**포스터 발표 대본 준비**

1. **Prophet이 무엇인지**

* 시계열 데이터 분석 모델로 현재 데이터를 토대로 미래의 데이터를 예측하는 모델
* This a time series data analysis model that predicts future data based on current data.
* Meta의 전신인 Facebook에서 개발한 모델로 해당 논문을 참고해 모델링 구현
* Model implementation based on the thesis developed by Meta's predecessor, Facebook.
* 날짜와 값 2가지의 데이터를 사용해 분석 가능
* It can be analyzed using two types of data: date and value.
* 적은 양의 데이터에 대해서도 파라미터들을 통해 보정 가능
* It can be calibrated through parameters for even small amounts of data.

1. **어떤 실험을 했는지**

* 실제 커피매장의 일일 매출액을 활용해 데이터 분석 실시
* Data analysis is conducted using the actual daily sales of coffee shops.
* 모델을 구성하는 파라미터인 trend, seasonality, holiday를 반영해 주간 예상 매출액을 추정함
* Estimated weekly sales by reflecting trends, seasonality, and holidays, which are the parameters that make up the model.
* 첫번째 그림을 통해 trend의 파라미터를 조절해 전체적인 데이터의 흐름을 감지할 수 있었고, 빨간 점선들처럼 changepoint를 감지하는 부분은 조절 가능함
* Through the first figure, the overall data flow could be detected by adjusting the trend's parameters, and the part that detects the changepoint, like the red dotted lines, is adjustable.
* 크게 보았을 때 아래로 볼록한 모습의 형태를 보임
* When viewed in a big way, it looks convex downward.
* 두번째 그림에서는 seasonality와 holiday를 반영한 모습
* In the second picture, it reflects seasonality and holiday.
* 먼저 주기성을 보면, 해당 매장은 주말에 매출이 감소하고, 평일에는 매출이 증가하는 모습을 보이는데 그러한 부분을 잘 잡아낸 모습을 보임
* First, if you look at the periodicity, the store shows a decrease in sales on weekends and an increase in sales on weekdays, but it has caught that part well.
* 이전 그림보다는 좀 더 flexible하게 감지를 하는 모습
* It's more flexible than the previous one.
* Changepoint도 분명해졌고, 단순히 오목했던 형태에서 여러 부분에서 굴곡진 이상적인 형태를 보임
* Changepoint has also become clear, and it shows an ideal shape that is curved in many areas from a simple concave shape.
* 시각화를 통해 데이터의 흐름을 잘 잡아내는지 확인해보고, 추후에 평가지표들을 확인해가며 정확도를 검증함
* Visualization checks to see if the flow of data is well captured and checks the evaluation indicators later to verify accuracy.

1. **정확도가 어떻게 되는지**

* 전체적으로 10 ~ 20%의 오차율을 보임
* Overall, it has an error rate of 10 to 20%
* 금액으로 나타내면 3만원에서 많게는 20만원까지 차이를 보임
* In terms of amount, there is a difference from 30,000 won (23$) to 200,000 won (150$)
* 아무래도 매출에 영향을 미치는 다양한 요인들이 있는데, 그러한 부분들을 배제한 게 정확도에 영향을 준 것으로 해석됨
* There are various factors that affect sales, and excluding those parts is interpreted as affecting accuracy.
* 또한 모델들을 검증하는 평가지표에 잘 몰랐다 보니 초기에 이러한 부분을 활용하지 못함
* Also, I was not familiar with the evaluation indicators that verify the models, so I couldn't utilize them in the early stages.

1. **왜 이 모델을 써야 하는지**

* 전체적인 데이터 양이 적더라도 보정할 수 있는 여러 파라미터들이 존재하며, 단순히 날짜열과 데이터열만 이용하더라도 쉽게 분석이 가능
* There are several parameters that can be corrected even if the overall amount of data is small and can be easily analyzed by simply using the date and data columns.
* 학부생 입장에서 가장 접근하기 쉽고, 활용이 가능할 것이라 생각되어 실험
* Experiment because it is thought to be the most accessible and applicable from the perspective of an undergraduate student.
* 또한 전문적인 지식이 없는 사람도 쉽게 배우며 활용이 가능하기 때문
* Also, even people without professional knowledge can easily learn and utilize it.
* 이유는 메타에서 기본으로 제공하는 방법도 있으며, 기존에 활용되고 있는 정보도 많이 제공되기 때문
* The reason is that there is a basic method provided by Meta, and there is a lot of information that is currently being used.

1. **전제적으로 설명하는 부분**

* 앞선 내용들을 정리해가며 얘기하면 될 듯

1. **매장별로 차이가 나는 이유**

* 매출에는 유동인구, 거주인구, 교통 인프라 등 여러가지 요소를 고려해야 하는데 그러한 부분이 배제되어 있기 때문에 매장별로 차이가 발생
* Sales require consideration of various factors such as floating population, resident population, and transportation infrastructure, but such parts are excluded, resulting in differences between stores.
* 해당 부분을 고려해서 반영한다면 좀 더 정확도 높은 모델을 구현할 수 있을 것으로 예상됨
* A more accurate model is expected to be implemented if this is considered and reflected.

1. **다른 모델에서 변환하는 것에 대해 부족한 점은 없는지**

* 사실 이 부분에 있어서 해당 논문이 많이 부족한 부분임
* In fact, the paper is lacking a lot in this area.
* 시계열 데이터 분석에는 RNN이나 ALIMA 같은 다양한 모델이 존재하는데, 해당 모델들을 통해 구현할 수 있는 부분이나 본 논문에서 다루고 있는 prophet과의 차이점을 구체적인 통계를 통해 나타내지 못함.
* There are various models such as RNN and ALIMA in time series data analysis, but specific statistics do not show the difference between the parts that can be implemented through these models, or the profile covered in this paper.
* 이러한 비교는 향후 프로젝트 진행에 반영하도록 할 예정
* These comparisons will be reflected in future projects.