

관측 및 예보 분과 [P-251]

공군 기상 AI 모델 기상 학습 데이터셋 구축

이하룡, 함숙정, 승연수, 강하림, 김기남

공군기상단

공군기상단은 항공자동기상관측장비(AMOS, Automatic Meteorological Observation System) 등 9개 장비 및 기상관측체계를 통해 연간 약 8,800만 장의 방대한 기상관측자료를 기반으로 기상예측 정보를 제공중이며, 정확도 향상을 위해 초단기예·경보체계('19~'20), 안개탐지 및 예측체계('23), 인공지능 기상레이다('23~'24) 등 3종의 AI 기반 기상예측체계를 도입하였다. 그러나 기상관측자료를 AI 학습용 데이터로 활용하기 위한 결측치 제거, 단위 통일, 품질검증 등 일련의 과정 없이 저장 및 관리 중이며, 공군 기상관측정보에는 '군 항공기 운영을 위한 작전기상지원' 측면에서 기상청과는 독립적인 관측정보가 포함되어 독자적인 데이터셋 구축이 필요한 상황이다. 이를 개선하기 위해 공군기상단은 '24년부터 '25년까지「AI 기반 기상예측 모델 데이터셋 구축 사업」을 추진하고 있다.

데이터셋 구축은 기상데이터 수집-정제-가공 및 학습데이터 생성 3단계로 구성되어 수집단계는 분산 저장된 최대 10년치의 AMOS 등 9종의 원시데이터와 관측지점정보를 수집한다. 정제단계는 WMO(World Meteorological Organization, 세계기상기구) 및 대한민국 기상청 기상관측자료 품질관리 기준, NIA(National Information Society Agency, 한국지능정보사회진흥원) 품질관리 가이드라인, 국방분야 AI 학습용 데이터 구축사업 실무가이드, 공군 기상관측자료 특성 등을 고려한 품질관리 기준을 수립하고 시행하여 원천데이터를 생성한다. 마지막으로 원시/원천데이터에 결측 자료 보간, 표준화 등의 전처리과정과 레이블링을 포함 가공단계를 거쳐 학습데이터셋을 생성, 공군 기상 AI 모델에 적용하는 것을 목표로 하고 있다.

공군기상단은 기상 AI 모델 기상 학습 데이터셋 구축을 통해 데이터셋 구축 기간 확대 및 정제 과정 추가를 통한 기상업무용 AI 모델 성능 추가 향상, 추후 기상업무용 AI 모델 개발 시 구축된 데이터셋을 활용한 사업 능력 향상 등의 효과를 기대하고 있다.

Keywords: AI, 기상예측체계, 품질관리, 데이터셋

※ 이 연구는 24~25-F-AI 기반 기상예측 모델 데이터셋 구축 사업(24J052-A)의 지원으로 수행되었습니다.