

ARIMA & Prophet 모델비교 1 차

개요

- 현재 예천양수 열화상프로젝트에 ARIMA 모델이 사용되고 있는데 Prophet 모델과 성능을 비교해 추후 개발할 새로운 프로젝트나 데이터 특성별로 더 적합한 모델을 찾기 위함

모델 별 특징

	ARIMA	Prophet
데이터 요구사항	비정상성 데이터는 차분과정이 필요함 데이터가 적어도 사용이 가능함	비정상성 데이터 직접 사용 가능 데이터가 많을수록 예측성능과 속도가 향상됨
설정 난이도	ACF, PACF 분석등을 통해 적절한 p,d,q 값을 찾아야함	자동 조정 기능이 있어 크게 바꿀 파라미터 없음
해석 가능성	모델의 수학적 구조가 명확하여 결과 해석이 용이함	트렌드, 계절성, 이벤트 등의 요소를 시각화하여 직관적인 해석이 가능함

ARIMA 모델 설명

- ARIMA 는 Auto Regressive Intergrated Moving Average 로 시계열 데이터의 패턴을 파악하여 통계적 기법으로 미래값을 예측하는 모델이다. ARIMA 모델 이전 예측모델인 AR 과 MA 를 합친 후 차분하는 과정을 추가한 모델로 이전의 모델들과는 달리 비정상성을 가진 데이터도 차분을 통해 정상성을 가지게 한 후 예측이 가능하게 하였다.
- 구성요소
 - AR(Auto Regressive) : 과거 값들이 현재 값에 미치는 영향
 - I(Intergrated) : 시계열 데이터가 정상성을 가질 수 있게 차분을 적용
 - MA(Moving Average) : 과거 오차들이 현재 값에 미치는 영향
- 파라미터
 - P : 자기회귀(Auto Regressive)되는 항의 수를 의미
이전의 관측 값들이 현재의 값에 얼마나 영향을 미치는지 나타냄
 - D : 차분 차수를 의미
시계열의 비정상성을 제거하고 데이터의 평균이나 분산이 시간에 따라 일정하도록 함
시계열 데이터의 추세(Trend)를 제거하기 위해 사용됨
 - Q : 이동평균(Moving Average)되는 항의 수를 나타냄
예측 값이 예측 오차에 따라 조정되는 정도를 의미하고 시계열데이터의 잡음을
모델링하기 위해서는 Q 값을 설정하여 오차정도를 보정할 수 있음

Prophet 모델 설명

- Prophet 은 Facebook 에서 개발한 시계열 예측 모델로, 결측치가 있거나 계절성이 뚜렷한 데이터를 처리하는데 강점이 있다. 다양한 트렌드 변화를 자동으로 감지가 가능하다.

- Prophet 모델의 주요 분석 성분

- 추세(trend) 모델: 선형 또는 로그 성장 함수를 사용하여 장기적인 데이터 추세를 학습
- 계절성(seasonality) 모델: 주기적인 패턴을 캡처하여 요일별, 월별, 연간 변동성을 모델링
- 휴일(holiday) 효과: 사용자 정의 가능한 휴일 및 이벤트를 반영하여 예측 성능 향상

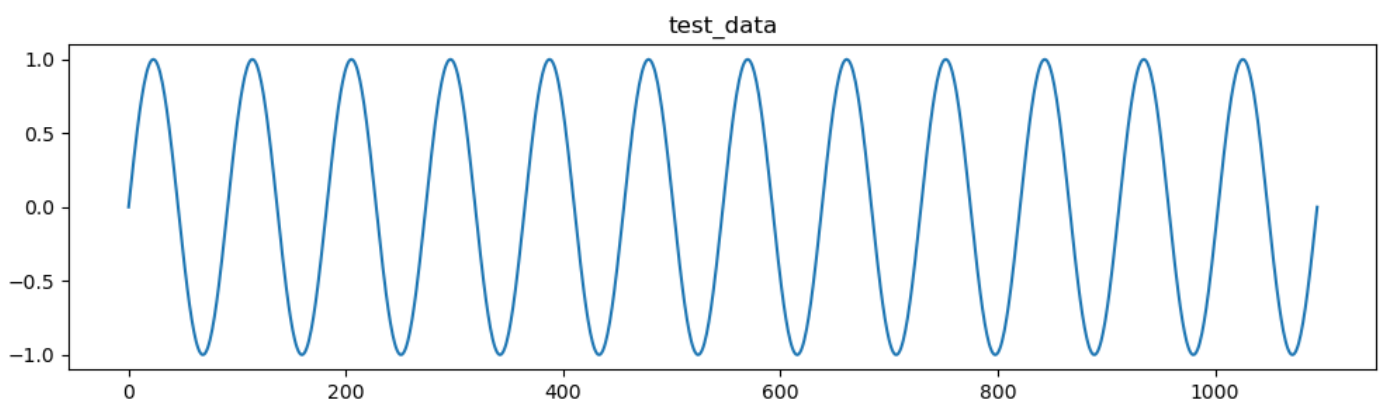
Prophet 은 데이터에 맞게 자동으로 적절한 매개변수를 조정하며, 이상치(outlier)에도 좋은 성능을 보인다.

- 파라미터

- changepoint_prior_scale: 추세 변화를 얼마나 유연하게 허용할지 결정
- weekly_seasonality, yearly_seasonality, daily_seasonality
- seasonality_mode: 계절성을 'additive'(덧셈) 또는 'multiplicative'(곱셈)로 설정
- seasonality_prior_scale: 계절성의 영향을 얼마나 크게 반영할지 조절
- holidays: 특정 이벤트 또는 공휴일을 예측에 반영

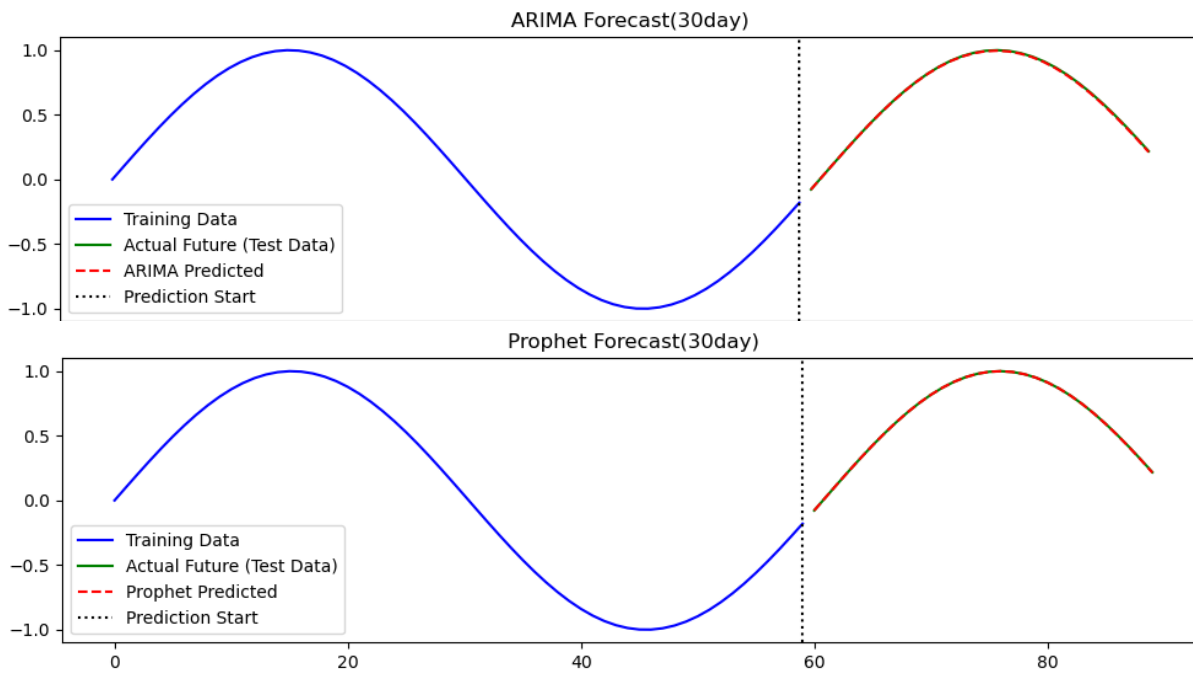
모델 성능 비교

- 모델 성능을 비교하기 위해 임시 데이터로 모델별 특성을 확인하고자 함



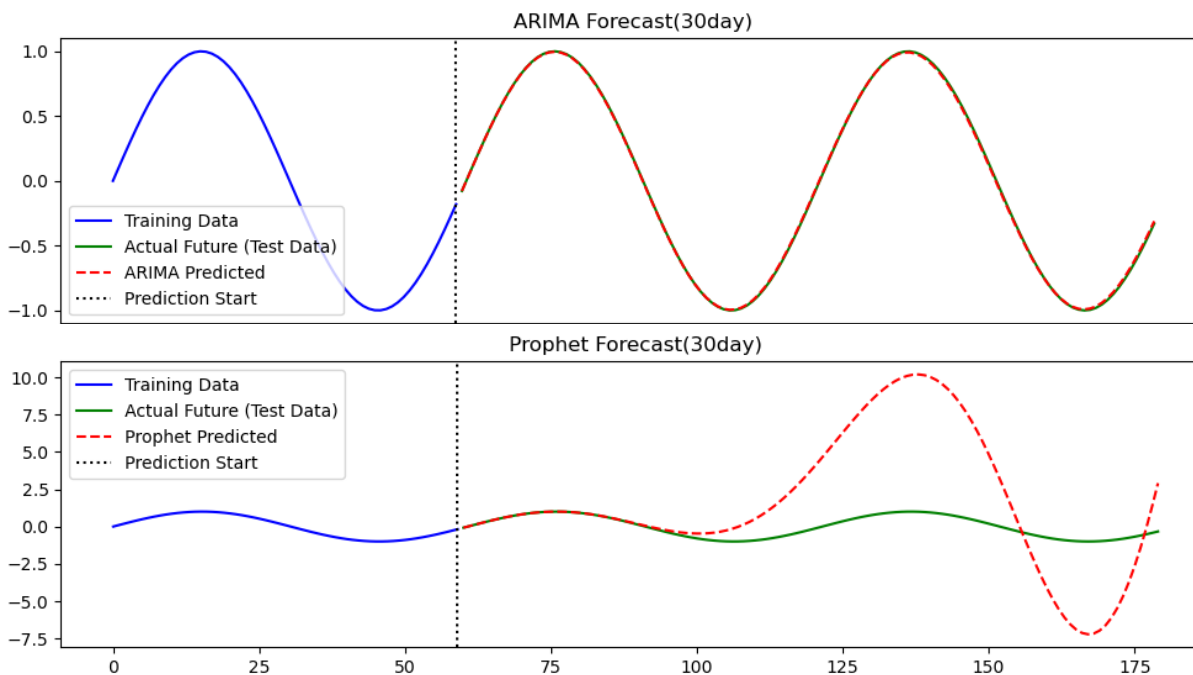
1. 장단기 예측성능 비교
2. 추세유무에 따른 예측성능 비교
3. 노이즈유무에 따른 예측성능 비교

1. 1 시간 데이터로 30 분 예측 (60, 30)



	ARIMA	Prophet
파라미터	p : 10, d : 1, q : 0	yearly_seasonality=True
소요시간	1.28 초	20.64 초

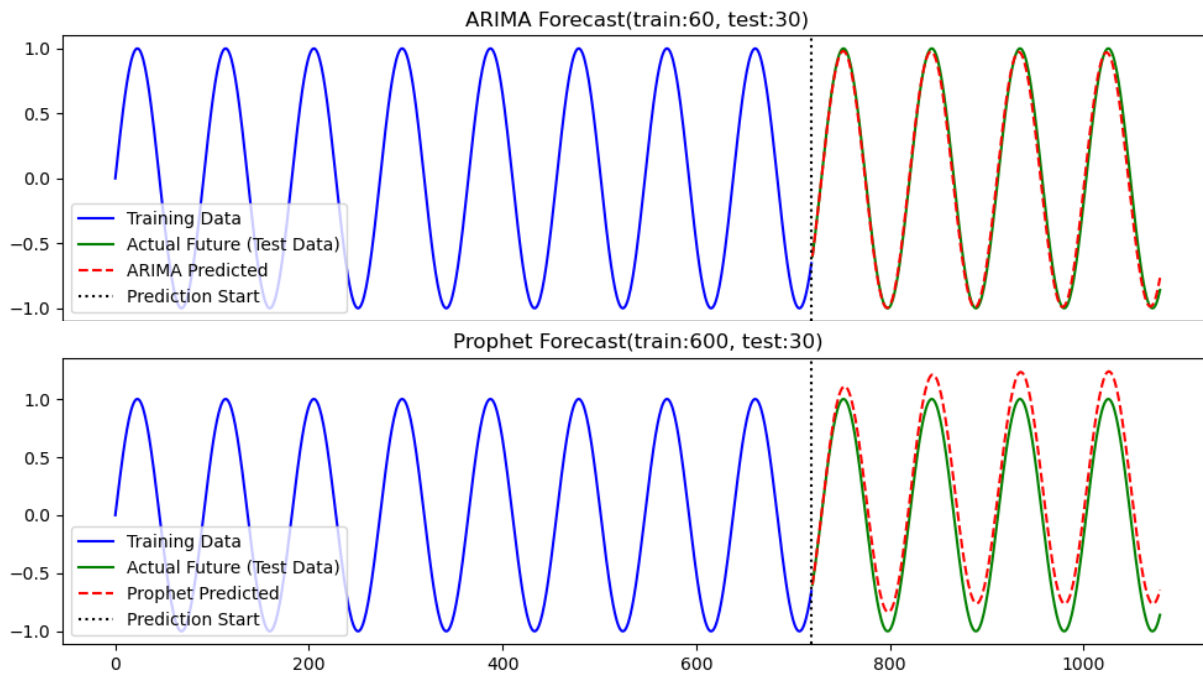
2. 1 시간 데이터로 2 시간 예측 (60, 120)



	ARIMA	Prophet
파라미터	p : 10, d : 1, q : 0	yearly_seasonality=True
소요시간	0.96 초	8.01 초

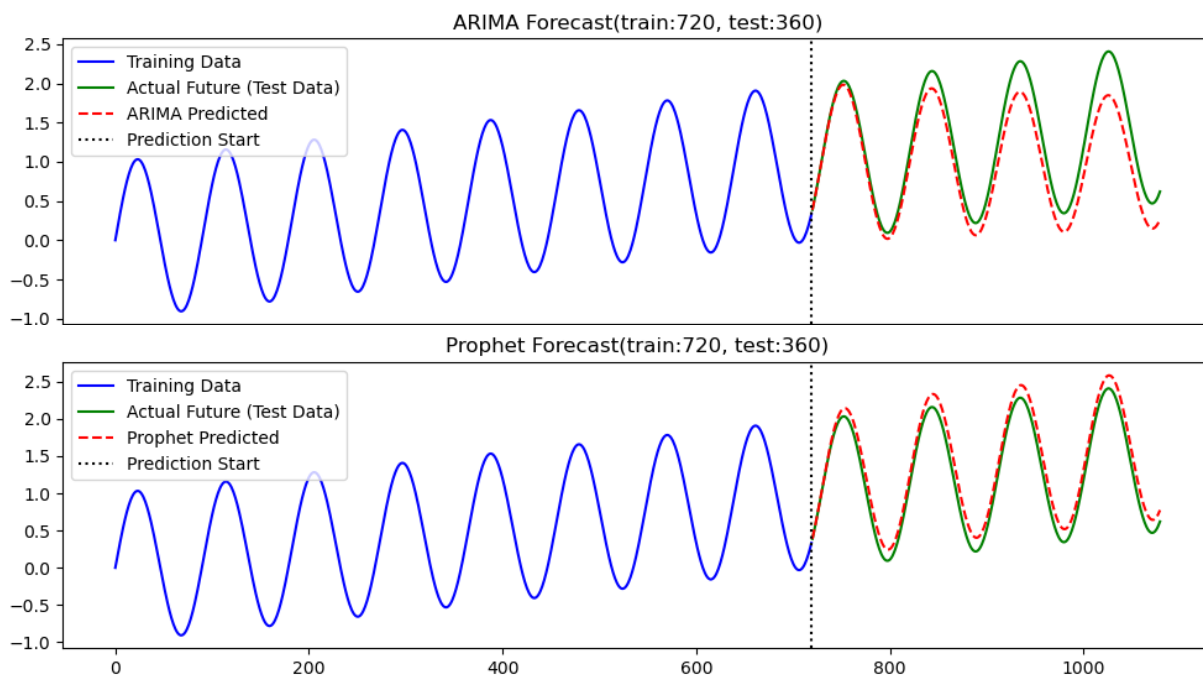
- 단기간의 데이터(60 개)로 미래를 예측한 결과 특정한 패턴이 반복된다면
ARIMA 모델의 성능이 좋은걸 확인할 수 있음 -> 적은 데이터로도 예측가능

- 3. 12 시간의 데이터로 6 시간 예측(720,360) – 추세 없음



	ARIMA	Prophet
파라미터	p : 10, d : 1, q : 0	yearly_seasonality=True
소요시간	1.86 초	0.96 초

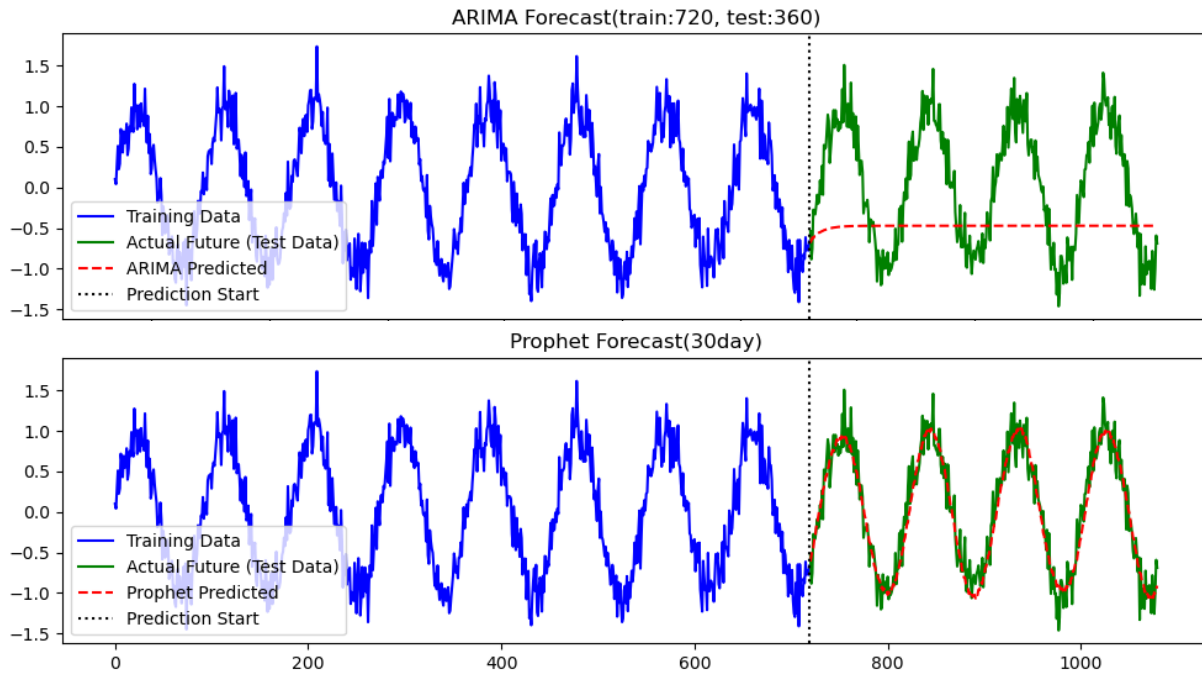
4. 12 시간의 데이터로 6 시간 예측(720,360) – 우상향 하는 추세



	ARIMA	Prophet
파라미터	p : 10, d : 1, q : 0	yearly_seasonality=True
소요시간	3.57 초	0.92 초

- 아무런 추세가 없고 동일한 패턴이 반복되는 경우 ARIMA 의 예측이 정확하지만 추세가 존재할때는 Prophet 은 따라가지만 ARIMA 는 따라가지 못함

5. 노이즈가 있는 12 시간 데이터로 6 시간 예측 (720, 360)



	ARIMA	Prophet
파라미터	p : 10, d : 1, q : 0	yearly_seasonality=True
소요시간	0.59 초	0.52 초

- 노이즈가 있는 데이터의 경우 ARIMA 는 예측을 전혀 하지 못하지만 Prophet 은 추세를 반영하여 예측하는 것을 확인

모델 성능 비교 결과

	ARIMA	Prophet
장점	적은 데이터로 예측가능 예측에 소요되는 평균시간이 전반적으로 빠름	노이즈, 추세를 반영해도 예측이 가능함
단점	학습데이터가 길어진다면 추세를 예측할 수 없음 노이즈가 추가된다면 예측이 불가능해짐	데이터가 적을 때 학습시간이 오래걸림 (데이터가 충분하면 ARIMA 보다 빠름)