캡스톤 18조 장우혁 멘토링 활용 보고서

프로젝트명	O24Sec	팀명	멜러리를 찾아서

팀 멘토링

* 모든 멘토링을 상세하게 적으면 분량이 너무 길어져, 간단히 미팅 주제와 개인별로 느낀점을 적었습니다.

◉ [01/22] 팀 미팅 - 윤명근 교수님
- 연구에 사용할 데이터 분석 및 도식화에 대한 피드백
=> 느낀점: 이러한 데이터를 다뤄보는 것이 처음이었는데, 데이터를 다루는 방법과 시각화의 필요성을 느꼈습니다.

◉ [01/26] 팀 미팅 - 윤명근 교수님
- 1차 주제 확립: 보안관제 이벤트 처리에서, 비지도학습 기반으로 기존기술의 한계와 알람피로, 데이터 오분류 등을 해결하고, 군집화를 통한 여러 머신러닝등의 데이터셋으로 활용할 수 있는 데이터 군집화 기술 개발
- 주제에 관한 토론 및 피드백
=> 느낀점: 산학지정주제에 대한 상세한 교수님의 설명과 토론을 진행하면서 주제와 배경지식에 대한 더 깊은 이해가 가능했습니다.

◉ [01/28] 윈스 미팅 - 최병환 팀장님, 윤명근 교수님, 박하명 교수님
- 주식회사 윈스에 방문하여, 기존 연구에 대한 피드백과 제안 기술에 대한 토론 진행.
- 데이터 셋에 대한 요구사항 및 개선사항 접수, 윈스와의 제안 기술 관련 협의
=> 느낀점: 산학지정주제 관련 회사와의 미팅을 진행했는데, 실제로 개발할 기술이 사용될 관제센터와 다른 여러 기술들을 보며 실무 정보와 자극을 얻을 수 있었습니다.

◉ [02/02] 팀 미팅 - 윤명근 교수님
- 목표사항인 비지도학습 군집화를 위한 단계별 계획 수립
- 단계의 목적 및 시나리오 설립, 알고리즘에 관한 토의
- 1차 목표사항: 보안관제 이벤트의 객체 분류(IPS장치 기준 내외부 구별, 서버와 클라이언트 구별)
- 2차 목표사항: 보안관제 이벤트의 페이로드 암호화 판별(암호화하지 않은 데이터셋을 구성)
- 3차 목표사항: 기존의 주제인 데이터 군집화 기술 개발
=> 이루고 싶은 목표에 대해 여러 단계를 세분화 과정을 거쳤는데, 문제에 대한 접근 방식을 공부할 수 있었습니다.

다음 장에 계속

팀 멘토링

- **[02/17] Ceeya 미팅 - 문성익 멘토님**
  - 문제 접근 방식, 문제 해결 방식, 프로젝트 진행 방식에 대한 피드백
  - 멘토님과의 통신 체계 마련
  - 1차 목표사항과 관련하여 여러 아이디어 토의:
  멘토님의 네트워크 지식을 활용하여 IPS장비 기준 내외부와 서버 클라이언트를 구별할 수 있는 아이디어를 같이 토론
  - 3차 목표사항과 관련하여 여러 아이디어 토의:
  멘토님의 클러스터링 기술 사용 경험을 통한 조언. 클러스터링 기술별 성능의 차이, 데이터 전처리와 피쳐 선정, 가중치관련 조언
  네트워크 데이터의 특수성을 활용한 클러스터 방안 조언. 페이로드의 특성을 활용하여 클러스터링
  => 느낀점: 미국 캘리포니아에 계시는 멘토님과 줌으로 미팅을 진행했는데, 주제에 대해 모르시는 전문가를 상대로 하는 미팅을 통해, 추후에 있을 캡스톤 평가에 대한 PT실력을 증진할 수 있었으며, 멘토님의 다양한 전문분야 지식을 통해 여러 방면의 도움과 아이디어를 받고 토론해볼 수 있는 소중한 기회였습니다.

- [02/26] 팀 미팅 - 윤명근 교수님
  - 1차, 2차 목표사항의 현재 진행 상황에 대한 PT 및 그에 대한 피드백:
  1차 목표사항에 있어 간결한 룰을 유지해야 할 필요성이 있으며, 강건한 규칙을 순차적으로 적용해야한다는 피드백
  2차 목표사항에 있어 엔트로피를 구하는 방식과 데이터 형 변환 관련 피드백
  => 느낀점: 실험 -> 구현 과정에서 실험 종료 후 구현의 순서가 아니라 실험 -> 구현 -> 검증 -> 추가 실험 등의 방식을 채택해야함을 느끼게 해주셨습니다. 또한 룰베이스기반 분류방식에서 더 나은 규칙 적용 방식에 대해서도 알 수 있었습니다.

- [03/12] 윈스담당 팀 미팅 - 윤명근 교수님
  - 윈스 프로젝트와 관련한 사람들과의 팀미팅 진행:
  현재까지 진행된 1차, 2차 목표사항의 구현에 대한 피드백 및 앞으로의 진행상황에 대한 계획 수립
  1차 목표사항에 있어 겪는 여러 예외사항에 대한 토론
  => 느낀점: 교수님 말고도 교수님 아래서 연구를 진행하는 분들의 피드백을 받을 수 있는 시간이었습니다. 지금까지 제가 했던 접근방식에 대한 반성과 보고 체계 및 도식화의 중요성을 알 수 있었습니다. 또한 제가 진행했던 과정을 정리하고 돌아보며 앞으로의 진행 계획을 수립할 수 있는 시간이었습니다.

다음 장에 계속

개인 멘토링	* 제가 맡은 담당 분야인 1차 목표사항 달성을 위해 교수님과 진행했던 1대1 미팅 내용 및 느낀점을 작성했습니다. ▸ [02/05] 1대1 미팅 - 윤명근 교수님 - 데이터셋에서 출발지IP, 도착지IP 빈도 및 IP를 비트(16~32bit)단위 클래스로 묶었을 경우를 분석하고 도식화 한 자료를 분석 - IPS장비 기준 내부, 외부에서의 데이터를 나누어 그래프로 도식화한 자료를 분석 => 느낀점: 교수님이 가장 많이 등장한 포트번호를 순위를 매겨서 출력하면 더 잘 보일 것이라는 피드백을 주셨는데, 이를 통해 해당 IP대역에서 많이 사용된 포트 통계를 보고 서버, 클라이언트의 단서를 얻을 수 있었습니다. 또한 데이터 분류 및 도식화를 진행하며 여러 프로그래밍 스킬을 공부하고 배울 수 있었습니다. ▸ [02/15] 1대1 미팅 - 윤명근 교수님 - 02/05 1대1 미팅에서의 피드백을 통해 추가 구현을 한 데이터를 분석 => 느낀점: 교수님의 피드백을 통해 추출한 자료로 알 수 있는 정보가 매우 많았습니다. 이러한 정보들을 실제로 그 분야의 지식을 활용하며 분석하는 방법을 교수님에게 직접 배울 수 있었고, 네트워크 분야에 대한 추가적인 공부의 필요성을 느낄 수 있었습니다. ▸ [02/19] 1대1 미팅 - 윤명근 교수님 - 데이터 셋을 기관별로 나누어 분석을 진행했습니다. 또한 분석용 데이터 셋을 구축하였으며, 앞으로의 ToDo리스트를 작성하여 이에대한 이야기를 나누었습니다. => 느낀점: 2000만건 이상의 네트워크 패킷 이벤트를 다루다보니 병렬처리등의 여러 프로그래밍 기술을 써야만 했는데, 교수님과 교수님아래 연구생분들의 조언으로 해결할 수 있었고, 앞으로의 코드 또한 스스로의 능력에 안주하지 않고 더 성능이 높은 코드를 위해 열심히 공부하고 조언을 구해야 겠다고 생각했습니다. 또한 데이터 분석에 있어서 시각화, 확장성, 간결함 등이 매우 중요한 요소라는 것을 깨달았습니다. ▸ [02/22] 1대1 미팅 - 윤명근 교수님 - SYN패킷을 활용한 새로운 분석 실험 결과에 대한 미팅 => 특정 소주제에 대한 실험과 분석 보고서를 작성하는 과정에 대한 피드백을 받을 수 있어서 실력 증진에 많은 도움이 됐습니다. ▸ [02/23] 1대1 미팅 - 윤명근 교수님 - 목표사항 1 중 서버 - 클라이언트 분류하는 방법에서 특이 케이스들을 찾아내었고, 이를 분석한 자료에 대한 교수님과의 미팅 => 느낀점: 저는 단지 데이터만을 분석하여 이상 데이터를 찾아냈습니다. 그 이상데이터가 무엇을 의미하는 지 까지는 알 수 없었는데 교수님이 이에 대한 부분의 이유를 하나하나 설명해주셔서 더욱 잘 이해할 수 있었습니다. 또한 제가 분석한 내용에 대해 틀린 점이 있었는데, 이를 토대로 생각 방식에 유연성을 더할 수 있었습니다. ▸ [02/25] 1대1 미팅 - 윤명근 교수님 - 목표사항 1 중 서버 - 클라이언트 분류하는 방법에서 well-known port를 활용한 rule1에 대한 장점 및 단점을 분석한 내용을 토대로 미팅 => 느낀점: 예외 케이스에 너무 집중하여 메인 주제가 진행이 되지 않을 수 있는 문제에 대해 경각심을 깨달을 수 있었습니다. 예외 케이스는 파악만 하고 실제 검증 과정 이후에 처리하는 방식을 도입할 수 있었습니다. ▸ [03/11] 1대1 미팅 - 윤명근 교수님 - 목표사항 1 전체에 대한 실험 진행 및 분석에 대한 미팅을 진행 => 느낀점: 각 목표사항에 대한 목적이 무엇인지, 제가 지금까지 해왔던 실험과 분석에 대한 정도가 적절한 것인지 확인할 수 있었습니다.