

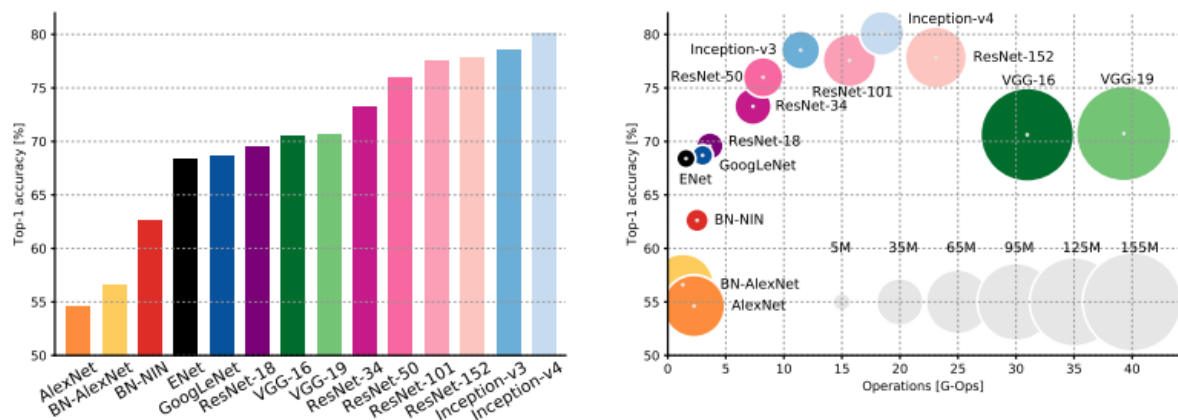
수요예측 알고리즘 경량화 및 서비스 구현 - 착수 보고서

팀명: 어쩔컴퓨터

팀 번호: 08

1. 과제 배경

기존의 수요 예측 알고리즘들은 매우 큰 데이터 셋을 학습하는 것을 전제로 개발되어 상당히 복잡하고 거대하기 때문에 상용화하기에는 학습 비용이 많이 들었다.



<그림 1.> CNN 기반 딥 러닝 모델들의 Parameter 수에 따른 Top-1 Accuracy 분석 차트, (AN ANALYSIS OF DEEP NEURAL NETWORK MODELS FOR PRACTICAL APPLICATIONS)

그리고 최근의 IoT 기기들이나 모바일 기기들 같이 상대적으로 저성능의 디바이스들에도 딥 러닝 모델을 탑재하려는 움직임들이 늘어나고 있다.

따라서, 이를 최대한 경량화하여 적절한 모델 크기를 가지고 일정 수준의 정확도를 보장하기 위한 경량화를 진행하는 것을 과제의 핵심 목표로 둘 것이다.

본 과제의 부가 목표는 다음과 같다.

- 기존 알고리즘과 경량화된 알고리즘간의 성능 분석
- 경량화된 알고리즘을 이용해 웹 기반 서비스 기획 및 배포

2. 연구 방향

2.1. 모델 경량화

과제의 핵심이 되는 모델 경량화는 정확도의 손실을 기존 모델 대비 최소화 시키면서 모델의 크기와 연산량을 줄이는 작업이다.

최근까지의 모델 경량화 기술들은 크게 이렇게 요약하자면 다음과 같다.

기술 분류	기술 요약	접근 방식
Pruning	기준 값 이하의 가중치나 뉴런을 제거하며 반복 학습	기 학습된 모델을 이용
Quantization	가중치, 활성화값, 기울기값을 보다 낮은 비트너비로 표현되는 값으로 대체	기 학습된 모델을 이용하거나, 모델 학습 때 적용
Knowledge Distillation	기 학습된 모델을 teacher 모델로 하여 경량화된 student 모델을 학습함	기 학습된 모델을 이용
Compact Network Design	연산 효율성을 향상을 위한 뉴럴 네트워크 구조를 변경하는 방식	모델 학습 때부터 적용

<표 1.> 모델 경량화 기술의 분류(딥러닝 모델 경량화 기술 분석, 2020, 지능형인프라기술 연구단)

모델을 선정함에 따라, 해당 모델에 맞는 접근 방식을 가지고 경량화를 진행할 것이다.

2.2. 데이터 수집

원래 계획한 데이터는 물품 보관 창고와 수요지를 연결해주는 일자별 보관 수요에 대한 데이터였다. 하지만 이 부분에 대한 데이터는 아직 확보가 어려웠다.

따라서, 다음 현재 단계에서 조사한 데이터 종류는 크게 2가지이다.

- 교통 수요 예측과 관련된 데이터 셋
- 주식 시계열 데이터 셋

해당 데이터를 수집하기 위해서 다음과 같은 공개 데이터를 활용하고자 한다.

- 인터넷에 공개된 Public Data Set 이용
- 국가교통데이터베이스 KTDB의 데이터 이용
- 금융기관의 OpenAPI를 이용하여 데이터 수집
 - 대신증권 CybosPlus API
 - 키움증권 OpenAPI+

2. 개발 일정

개발 일정

Aa 프로세스	📅 일정	▼ 진행상황
<u>경량화할 수요 예측 알고리즘 조사</u>	@2022년 5월 23일 → 2022년 8월 9일	완료
<u>데이터 종류 선정</u>	@2022년 5월 23일 → 2022년 7월 12일	완료
<u>알고리즘 경량화 작업</u>	@2022년 8월 8일 → 2022년 8월 31일	시작전
<u>중간 보고서 작성</u>	@2022년 7월 28일 → 2022년 8월 1일	진행중
<u>백엔드 개발</u>	@2022년 7월 12일 → 2022년 8월 31일	진행중
<u>프론트 앤드 개발</u>	@2022년 7월 12일 → 2022년 8월 31일	진행중
<u>최종 보고서 작성</u>	@2022년 9월 1일 → 2022년 9월 30일	시작전
<u>알고리즘 구현</u>	@2022년 7월 12일 → 2022년 8월 9일	진행중

Aa 프로세스	📅 일정	▼ 진행상황
<u>서비스 테스트</u>	@2022년 9월 1일 → 2022년 9월 30일	시작전

3. 역할 분담

이름	역할
서준호	수요 예측 알고리즘 조사, 데이터 수집 및 전처리, 알고리즘 경량화 작업, 서비스 기획, 백엔드 개발
윤정현	수요 예측 알고리즘 조사, 알고리즘 경량화 작업, 서비스 기획, 백엔드 개발
김민태	알고리즘 경량화 작업, 서비스 기획, 유저 인터페이스 개발

4. 참고 문헌

<그림 1.> AN ANALYSIS OF DEEP NEURAL NETWORK MODELS FOR PRACTICAL APPLICATIONS

<표 1.> 딥러닝 모델 경량화 기술 분석(지능형인프라기술연구단)