

[참가자 명단]**2025 날씨 빅데이터 콘테스트**

참가단위	개인 <input checked="" type="checkbox"/> 팀 <input type="checkbox"/>	참가과제	주제1 <input type="checkbox"/>	주제2 <input checked="" type="checkbox"/>
			주제3 <input type="checkbox"/>	
접수번호	250664	소속	서울과학기술대학교	
팀명	핀토레스트	팀원 수	3명 (대표자 포함)	

○ 참가자 정보

※ 필요시 참가자 정보 셀 추가

성명 (대표자)	권대훈	생년월일	01.07.31
전화번호	010-8909-7359	이메일	dnwn7359@naver.com

성명 (팀원1)	이정윤	생년월일	02.06.05
-------------	-----	------	----------

성명 (팀원2)	주영진	생년월일	02.07.25
-------------	-----	------	----------

위와 같은 팀 구성으로 「2025 날씨 빅데이터 콘테스트」에 참가하고자 합니다.

2025년 06월 27일
참가자 대표 : 권대훈 (서명)

[참가 서약서] 2025 날씨 빅데이터 콘테스트

본인/본 단체는 '2025 날씨 빅데이터 콘테스트'를 위한 공모 참가와 관련하여 아래 사항에 대하여 서약합니다.

1. 경진대회 추진과 관련한 행사 일정을 준수합니다.
2. 행사 추진과 관련하여, 행사 진행 요원과 불미스런 일이 발생하지 않도록 상호 협력합니다.
3. 기상청 경진대회 기 수상작 및 유사 분야의 타 경진대회 수상작, 타인/타사의 아이디어 또는 제품 내용 등을 표절, 도용 시 심사대상에서 제외됨에 동의합니다. 또한, 상금이 기 지급되었을 경우에는 즉시 반환하며, 위 사실로 인한 민형사상의 책임질 것을 서약합니다.
4. 공모작 심사 평가방법 및 절차, 최종 결과는 공개되지 않음에 동의합니다.
5. 대회 진행과정에서 습득한 기술정보를 상대방의 사전 서면동의 없이 제3자에게 누설하지 않고 비밀유지에 동의하며, 이를 위반할 경우 민형사상의 책임질 것을 서약합니다.
6. 출품된 내용 일체에 대한 원 저작권은 응모자 본인에게 귀속되나, 저작권자와의 협의를 통해 양도될 수 있음에 동의합니다.
7. 입상작의 경우 대회의 원활한 홍보와 진행, 공동사업 추진 및 대회 목적 달성 등을 위하여 주최·후원 측이 공동 홍보, 배포, 복제, 이용할 수 있음에 동의합니다.
※ 기타 저작권 관련 사항은 문화체육관광부 창작물 공모전 가이드라인 준용
8. 참가자가 경진대회 참여를 위해 활용하는 데이터는 데이터 제공 기관의 이용 조건을 준수하여야 하며, 본 경진대회 이외의 타 목적으로 활용하지 않을 것을 동의합니다. 데이터 무단 이용에 따른 분쟁 발생시, 민형사상의 책임을 질 것을 서약합니다.
9. 기타 참가 유의사항을 읽고 충분히 이해·동의합니다.

2025년 06월 27일

참가자(대표자명) : 권 대 훈 (인)

기타 참가 유의사항

1. 참가 자격에 결격사항 발생 시 자격 자동박탈이 될 수 있으며, 수상금이 지급되었을 경우 회수 조치
2. 본선, 시상, 시연 등에 응하지 않을 경우 수상자 선정에서 제외
3. 제안한 내용이 평가위원단의 최저 심사기준에 미달할 경우, 해당 부문 시상별 수상자가 없을 수 있음
4. 평가 및 수상자 선정 등 주최기관의 결정에 대해 이의 제기 불가
5. 수상 관련 제세공과금(상금액의 4.4%)은 수상자 부담
6. 제출된 공모작에 대한 내용은 접수 및 평가과정에서 비밀 유지

[공모 요약서] 2025 날씨 빅데이터 콘테스트

공 모 분 야	안전방재	참 가 번 호	250664
팀 명	핀토레스트	대 표 자 명	권 대 훈

공 모 제 목	시계열 기반 지역별 기상 정보를 활용한 119 신고 건수 예측을 위한 CatBoost 모델 제안
분석 배경 및 목표	최근 기후 변화로 인한 자연재해의 빈도 및 규모 증가는 인명 및 재산 피해의 심각성을 가중시키고 있으며, 현행 119 신고 대응 방식은 사후적 특성을 지녀 골든타임 확보 및 자원 배분의 비효율성 문제를 내포하고 있다. 본 연구는 이러한 문제의식에 기반하여 기상 요인과 119 신고 건수 간의 유의미한 상관관계를 규명하고, 이를 활용한 지역별 119 신고 건수 예측 모델을 개발하는 것을 목표로 한다. 이는 소방 인력 및 자원의 효율적 배분 및 투입을 통해 재난 피해를 최소화하고, 소방 서비스의 질적 향상 및 국민 안전의 선제적 확보에 기여할 것으로 기대된다.
분석 데이터 및 전처리	본 연구는 기상청의 지역별 기상 데이터와 119 신고 접수 데이터를 활용했습니다. tm(날짜), address(지역 정보), ta_max/min(기온), hm_max/min(습도), ws_max/ins_max(풍속), rn_day(강수량), 그리고 예측 목표인 call_count(119 신고 건수) 등의 변수를 사용했습니다. 데이터 전처리 과정에서는 날짜 정보를 datetime 형식으로 변환하고 요일, 월 등의 주기성 정보를 사인/코사인 변환을 통해 파생했습니다. 또한, 일교차, 습도차, 그리고 폭염 및 폭우 여부를 나타내는 이진 변수들을 생성하여 기상 특성을 강화했습니다. 마지막으로, 지역 변수들은 One-Hot Encoding을 적용했으며, 불균형한 분포를 가진 call_count는 np.log1p 변환을 통해 모델 학습의 안정성을 높였습니다.
분석방법	119 신고 건수 예측을 위해 Gradient Boosting Decision Tree(GBDT) 계열의 CatBoost Regressor 모델을 적용하였다. CatBoost는 다수의 약한 결정 트리를 순차적으로 학습

	<p>시켜 예측 오차를 점진적으로 감소시키는 앙상블 기법을 활용하며, 특히 범주형 변수 처리 및 과적합 방지에 강점을 지닌다. 모델 학습 시에는 iterations 500, learning_rate 0.05, depth 6 등의 최적화된 하이퍼파라미터를 설정하여 RMSE(Root Mean Squared Error) 손실 함수를 최소화하도록 훈련을 수행하였다. 모델을 통해 로그 변환된 신고 건수를 예측한 후에는 역변환을 통해 원래의 신고 건수 스케일로 최종 예측값을 도출하였다.</p>
결과 및 검증	<p>개발된 모델은 신고 건수가 적은 구간에서 비교적 높은 예측 정확도를 나타내었다. 신고 건수 5건 이상인 고신고 구간에서는 실제 값보다 다소 낮게 예측하는 경향이 관찰되었으나, 일부 고신고일 패턴을 성공적으로 학습하여 자원 배치에 활용될 수 있는 잠재력을 확인하였다. 2024년 데이터에 대한 예측 정확도(RMSE)는 1.04로, 이는 기존 검증 성능인 1.2767 대비 우수한 결과로서 모델의 일반화 성능이 안정적으로 유지됨을 시사한다.</p>
활용방안	<p>본 모델을 통해 예측된 119 신고 건수는 소방 자원 운용의 효율성을 제고하고 국민 안전을 강화하는 데 다양하게 기여할 수 있다. 구체적으로, 예측된 신고 증가 지역에 소방 인력 및 장비를 선제적으로 배치하여 긴급 상황 발생 시 골든타임을 확보하고 신속한 초동 대응을 가능하게 한다. 또한, 예측 정보를 활용하여 재난 취약 지역에 대한 예방 활동을 강화하고, 대국민 안전 수칙 홍보를 통해 선제적인 재난 예방 체계를 구축할 수 있다. 더 나아가, 예측 결과는 재난 대응 전략 수립 및 소방관 훈련 프로그램 개발에 활용될 수 있으며, 유관기관과의 협력을 통한 통합적인 재난 대응 체계 구축에도 기여할 수 있다. 궁극적으로 본 모델은 수동적이고 사후적인 대응 방식을 벗어나 능동적인 재난 관리를 실현하는 데 핵심적인 역할을 수행할 것으로 기대된다.</p>