

# *Capstone Design* : 산업수학 프로젝트 정례 보고서

Team 복소수

May 11, 2025

## 1 What we have done

### 1.1 주제 세분화 - 공동

중간보고회 과정에서 받은 피드백을 바탕으로 주제를 세분화하였다.

- 주제: VPN 트래픽 예측 기반 동적 로드밸런싱 알고리즘 개발  
세부 구현 1: VPN/NON-VPN 분류 결과 기반 로드밸런싱  
세부 구현 2: 시계열 예측을 통한 시간대별 루트 변경

### 1.2 분류 모델 수정 - 전용진

전처리와 분류 과정에 대한 수정이 필요하다는 피드백을 받아 수정하였다.

- 전처리: 전처리 과정에서 이상치 제거가 과하게 이루어져서 데이터가 손실되는 경우가 발생하였다.  
해결: Outlier가 발생하는 경우는 데이터의 특성상 자연스러운 경우가 많기 때문에, 이상치 제거를 생략해 보았다. 추후 데이터에 대한 분석을 통해 올바른 방법을 찾아보도록 하겠다.
- 분류모델: 상관관계 분석을 통한 feature selection을 진행하였으나, 이는 선형 관계만을 고려한 방법이기 때문에 부적절할 수 있다.  
해결: t-test 가설 검정 기법을 사용하여 feature selection을 진행하였다. t-test는 두 집단 간의 평균 차이를 검정하는 방법으로, 비모수적 방법이기 때문에 선형 관계를 가정하지 않는다.

### 1.3 VPN 트래픽 캡처 - 김재영

분류 모델이 실제로도 잘 작동하는지 확인하기 위해, VPN 트래픽을 캡처하였다. 추후 알고리즘이 완성된 이후 성능의 평가에도 해당 작업은 필수적인 과정이다. VPNBOOK을 사용하여 VPN 트래픽을 캡처하였고, 캡처한 트래픽은 pcap 파일로 저장하였다.

## 2 로드밸런싱 테스트 환경 구축 - 주민찬

간단한 로드밸런서 코드와 백엔드 서버를 구축하였다.

- 파이썬의 fastapi 프레임워크 이용
- 백엔드 서버의 엔드포인트에 /health를 통해 cpu 사용률을 확인 가능
- 백엔드 서버를 여러 포트번호를 통해 실행시키고 로드밸런서를 통해 현재 가장 빠른 서버로 요청을 보내주는 역할로 실행
- 로드밸런서는 현재 실행중인 요청처리수와 평균 응답 시간을 기준으로 작은 서버를 우선적으로 보냄
- 동시에 100개, 총 1000개의 요청을 보내서 서버 부하 테스트 진행

## 3 What we will do

- 분류 모델 성능 평가: 캡처한 VPN 트래픽을 사용하여 분류 모델의 성능을 평가할 예정이다.
- 로드밸런싱 알고리즘 고안: 현존하는 로드밸런싱 알고리즘을 분석하고, 이를 바탕으로 동적 로드밸런싱 알고리즘 아이디어를 고민해 보겠다.