



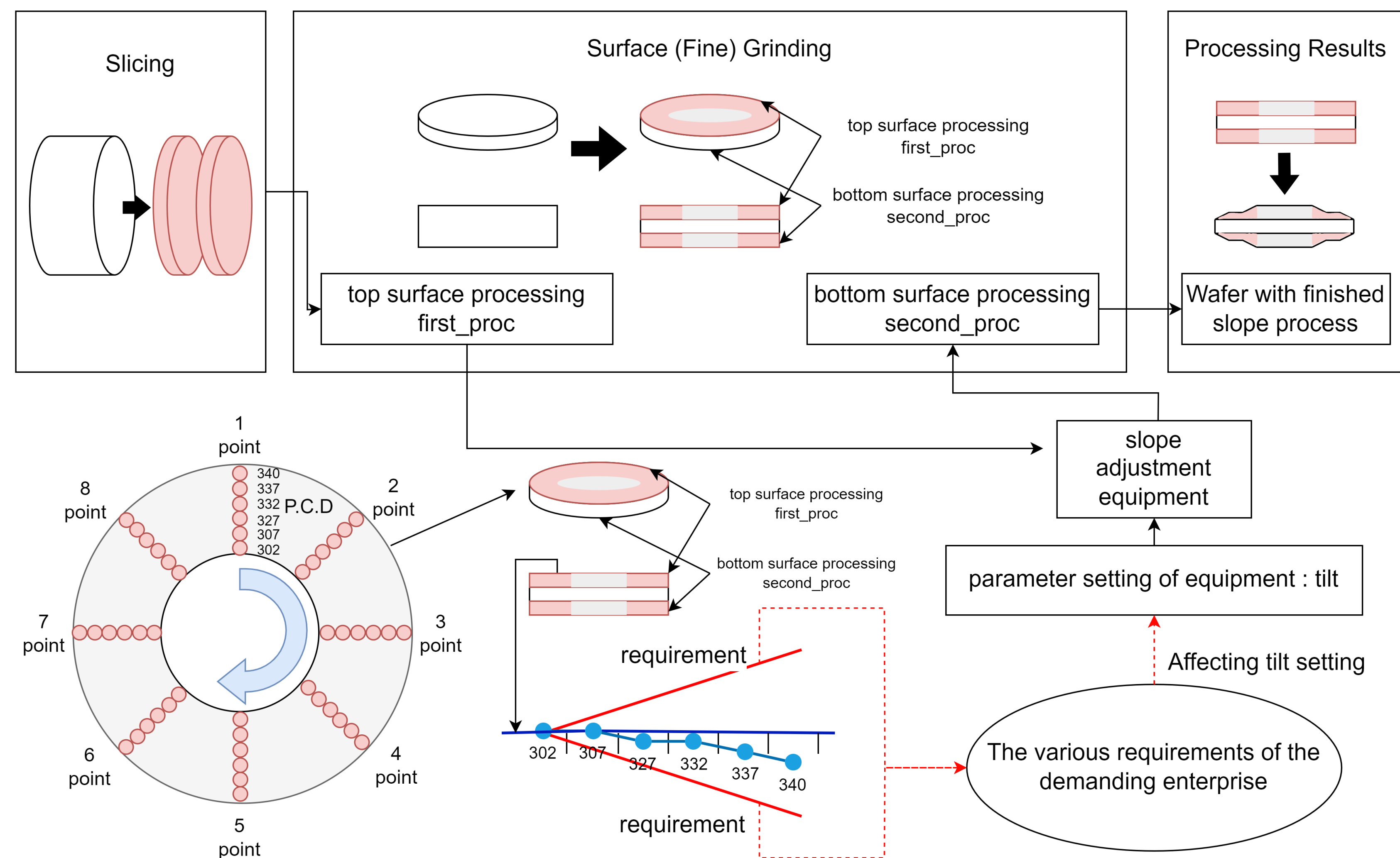
빅데이터 기반 웨이퍼 제조 과정의 불량 관련 주요 인자 분석 연구

길준혁 (한국기술교육대학교), 배장원 (한국기술교육대학교)

배경 및 목적

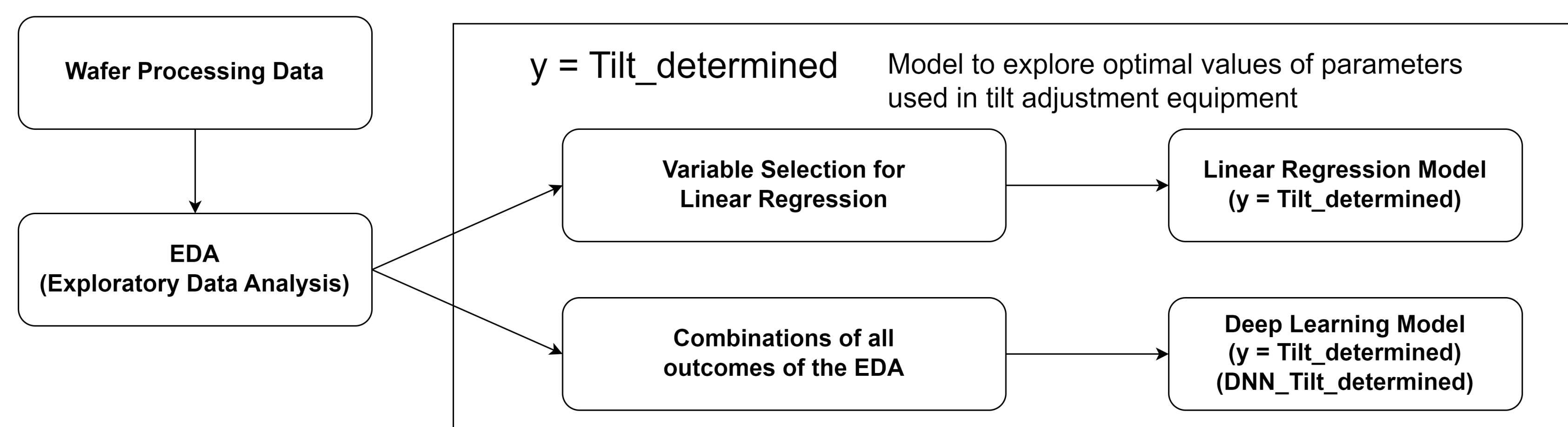
현대 사회는 모바일 기기의 고성능화와 IT 분야의 급속 성장 속에서 반도체에 대한 다양한 수요가 증가하고 있음. 특히 웨이퍼 연마 프로세스 기술 뿐만 아니라 웨이퍼의 평탄도 및 기물기에 대한 정밀한 요구사항을 반영할 수 있는 능력이 주요한 부분으로 떠오르고 있음. 본 연구는 반도체 제조 관련 빅데이터를 이용하여 수요 기업의 다양한 요구사항에 능동적으로 대응할 수 있는 인사이트를 도출하여 불량률을 낮추는 것을 목적으로 함.

웨이퍼 두께 가공 공정



- 4가지 종류의 제품을 대상으로 웨이퍼 제조 공정 중 웨이퍼의 두께를 조절할 수 있는 표면 연삭(Surface Grinding) 과정을 고려.
- 공정 과정
 - 1차 공정(first_proc): 웨이퍼의 상면에 대한 가공(top surface processing)으로 가공 이후 웨이퍼의 각 PCD마다 두께를 측정.
 - ※ PCD(Pitch Circle Diameter)란 웨이퍼의 중앙에서 외곽으로 이동하면서 두께 값을 측정할 때 측정 구간의 명칭으로, 제품에 따라 개수 및 기물기 요구사항이 다름.
 - 이를 기반으로 웨이퍼에 대한 정밀한 요구사항을 달성하기 위해 기물기 조절 연삭 장비의 파라미터값 (본 연구에서는 이를 tilt라 칭함)을 적절히 설정.
 - 2차 공정(second_proc): 웨이퍼의 하면에 대한 가공(bottom surface processing)으로 최종 웨이퍼 제조.

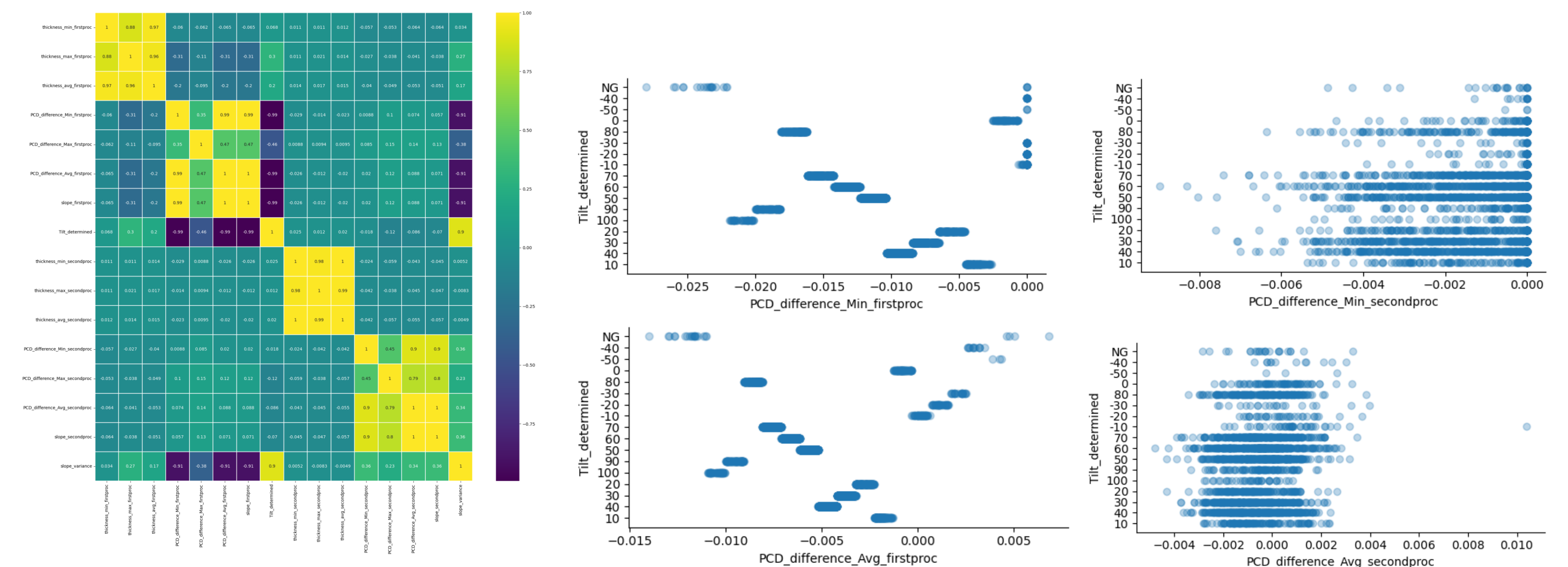
전체 분석 프로세스



종속변수 (Tilt_determined): 최적의 tilt 결정값.

- 반도체 제조 과정에서 수집한 데이터 전처리.
 - 웨이퍼 기물기와 관련 있는 주요 요인 도출을 위한 EDA.
 - heatmap: 상관계수 0.5 이상 / scatterplot: 비교적 명확히 구분되는 변수
- 1) 종속변수를 잘 표현하는 변수 조합 탐색을 위해 변수 선택법 후 선형 회귀 모델 사용.
2) 새로운 변수 조합의 존재 가능성 때문에 EDA 결과의 모든 조합으로 딥러닝 모델 사용.

방법론 및 결과



	Data Size	Data Split (8:2)		y = Tilt_determined		
		Train	Test	EDA (Exploratory Data Analysis)	Variable Selection	AIC
A	3,086	2,468	618	slope_firstproc slope_variance PCD_difference_Min_firstproc PCD_difference_Max_firstproc PCD_difference_Avg_firstproc	slope_firstproc slope_variance PCD_difference_Avg_firstproc	12220.03
B	2,516	2,012	504	slope_firstproc slope_variance PCD_difference_Min_firstproc PCD_difference_Avg_firstproc	PCD_difference_Avg_firstproc	9962.66
C	487	389	98	slope_firstproc slope_variance PCD_difference_Min_firstproc PCD_difference_Max_firstproc PCD_difference_Avg_firstproc	slope_firstproc slope_variance	1944.99
D	2,656	2,124	532	slope_firstproc slope_variance PCD_difference_Min_firstproc PCD_difference_Max_firstproc PCD_difference_Avg_firstproc	PCD_difference_Max_firstproc PCD_difference_Avg_firstproc	10635.43

	y = Tilt_determined					
	Linear Regression (LR)		DNN (DNN_Tilt_determined)			
	Best Variables Combination	RMSE	Variables used in LR	RMSE	Best Variables Combination	RMSE
A	slope_firstproc slope_variance PCD_difference_Avg_firstproc	2.792	slope_firstproc slope_variance PCD_difference_Avg_firstproc	2.883	slope_firstproc slope_variance PCD_difference_Min_firstproc PCD_difference_Avg_firstproc	2.849
B	PCD_difference_Max_firstproc PCD_difference_Avg_firstproc	2.914	PCD_difference_Max_firstproc PCD_difference_Avg_firstproc	3.027	PCD_difference_Avg_firstproc	3.012
C	slope_firstproc slope_variance PCD_difference_Avg_firstproc	2.788	slope_firstproc slope_variance PCD_difference_Avg_firstproc	2.752	slope_variance PCD_difference_Max_firstproc PCD_difference_Avg_firstproc	2.651
D	PCD_difference_Avg_firstproc	2.883	PCD_difference_Avg_firstproc	3.025	slope_firstproc PCD_difference_Max_firstproc	2.940
	slope_firstproc slope_variance		slope_firstproc slope_variance	3.054		

- 선형 회귀모델을 통해 도출한 변수 조합을 딥러닝 모델에 적용하여 RMSE값을 산출
- 제품별로 유의미한 변수 조합이 동일하지 않은 경우가 많음
- RMSE값은 선형 회귀모델에서 대체로 더 좋은 값을 보였지만, 딥러닝 모델로 사용될 경우에는 그렇지 않았음
- 제품 혹은 분석 방법에 상관 없이 대부분의 조합에서 공통적으로 포함되는 특정 변수가 있음 (ex. PCD_difference_Avg_firstproc)
- 특정 제품에서만 주요하게 판단된 변수도 있음 (ex. PCD_difference_Max_firstproc)

결론

- 제품별로 종속변수에 영향을 주는 요인들의 조합과 영향력이 다름.
 - 제품 특성에 따라 조합되는 변수들이 다르기에 요구사항별로 심도있는 연구 및 분석이 필요하며, 각 제품에 맞춤형 공정이 필요함.
 - 공정 프로세스의 특성으로 인한 중요 변수가 존재할 수 있음.
- 두 번째는 분석 방법에 따라 주요 인자로 선택된 조합이 다르게 도출됨.
 - 특정 분석 방법이 더욱 좋다고 판단하기는 어려움.
 - 다양한 방법을 통해 유의미한 변수 조합들을 도출하고 여러 상황을 고려한 조합을 사용하는 것이 필요함.
- 웨이퍼 제조 의사결정 상황에서 유용하고 실용적인 정보를 제공할 수 있음.