만드실분		
선발 방식	서류 + 면접 + 코딩 테스트		
<b>주차별 계획 (모든 과정은 ML/NLP/service 실무진의 멘토링과 함께 진행됩니다)</b>			
1주차	미디어인텔리전스에서 사용하는 ML Data flow 아키텍처의 이해		
2주차	ML/DL Data pipeline(수집-추출-태깅-학습-평가)을 위한 database 및 stream 설계. 기존 Classification 모델 개발 프로세스 분석		
3주차	Data warehousing service(REST API)설계 및 개발		
4주차	Data warehousing service 연동을 위한 FE 개발		
5주차	Data warehousing service 연동을 위한 FE 개발		
6주차	Pipeline manager 및 data wharehousing admin 개발		
7주차	Data 학습 및 평가를 위한 Naver ML Platform 과 연동 개발		
8주차	Data 학습 및 평가를 위한 Naver ML Platform 과 연동 개발		