

Project 예상질문

개인적 PPT 수정사항

플라스틱 몰드(상,하형 몰드)가 변형이 된다고 되어있는데, 그림 상에서는 상,하형 몰드가 금속부분을 가리키고 있음. 정의를 제대로 하여주어야 혼란이 없을 것이라고 생각.

01 Concept

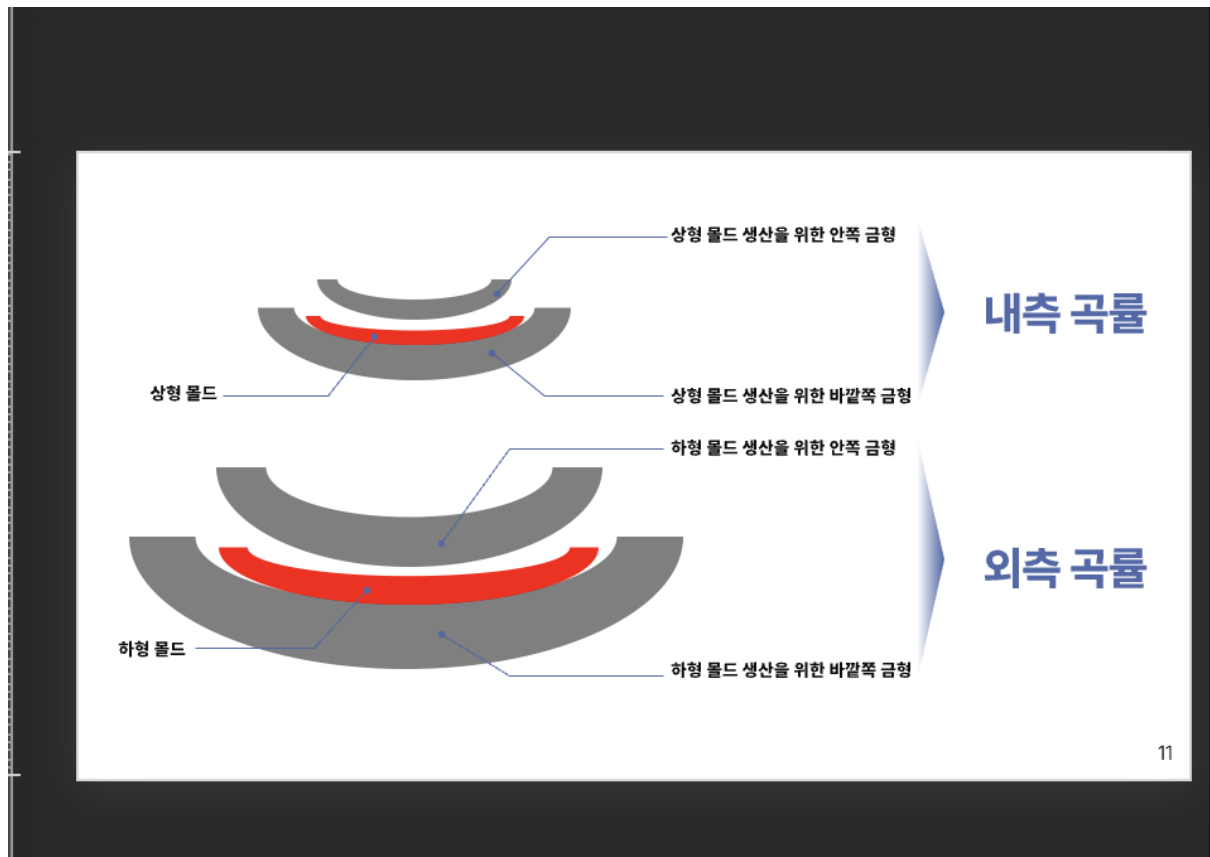


위의 그림과 아래 그림과 설명이 좀 다름..(아래의 그림은 몰드가 플라스틱을 가리키고 있음)
아래의 그림도 상형몰드와 하형몰드가 가리키는 것이 동일한 플라스틱 몰드를 가리키는 것 같아서 혼란이 있는 것 같다고 생각.

상형 몰드, 하형 몰드를 따로따로 생산하는 것이 아니라 한번에 생산하는 것으로 알고 있음
(잘못되었다면 알려주세요....)

아래의 그림을 보면 2개를 따로 생산하는 것 같이 표현이 되어있음.

플라스틱 몰드의 윗부분을 내측곡률이라고 표현하고, 아래쪽을 외측곡률 이런식으로 하나의 그림에 표현하여주면 이해가 더 잘 될 것 같습니다.



∴ 상형 몰드, 하형 몰드, 플라스틱 몰드, 금형 등의 정의를 정확하게 내려주는 것이 필요하다고 생각이 됨. 내측곡률 및 외측곡률을 하나의 그림에 표현이 가능하면 더 좋을 것 같습니다.

질문

1. 내측곡률 및 외측곡률이 가장 몰드의 곡률에 영향을 준다고 하였는데, 설비번호와 금형 위치는 왜 필요한 특성인지??
 - a. 해당 설비번호와 금형위치에 따라 같은 곡률의 금형을 넣어도 다른 도수가 생산될 수 있기 때문에 오차를 최대한 줄이기 위하여 특성으로 사용
2. 3가지 모델에 대한 성능이 좋지 않게 나온다면, 어떠한 방법을 사용할 것인지?
 - a. 사용하지 않은 특성을 사용 or Decision Tree를 앙상블하는 방법을 사용한 Random Forest와 같은 다른 모델들도 팀원들과 스터디 후 사용해볼 예정
3. 파라미터의 튜닝과정은 어떻게 할 것 인지?
 - a. Logistic Regression의 규제 값인 C, Decision Tree의 max depth, SVM의 kernel, 규제강도인 C와 gamma값을 Grid Search를 통해 validaion set에 대해

최적의 F1 score를 가지는 파라미터를 찾을 예정

4. Train-Validaion-Test 비율은?

a. 8:1:1

5. 이상치의 분포는 어떻게 됨? 산업에서 불량품의 비율은 상당히 적은 것으로 알고있는데 데이터가 불균형하면 지도학습이 제대로 진행되지 않을 확률이 큰데?

a. 아직 분석은 거치지 않았습니다.

b. 데이터 분석을 진행해본결과 심각한 불균형을 이루고 있습니다. → 지도학습을 통해 적절한 성능이 도출되는지 확인해볼 예정

만약 성능이 도출되지 않는다면 비지도학습과 같은 방법을 도입해볼 예정입니다.