

빅데이터 사이언스 실무자 양성과정

< 프로젝트 착수보고서 >

주 제	CIFAR10 응용모델 만들기 프로젝트 - 운송수단을 중심으로		
팀 명	운송수단팀	일 자	2019년 10월 22일
팀 장	김지은 (tel0r0@naver.com)	팀 원	김은비, 김지은, 성현민, 홍성민

1. 과제 개요

운송수단의 발달은 세계화의 진행 과정에서 빼놓을 수 없는 핵심 요소이며, 이를 통해 상이하고 이국적인 문화가 각 나라에 소개되며 인류 문화 및 전체적인 인류의 발전에 기여하였다. 이에 팀은 컴퓨터가 이러한 운송수단을 인지하고, 구분할 수 있는 것이 의미 있을 것이라는 판단 하에 CIFAR10을 벤치마킹하여, 현대 사회에서 많이 사용되는 운송수단 중, 비슷한 특성을 가지거나, 구별하기 힘든 10개의 운송수단을 선정하여 CNN모델을 구성을 진행 해 보았다.

2. 과제 진행 방향

충분한 학습이 정확도에 가장 큰 영향을 미칠 것 이라는 가정 하에 각 카테고리별로 최소 500장 가량의 학습 이미지를 웹크롤링을 통해 확보 후, Keras를 사용한 CNN학습을 Relu/Sigmoid 등의 다양한 활성화함수를 적용해 보며 정확도를 높이는 방식으로 진행 할 예정이다.

3. 결과 보고 예측 및 실행 계획

- 역할 분담

김은비	배, 헬리콥터 학습 이미지 확보
김지은	자전거, 버스, 오토바이 학습 이미지 확보
성현민	웹크롤링 모델 구성, 트럭, 자동차 학습 이미지 확보
홍성민	착수보고서 및 PPT Draft 작성. 비행기, 기차, 열기구 학습 이미지 확보

- 결과 도출 전략 및 가설

뚜렷한 차이점을 보이는 운송수단은 컴퓨터가 쉽게 학습하고 구별할 수 있을 것이며, 비슷한 특성을 공유하는 이동수단은 컴퓨터의 학습과 판단이 어려울 것이라고 판단된다. 즉, 하늘에서 이동하는 수단인 비행기/헬리콥터/열기구 및 도로를 달리는 버스/트럭/자동차를 컴퓨터가 구분하기 쉽지 않을 것이라고 생각하였다. 따라서 팀은 많은 학습이미지를 확보해 이를 극복하려고 하였다.

- 프로젝트를 통해 얻고 싶은 것

컴퓨터의 학습능력이 인간의 뇌의 학습 능력과 비등해지고 있다는, 혹은 더 뛰어나다는 시대를 살고 있는 가운데, 컴퓨터의 학습능력을 개인이 짠 코드로 실험해 볼 수 있는 기회로 삼고 싶다. 아울러, 이번 프로젝트를 통해 CNN 학습을 다시 리뷰할 수 있는 시간을 가졌으면 한다.

감사합니다

- 결과보고서 제출 과제

- 1) 발표자료 : 프리스타일, 파워포인트 12page 내외
- 2) 소스코드 : 데이터수집, 데이터분석은 모듈별 별도 파일 작업후 (폴드를 하나의 파일로 압축)
- 3) 기타자료 : 그외 프로젝트 산출물