

국외 출장 결과보고서

결	담당자	팀장	실장	사업단장
	6월 7일	6월 7일	6월 7일	6월 7일
재	김은진	강전호	권인혁	권인혁

1. 출장개요

출장목적	European Space Agency (ESA)–ECMWF Workshop 2024 참석 및 발표		
출장일정	2024. 5. 6.~2024. 5. 10. (3박 5일)		
성명		소속	직급(직위)
김은진		자료전처리팀	원급연구원
방문기관 (또는 회의장소)	유럽 우주국 (ESA/ESRIN)		
출발지	인천	도착지	로마

2. 업무수행내용

가. 출장 배경 및 목적

- o 4th ESA-ECMWF Workshop on Machine Learning for Earth System Observation and Prediction (ESOP)는 유럽우주국(ESA)과 유럽중기예보센터(ECMWF)에서 개최하는 워크샵으로, 전통적인 ESOP 기술과 머신러닝 및 딥러닝 기술을 융합한 연구에 대해 발표하고 토론을 진행하는 학술 워크샵이다. 본 워크숍에서는 위성 관측, 기상 및 기후 모델, 후처리 모델 개발을 위하여 머신러닝/딥러닝 기술을 접목하고, 기존의 도구와 머신러닝/딥러닝 간의 개발 진행 상황 및 시너지 효과를 논의하는 것을 목표로 한다. 또한 복잡한 알고리즘과 시뮬레이션을 고속으로 실행할 수 있도록 하는 시스템과 전지구 규모에서의 디지털 모델 개발을 중점 주제로 다룰 예정이다.

나. 주요 업무 수행 사항

- ※ 학술회의에서의 연구개발성과 발표 내용, 최신 연구개발 기술동향 습득 내용, 외부 전문가와의 토의 내용 등 기술, 기관 방문 시 회의내용에 대한 기술
- o 마이크로파 위성자료에 해당하는 ATMS 위성자료의 청천 자료동화를 위해 구름탐지 과정을 머신러닝으로 수행한 결과들을 발표하였다. 비지도학습에 해당하는 Gaussian Mixture Model (GMM) 모델의 결과와 한계점, 이후 개발하였던 지도학습에 해당하는 Neural Network (NN) 기반의 Binary Classification Model 적용 결과에 대해 소개하였다. 딥러닝 기반의 모델 구축에 있어 입력 자료에 ATMS window 채널 자료뿐만 아니라, temperature/humidity sounding channel 모두 사용하는

것, 딥러닝 모델 외에 XGBoost 모델을 사용해보는 것을 제안 받았으며, 이는 앞으로의 연구 고도화 및 확장에 기여할 수 있을 것으로 기대된다. 또한, ECMWF에서 발표한 수치모델을 대체한 ML 모델, Hybrid 모델 연구를 비롯하여 ML weather prediction (PanguWeather, FourCastNet, GraphCast, FengWu 등)의 한계점 분석, NN을 이용한 온라인 모델 오차 보정, AutoEncoder 기반의 3D-Var 자료동화, ASCAT & SIF 위성자료 관측연산자 개발, 위성자료를 이용한 특정 현상에 대한 Detection 연구 등 기상 분야에 적용할 수 있는 다양한 인공지능 기술을 접목한 연구를 통해 최신 인공지능 동향을 파악할 수 있었다.

3. 출장성과·시사점 및 향후 업무 활용 계획

가. 출장성과·시사점

- 포스터 발표에서 타기관 연구자들로부터 몇 가지 의견을 제안 받았다. 1) 딥러닝 기반의 ATMS 구름탐지 모델을 개발할 때, window 채널뿐만이 아니라 수증기 채널 자료를 추가하면 수증기 정보를 충분히 줄 수 있을 것. 2) ATMS는 sounder 자료이기 때문에 구름탐지와 관련될 수 있는 모든 자료(온도 채널, 수증기 채널, 위/경도)를 input data로 1년 이상 장기간 학습시키거나, 3) input 자료를 다양하게 넣을 계획이 아니라면 잘 학습하지 못한 부분을 강화시킬 수 있는 "XGBoost"라는 머신러닝 모델을 사용해보는 것을 제안 받았다.

나. 향후 업무 활용 계획

- 향후 ML 기반의 ATMS 구름탐지 모델의 성능을 향상시키거나 다른 관측종의 품질관리를 위한 ML 모델 구축을 위해 제안 받은 다양한 의견을 참고하여 머신러닝 기반의 품질관리기법 기술 연구를 확장시킬 예정이다. 또한, 이번 ESA-ECWMF 워크숍을 통해 접하게 된 선행연구에 사용된 ML 모델, 한계점 및 구축 시 참고해야 할 사항들은 차년도에 계획된 오차 보정과 관측연산자 개발 연구에 필요한 핵심적인 정보가 될 것으로 기대한다. 더 나아가 머신러닝을 이용한 하이브리드 자료동화 시스템 개발 연구의 동향을 파악하여 향후 해당 분야의 ML 연구를 확장시킬 수 있을 것으로 기대된다.