

# 머신러닝을 이용한 반도체 웨이퍼 평탄화 공정품질 예측 및 해석 모형 개발

Predicting and Interpreting Quality of CMP Process for Semiconductor  
Wafers Using Machine Learning

안정언·정재윤<sup>†</sup>

경희대학교 산업경영공학과

## 요 약

반도체 웨이퍼의 표면을 연마하여 평탄화하는 Chemical Mechanical Planarization(CMP) 공정은 다양한 화학물질과 물리적인 기계장치에 의한 작용을 받기 때문에 공정을 안정적으로 관리하기 힘들다. CMP 공정에서 품질 지표로는 Material Removal Rate(MRR)를 많이 사용하고, CMP 공정의 안정적 관리를 위해서는 MRR을 예측하는 것이 중요하다. 본 연구에서는 머신러닝 기법들을 이용하여 CMP 공정에서 수집된 시계열 센서 데이터를 분석하여 MRR을 예측하는 모형과 공정 품질을 해석하기 위한 분류 모형을 개발한다. 나아가 분류 결과를 분석하여, CMP 공정 품질에 영향을 미치는 유의미한 변수를 파악하고 고품질을 유지하기 위한 공정 조건을 설명한다.

■ 중심어 : 머신러닝, 예측, 분류, 공정 품질 해석, CMP, 웨이퍼

## Abstract

Chemical Mechanical Planarization (CMP) process that planarizes semiconductor wafer's surface by polishing is difficult to manage reliably since it is under various chemicals and physical machinery. In CMP process, Material Removal Rate (MRR) is often used for a quality indicator, and it is important to predict MRR in managing CMP process stably. In this study, we introduce prediction models using machine learning techniques of analyzing time-series sensor data collected in CMP process, and the classification models that are used to interpret process quality conditions. In addition, we find meaningful variables affecting process quality and explain process variables' conditions to keep process quality high by analyzing classification result.

■ Keyword : Machine Learning, Prediction, Classification, Process Quality Interpretation, CMP, Wafer

## I. 서론

최근 제조 공장에서는 기업 경쟁력을 유지하고 생산성을 높게 유지하기 위해 품질을 주요한 화두로 삼아 왔다. 품질은 기업 이익뿐 아니라 기업 이미지에 방대한 영향을 끼치기 때문이다. 이를 위해 스마트 팩토리에서는 IT (Information Technology) 시스템과 빅데이터, 인공지능 기술을 활용하고, 머신러닝 및 딥러닝을 사용하여 품질을 예측하고 해석하며 품질을 관리한다.

국내 제조업 분야 중에서는 반도체 시장이 지속적인 성장을 이루면서 적층 구조를 갖는 고집적 반도체 생산의 필요성이 커지게 되었다. 이러한 필요성에 의해 웨이퍼를 평탄화하는 CMP 공정이 도입되었고 그 중요성이 증가하는 상황이다 [1]. CMP 공정은 물리·화학적 작용을 받기 때문에 다양한 공정 변수가 존재하고[2], 이 변수들로 인해 CMP 공정을 안정적으로 관리하기 어려워 CMP 공정의 품질을 예측하는 것은 중요한 문제이다.

단순히 품질을 예측하는 문제뿐만 아니라 높은 품질을 유지하기 위한 공정 변수의 조건을 파악하여 낮은 품질을 달성하기 전에 미리 방지하는 문제 또한 중요해지는 추세이다. 이를 위해 공정 데이터를 머신러닝을 기반으로 하는 분류 기법 중 트리 기반의 모형들을 주로 사용한다.

본 연구에서는 시계열 센서 데이터로부터 통계적 지표들로 특징을 추출한 후, 머신러닝 기법들을 이용하여 CMP 공정에서 품질 지표로 사용되는 MRR을 예측하고 높은 MRR을 달성하기 위한 공정 변수의 조건을 해석하는 방법을 소개한다. 2016 PHM(Prognostics and Health Management) Data Challenge에서 제공된 반도체 CMP 공정 데이터를 대상으로 제시된 방법을 적용 실험하였다.

본 논문의 구성은 다음과 같다. 2장에서는 MRR을 예측하는 방법론에 관한 관련 연구를 소개한다. 3장에서는 본 논문에서 사용한 데이터를

소개한다. 4장에서는 MRR 예측 및 분류를 하기 위한 실험 설계를 설명하고, 5장에서는 실험 결과를 기술한다. 6장에서는 분류 모형의 결과를 분석하여 공정 품질을 해석하는 방법론을 소개하고, 7장에서 결론을 맺는다.

## II. 관련 연구

기존의 MRR을 예측하는 모델링 방법은 크게 두 가지로, 물리적 수식 기반 모델과 데이터 기반 모델이 있다.

물리적 수식 기반 모델로는 Preston 방정식을 사용하여 MRR을 예측한 연구 사례가 있다[1]. 이외에 운동학적 분석을 사용하여 수학적 모델을 구축한 연구도 있다[2]. 뿐만 아니라 CMP 공정에서 polishing pad wear의 조건은 어떠한 영향을 끼치는지 연구하기 위해 분석적 모델을 구축한 연구도 있다[3].

데이터 기반 모델에 관한 연구는 의사 결정 나무 기반의 앙상블(Decision Tree-Based Ensemble) 학습을 통해 본 연구와 동일한 CMP 공정 데이터를 사용하여 MRR을 예측한 연구가 있다[2]. 이 연구에서는 네 가지 통계 지표인 표준편차, central moment, skewness, kurtosis를 사용하여 시계열 데이터로부터 특징을 추출하였다. 머신러닝 기법으로는 Gradient Boosting Tree(GBT), Extremely Randomized Tree(ERT), Random Forest를 사용하였다. 또 다른 데이터 기반 모델에 관한 연구로는 딥러닝을 사용한 연구가 있다[3]. 이 연구에서는 MRR을 기준으로 값이 낮고 높은 그룹으로 웨이퍼를 나누고, 각 그룹에 대해 Deep Belief Network(DBN)을 사용하여 MRR을 예측한 사례도 있다[3].

또한, 물리적 기반 모델과 데이터 기반 모델을 결합하여 MRR을 예측한 연구가 있다[4]. 이 연구에서는 시계열 데이터에서 특징을 추출하기 위해 한 가지 통계 지표로 평균을 사용하였다. 물리

적 수식 모델로는 Preston 방정식을 사용하였고, Preson 방정식에 포함된 변수 중 사용된 데이터 안에서 구할 수 없는 값을 머신러닝을 사용해 유추하고, 유추된 값을 사용하여 다시 MRR을 예측하는 구조를 갖는다.

### III. 데이터 소개

#### 3.1 CMP 공정

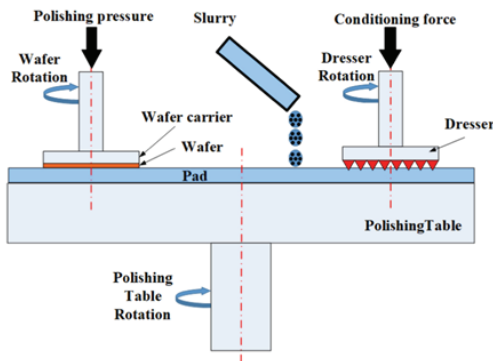
CMP 공정은 연마 공정을 거쳐 웨이퍼의 표면으로부터 물질을 제거하는 웨이퍼 평탄화 공정이다. 이 과정에서 <그림 1>처럼 Pad, Wafer Carrier와 같은 다양한 물리적 기계장치의 압력과 slurry라는 화학물질에 의한 작용을 받기 때문에 공정을 안정적으로 관리하는 것이 어렵다. CMP 공정의 품질은 MRR로 측정되며, 높은 MRR을 달성

하고 낮은 불균일성을 유지하는 것이 CMP 공정의 목표이다. 주 품질 지표인 MRR은 물질 제거율을 의미한다. 웨이퍼가 평탄화되기 전과 후의 두께 차이로 측정된다.

#### 3.2 2016 PHM Data Challenge

PHM Data Challenge는 PHM Society에서 매년 컨퍼런스에 참가할 가능성이 있는 사람들이 참가할 수 있도록 주최하는 데이터 분석 경진대회이다.

본 연구에서는 2016 PHM Data Challenge에서 제공된 CMP 공정 데이터를 사용하였다. Training, Validation, Test set 세 개로 분리되어 제공되며, 이 데이터는 A와 B 두 개의 stage에 해당하는 2829 개의 웨이퍼에 대해 초 단위로 관측된 값이다. 데이터 구성은 <표 1>과 같다. 각각의 웨이퍼는 웨이퍼 ID를 비롯하여 dresser, chamber 등의 공정 변수와 retainer ring 압력 등 총 25개 변수를 포함하고 있고 데이터의 변수와 이에 대한 설명은 <표 2>에 있다.



<그림 1> 웨이퍼 평탄화(CMP) 공정[5]

### IV. 실험 설계

#### 4.1 데이터 전처리

본 연구에서는 예측 모형과 분류 모형을 사용하기 때문에 각 모형에 맞는 데이터 전처리 과정이 필요하다.

예측 모형에 사용되는 데이터를 만드는 과정은

<표 1> 2016 PHM Data Challenge 데이터 구성

	Training	Validation	Test
Total number of observations	672,744	144,148	156,262
Number of wafers	1,981	424	424
Number of wafers under stage A	1,166	252	238
Number of observations under stage A	376,859	82,984	91,798
Number of wafers under stage B	815	172	186
Number of observations under stage B	295,885	61,164	64,464

<표 2> 2016 PHM Data Challenge 데이터 변수 및 변수 설명

MACHINE_ID	Numeric ID of machine	RETAINER_RING_PRESSURE	Pressure related to wafer placement
MACHINE_DATA	Numeric ID of wafer ring location in machine	RIPPLE_AIR_BAG_PRESSURE	Pressure related to wafer placement
TIMESTAMP	Seconds	USAGE_OF_MEMBRANE	A usage measure of polishing membrane
WAFER_ID	Number representing ID of wafer	USAGE_OF_PRESSURIZED_SHEET	A usage measure of wafer carrier flexible sheet
STAGE	A or B representing a different type of processing stage	SLURRY_FLOW_LINE_A	Flow rate of slurry type A
CHAMBER	Chamber in machine for wafer processing	SLURRY_FLOW_LINE_B	Flow rate of slurry type B
USAGE_OF_BACKING_FILM	A usage measure of polish-pad backing film	SLURRY_FLOW_LINE_C	Flow rate of slurry type C
USAGE_OF_DRESSER	A usage measure of dresser	WAFER_ROTATION	Rotation rate of wafer
USAGE_OF_POLISHING_TABLE	A usage measure of polishing table	STAGE_ROTATION	Rotation rate of stage
USAGE_OF_DRESSER_TABLE	A usage measure of dresser table	HEAD_ROTATION	Rotation rate of head
PRESSURIZED_CHAMBER_PRESSURE	Chamber pressure	DRESSING_WATER_STATUS	Status of dressing water
MAIN_OUTER_AIR_BAG_PRESSURE	Pressure related to wafer placement	EDGE_AIR_BAG_PRESSURE	Pressure of bag on edge of wafer
CEBTER_AIR_BAG_PRESSURE	Pressure related to wafer placement	AVG_REMOVAL_RATE	The average rate of material removal (the measure to predict)

다음과 같다. 먼저 training, validation, test set으로 나뉘어 있는 데이터를 통합하고, A stage, B stage에 해당하는 관측치로 분리해서 두 개의 데이터를 생성하였다. 두 데이터에 대해 MRR 예측 시 사용하지 않는 MACHINE\_ID, MACHINE\_DATA, STAGE, CHAMBER 네 개의 변수를 제거하였다. 다음으로 TIMESTAMP, WAFER\_ID 외 제거되지 않은 19가지 변수에 대해 10가지 통계적 지표로 평균, 분산 (variation), 표준편차, 중간값, 왜도(skewness), 최댓값(max), 최솟값(min), RMS(Root Mean Square), P2P(Peak-to-Peak), crest factor를 사용하여 시계열 특징을 추출하였다. 이 때 사용된 통계 지표에 대한 설명은 <표 3>이다. 각각의 웨이퍼에 대해 추출된 190개 변수를 모델의 입력으로 사용하였고, MRR이 매우 높은 네 개의 웨이퍼를 이상치로 제거하는 과정을 거쳤다[3]. 마지막으로 Min-Max scaling을 사용한 정규화 과정을 거친 후에, Training set과 Test set을 4:1의 비율로 나누었다.

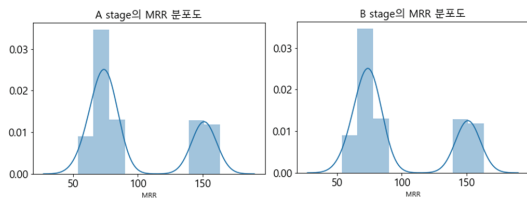
분류 모형에 사용된 데이터를 만들기 위해 데

이터 재구성 과정을 거친다. 예측 모형에 필요한 데이터를 만드는 과정과 마찬가지로 training, validation, test set으로 분리된 데이터를 통합하고, 사용되지 않는 변수를 제외하고 10가지 통계 지표를 적용하였다. 예측 모형에 필요한 데이터를 만들 때 정규화 과정을 거친 것과 달리 분류 모형에 필요한 데이터를 구성할 때에는 정규화 과정을 거치지 않았다. 그 이유는 추후에 이 모델을 사용하여 공정 변수의 조건에 대한 해석을 진행하기 위함이다. 그 후, MRR을 기준으로 분류에 알맞은 데이터를 구성하기 위해 MRR 분포도를 stage마다 그렸고, 그 결과는 <그림 2>이다. A stage에 속하는 모든 웨이퍼는 MRR이 55-100nm, 140-165nm인 그룹으로 양분된다. 이와 달리 B stage에 속하는 모든 웨이퍼는 MRR이 55-165nm에 해당하는 한 그룹으로 나타나기 때문에 A stage에 속하는 웨이퍼를 분류 모델에 적합한 데이터로 선별하였고, MRR이 55-100nm인 웨이퍼는 class 0, 140-165nm인 웨이퍼는 class 1로 범주

를 두었다. 이 과정을 거쳐 생성된 분류 모델에 사용할 최종 데이터 구성은 <표 4>와 같다.

<표 3> 통계 지표

통계 지표	수식
skewness	$x_{skew} = \frac{1}{n\sigma^3} \sum_{i=1}^n x_i$
Root Mean Square	$x_{rms} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$
P2P	$x_{p2p} = x_{max} - x_{min}$
crest factor	$x_{crest} = \frac{x_{p2p}}{x_s}$



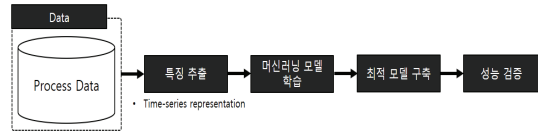
<그림 2> A, B stage의 MRR 분포도

<표 4> 분류 모형에 사용되는 데이터 구성도

MRR	웨이퍼 개수	분류
55-100nm	1148	0
140-165nm	503	1

### 4.2 품질 예측 모형

본 연구에서는 <그림 3>과 같이 공정품질을 예측하기 위한 머신러닝 적용 과정을 제시하였다. 먼저 4.1절에서 설명한 데이터 전처리 과정 중에서 10가지 통계 지표를 사용하여 시계열 특징을 추출한다. 추출된 특징들로 구성된 Training set을 사용하여 머신러닝 모델을 학습한다. 학습된 모델 내의 하이퍼파라미터를 튜닝하여 최적 모델을 구축한 후, Test set을 사용하여 성능을 검증하는 단계를 진행하게 된다.



<그림 3> 예측 모형 적용 절차

전처리 과정을 거친 데이터를 머신러닝을 기반으로 하는 Regression 기법들의 입력으로 사용하였고, 두 가지 stage를 분리하여 각 기법에 대해 두 가지 모델을 만들었다. 본 연구에서는 선형회귀기법인 ElasticNet, 비선형회귀기법에 해당하는 SVR(Support Vector Regression), MLP(Multi-Layer Perceptron), k-NN(k-Nearest Neighbor)과 회귀트리 기법인 Decision Tree, Random Forest, Bagging Tree, Gradient Boosting Tree, AdaBoost, XGBoost를 사용하여 MRR 예측 모형들을 개발하였다.

### 4.3 품질 분류 모형

본 연구에서는 목표로 하는 공정 변수 및 공정 조건을 해석하기 위하여 분류 모형을 구축한다. 분류 모형을 구축하기 위한 적용 과정은 <그림 4>와 같다.



<그림 4> 분류 모형 적용 절차

4.1절에서 설명한 데이터 전처리 과정 중 데이터 재구성 단계를 거치고, 이 데이터를 머신러닝을 기반으로 하는 Classification 기법들의 입력으로 사용하였다. 머신러닝 기법으로는 설명할 수 있는 모델로 잘 알려진 트리 기반의 두 가지 알고리즘, Random Forest, XGBoost를 사용하여 분류 모형을 개발하였다.

## V. 실험 결과

### 5.1 품질 예측 모형

10가지 머신러닝 기법을 사용한 예측 모형의 실험 결과는 <표 5>와 같다. R-squared 값은 A, B stage 모두 XGBoost에서 0.974, 0.855로 가장 높게 나타났고, RMSE(Root Mean Square Error)는 A, B stage 모두 5.9286, 3.545로 모두 XGBoost에서 가장 낮게 나타났다. 가장 높은 성능을 나타내는 XGBoost 모형의 하이퍼파라미터는 <표 6>과 같다.

### 5.2 품질 분류 모형

Random Forest, XGBoost를 사용한 분류 모형

의 성능을 비교하기 위해서 <표 7>의 분류 결과 표(Confusion Matrix)의 값을 사용하는 다음과 같은 성능 지표를 사용하였다.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

그 실험 결과는 <표 8>로 나타난다. XGBoost의 성능이 모든 성능 지표 측면에서 가장 좋게 나타났고, 최적 모형의 하이퍼파라미터는 <표 9>와 같다.

<표 5> A stage, B stage 실험 결과

머신러닝 기법	A stage		B stage	
	R-squared	RMSE	R-squared	RMSE
ElasticNet	0.812	15.5149	0.616	5.496
SVR	0.812	15.5149	0.604	5.580
MLP	0.788	16.9243	0.788	4.086
k-NN	0.755	17.6913	0.567	5.835
Decision Tree	0.880	13.2589	0.572	5.801
Random Forest	0.953	7.7914	0.816	3.802
Bagging Tree	0.934	9.1674	0.803	3.934
Gradient Boosting Tree	0.945	8.4532	0.835	3.606
AdaBoost	0.935	9.1465	0.800	3.965
XGBoost	0.974	5.9286	0.855	3.545

<표 6> XGBoost 최적 모델 하이퍼파라미터

	colsample_bytree	gamma	learning_rate	max_depth	max_child_weight	n_estimators	reg_alpha	reg_lambda	sub_sample	tree_method
A stage	0.6	0.5	0.05	3	1	1000	0.6	0.1	0.9	hist
B stage	0.8	0.05	0.05	3	1.5	1500	0.5	0.45	0.8	auto

<표 7> 분류 결과표

		실제 정답	
		True	False
분류 결과	True	True Positive (TP)	False Positive (FP)
	False	False Negative (FN)	True Negative (TN)

〈표 8〉 분류 모델 실험 결과

	accuracy	precision	recall	F1-score
Random Forest	0.97651	0.974138	0.991228	0.982609
<b>XGBoost</b>	<b>0.9993902</b>	<b>0.991304</b>	<b>1</b>	<b>0.995633</b>

〈표 9〉 XGBoost 최적 모델 하이퍼파라미터

max_depth	n_estimators
3	100

## VI. 공정 품질 해석

### 6.1 RuleCOSI

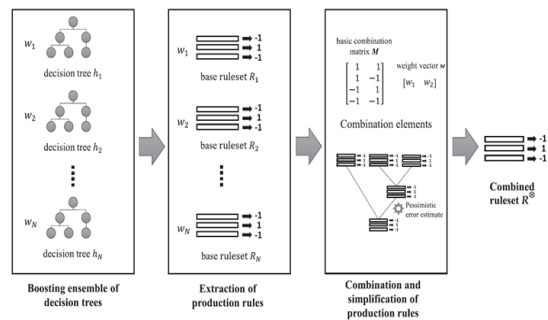
RuleCOSI(Rule Combination and Simplification) 알고리즘은 부스팅 앙상블 알고리즘이 방대한 개수의 의사 결정 나무를 생성할 때 모델을 설명하기 어렵다는 단점을 보완하기 위해 개발되었다. 또한, 기존의 알고리즘들이 가중치를 사용하여 최종 결과를 얻기 어렵다는 단점을 보완하기 위해서 이진 분류 트리의 결과물을 결합하고 단순화하는 방법을 제안하였다[7].

RuleCOSI 알고리즘의 작동 방식은 <그림 5>와 같다. 부스팅 알고리즘이  $n$ 개의 트리를 생성하면, 각 트리에서 한 개의 ruleset을 생성한다. 이때, 한 개의 ruleset에는 한 트리의 최상위 노드부터 최하위 노드까지 가는 모든 경우의 수가 포함되어 있다. 이렇게  $n$ 개의 ruleset이 생성되면, 두 개씩 묶어 단순화하는 과정을 거쳐 총 하나의 결합된 ruleset을 생성하게 된다.

단순화 과정은  $n$ 개의 ruleset 중 두 개씩 묶고, 두 개의 ruleset에 포함되어 있는 모든 rule들 중 한 개를 번갈아 가면서 제거하고 제거했을 때의 성능과 제거하지 않았을 때의 성능을 비교한다. 만약 한 개의 rule을 제거하였을 때 성능이 감소하지 않는다면, 이 rule을 제거하고 그렇지 않다면 rule을 제거하지 않는 과정이다. 이렇게  $n$ 개의 ruleset을 모두 비교하여 더 이상 성능 개선이 없을

때까지 rule들이 제거되면, 한 개의 분류 트리 모형으로부터 최종적인 한 개의 ruleset이 생성된다.

이 알고리즘은 10개의 데이터를 사용하여 RuleCOSI 알고리즘을 통한 rule 제거 과정을 거친 후의 성능을 기존 트리 모형의 성능과 비교를 해 본 결과 대부분의 경우에 RuleCOSI 알고리즘을 사용했을 때 성능이 유사하거나 더 좋은 성능을 보인다는 것을 확인하였다.



〈그림 5〉 RuleCOSI 작동 방식 [7]

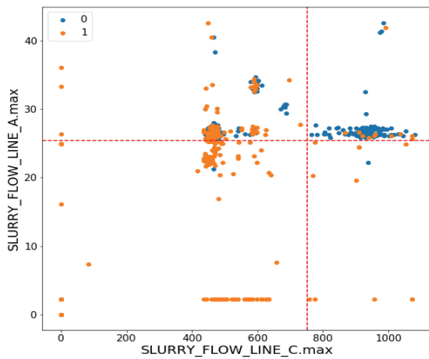
### 6.2 분석 결과

RuleCOSI를 4.3에서 구축한 분류 모형들에 적용하여 생성된 높은 MRR을 달성하기 위한 ruleset은 <표 10>과 같이 15개로 나타났다.

<표 10>을 보면 본 연구에서 사용한 데이터에 포함된 19가지의 변수 중 8가지의 변수만 포함되어 있다는 것을 확인할 수 있다. 이 ruleset 중 Rule 1을 예시로 보자면, USAGE\_OF\_POLISHING\_TABLE의 skewness가 0.537 초과, SLURRY\_FLOW\_LINE\_C의 max가 751.8 이하, RETAINER\_RING\_PRESSURE의 skewness가 2.682 이하, USAGE\_OF\_BACKING\_FILM의 p2p가 10.417 이

하라면 높은 MRR을 달성할 수 있다고 해석할 수 있다. 따라서 CMP 공정의 품질을 분석하기 위해 사용된 19가지의 변수 중에서 8가지의 변수가 품질에 가장 큰 영향을 끼치는 유의미한 변수로 판명되었다고 볼 수 있다.

이 8가지 변수 중 USAGE\_OF\_POLISHING\_TABLE, SLURRY\_FLOW\_LINE\_C, RETAINER\_



<그림 6> Rule 8을 기준으로 웨이퍼를 나타낸 그래프

RING\_PRESSURE, SLURRY\_FLOW\_LINE\_A, HEAD\_ROTATION, USAGE\_OF\_DRESSER 이렇게 6가지 변수는 Preston 방정식에 포함되어 있는 7가지 변수 안에 속한다.

또한, 데이터에서 특징을 추출하기 위해 사용된 10가지 통계 지표 중에서는 5가지 지표인 skewness, p2p, max, variation, min만 MRR에 영향을 끼치는 조건으로 나타났다. 예를 들어, USAGE\_OF\_POLISHING\_TABLE 변수는 이 변수에 적용된 10가지 통계 지표 중에서는 skewness와 p2p에 유의하여 공정이 진행되어야 한다는 것을 파악할 수 있다.

<표 10>의 Rule 8은 SLURRY\_FLOW\_LINE\_C의 최댓값이 751.8 이하이고 SLURRY\_FLOW\_LINE\_A의 최댓값이 25.417 이하이면 높은 MRR을 유지할 수 있다는 의미이다. Rule 8을 해석하기 위하여 포함된 두 가지 변수의 최댓값을 기준으로 모든 웨이퍼를 시각화한 결과는 <그림 6>과 같다. 그래프의 점선은 Rule 8의 조건을 나타낸 경계선이고, Rule 8의 조건을 만족하는 범위 내에는 class

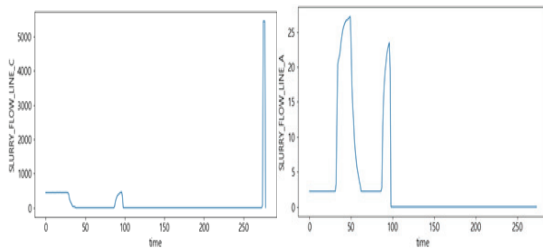
<표 10> 고품질 MRR 유지 조건에 관한 ruleset

	USAGE_OF_POLISHING_TABLE	SLURRY_FLOW_LINE_C	RETAINER_RING_PRESSURE	USAGE_OF_BACKING_FILM	SLURRY_FLOW_LINE_A	PRESSURIZED_CHAMBER_PRESSURE	HEAD_ROTATION	USAGE_OF_DRESSER
Rule 1	skewness > -0.537	max <= 751.8	skewness <= 2.682	p2p <= 10.417				
Rule 2	skewness > -0.731			var <= 10.234	p2p <= 25.417			
Rule 3	p2p <= 10.0					max <= 147.143 var <= 1391.627		
Rule 4	skewness > -0.537 p2p <= 10.0	max <= 788.2					var <= 46.807	
Rule 5	skewness > -0.272			p2p <= 10.417			var <= 46.807	
Rule 6				p2p <= 10.417		max <= 147.143 var <= 1391.627		
Rule 7	skewness > -0.537	max <= 751.8		p2p <= 10.417				skewness <= 0.261
Rule 8		max <= 751.8			max <= 25.417			
Rule 9		max <= 788.2		var <= 4.22		max <= 147.143		
Rule 10				var <= 4.22		var <= 1391.627	min >= 137.6	
Rule 11	skewness > -0.731			var <= 10.234			skewness <= 3.286	
Rule 12				var <= 4.22			skewness <= 3.286 min > 137.6	
Rule 13	skewness > -0.731			var <= 10.234			skewness <= 3.286	skewness <= 0.261
Rule 14				var <= 4.22	skewness <= 2.667			
Rule 15	skewness > -0.731					max <= 147.143	skewness <= 3.286	

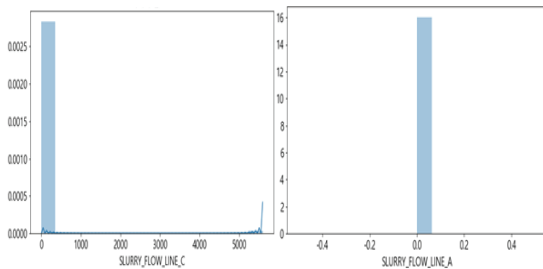


1, 즉 높은 MRR을 갖는 웨이퍼만 포함되어 있는 것을 확인할 수 있다. 따라서 공정이 진행될 때 Rule 8에 포함된 두 가지 변수에 적용된 최댓값이 이 범위 내에 들어가게 처리된다면 공정 품질을 좋게 유지할 수 있다고 해석한다. 이는 전체 15가지 rule에 같은 방법으로 적용할 수 있다.

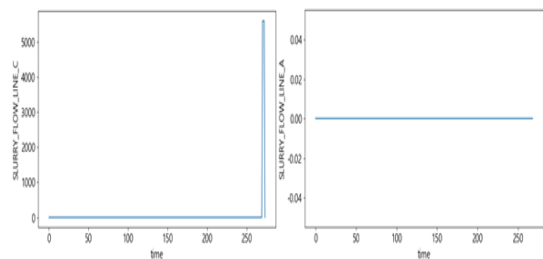
또한, <그림 7>은 모든 웨이퍼에 대해 Rule 8에 포함된 두 가지 변수, SLURRY\_FLOW\_LINE\_A와 SLURRY\_FLOW\_LINE\_C가 시간에 따라 어떻게 변하는지를 나타낸 그래프이고, <그림 8>은 모든 웨이퍼에 대해 이 두 가지 변수가 어떤 분포를 그리는지 보기 위해 시각화한 그래프이다.



<그림 7> 일부 웨이퍼의 특이한 경우



<그림 8> 웨이퍼의 Rule 8에 포함된 변수의 분포도



<그림 9> Rule 8에 포함된 변수를 시간에 따라 나타낸 그래프

class 1로 분류된 웨이퍼 중 대부분은 <그림 7>, <그림 8>과 같은 형태를 띠는 반면, 일부 웨이퍼가 <그림 9>처럼 시간에 따라 특이한 형태를 나타내는 것을 확인하였다.

Rule 8에 포함된 두 가지 변수 모두 유의미한 통계 지표는 최댓값인 것으로 나타났는데, <그림 9>의 두 가지 그래프에서 최댓값을 보면 대략 각각 751.8, 25.417 정도인 것을 확인할 수 있다.

대부분의 웨이퍼에 대해 Rule 8에 포함된 두 가지 변수가 <그림 7>, <그림 8>의 형태를 띠고 높은 MRR을 갖는 정상인 경우이다. 그러나 <그림 9>와 같은 형태를 보이는 변수를 갖는 웨이퍼는 평소와 다르게 공정 변수가 처리되는 과정에서 이상이 생겼지만 높은 MRR을 갖는 경우이다. 이러한 경우는 평소와 다르게 SLURRY\_FLOW\_LINE\_C 변수가 공정 도중에 751.8만큼 투입되었을 때, 이 변동을 완화하고 높은 품질을 계속 유지하기 위해서 SLURRY\_FLOW\_LINE\_A 변수 또한 25.417만큼 공정 도중 투입한다고 해석한다. 따라서, 이 두 가지 변수에 대한 변동이 생긴다면 최대 허용 수준이 이 값이라고 판단한다. 이는 전체 15가지 rule에 같은 방법으로 적용할 수 있다.

## VII. 결론 및 향후 연구

본 연구에서는 CMP 공정 데이터를 이용하여 MRR을 예측하기 위해 머신러닝을 사용한 예측 모델을 개발하였다. 시계열 데이터에서 특징을 추출하기 위하여 평균, 분산, RMS와 같은 통계 지표들을 적용하여 전처리하였고 추출된 값들을 사용하여 SVR, XGBoost 등 머신러닝 기반의 모델을 구축하고 성능을 비교 분석하였다.

앙상블 기법 중 Boosting을 사용하는 머신러닝 기반 모델의 성능이 높게 나타났고, XGBoost 모델의 성능이 가장 높게 나타났다. XGBoost는 그 리디 알고리즘과 Boosting 알고리즘을 분산 환경

에서도 실행이 되도록 구현되어 성능과 자원 효율이 높아서 속도가 빠르고 성능이 좋기 때문에 최근 각광을 받는 기법이다.

RuleCOSI 알고리즘을 적용하여 추출된 rule들로부터 CMP 공정 품질에 큰 영향을 미치는 8개의 변수와 공정을 진행하는 과정에서 이 변수들을 처리할 때 어떠한 통계 지표에 유의해야 하는지 파악하였고, 총 10가지 중 5가지의 지표가 유의하다고 나타났다. 또한, 어떠한 공정 변수에 평상시와 다른 변동이 생겼을 때 높은 공정 품질을 유지하기 위해서 어느 정도까지의 허용이 가능한지 최대 허용 수준을 파악할 수 있었다.

이렇게 구해진 유의미한 변수와 통계 지표만을 사용하거나, 기존 연구와 같이 Preston 방정식을 결합하여 예측 모델을 구축하고 성능을 높여볼 수 있을 것이다. 뿐만 아니라 현재는 RuleCOSI 알고리즘을 적용하여 추출된 rule들을 해석하기 위해 각 rule들에 대한 그래프를 그려보았지만, 이를 더 효율적으로 해석할 방법에 대해 구상해 볼 필요성이 존재한다. 최근 머신러닝 및 딥러닝 모델을 해석하는 XAI(eXplainable Artificial Intelligence)에 대한 관심이 대두되었는데, 잘 알려진 LIME(Local Interpretable Model-agnostic Explanations)을 적용하여 CMP 공정을 해석해볼 수 있을 것이다.

## 참 고 문 헌

- [1] Zhixiong Li, Dazhong Wu, Tianyu Yu, "Prediction of Material Removal Rate for Chemical Mechanical Planarization Using Decision Tree-Based Ensemble Learning", *Journal of Manufacturing Science and Engineering*, Vol. 141 Issue. 3 : 031003, March 2019.
- [2] Peng Wang, Robert X.Gao, Ruqiang Yan, "A Deep Learning-based Approach to Material Removal Rate Prediction in Polishing", *CIRP Journal of Manufacturing Science and Technology*, Vol. 66, April 2017, pp. 429-432.
- [3] Hyunseop Lee, Haedo Jeong, "A Wafer-Scale Material Removal Rate Profile Model for Copper Chemical Mechanical Planarization", *International Journal of Machine Tools & Manufacture*, Vol. 51, January 2011, pp. 395-403.
- [4] Jianfeng Luo, David Dornfeld, "Material Removal Mechanism in Chemical Mechanical Polishing: Theory and Modeling", *IEEE Transactions on Semiconductor Manufacturing*, Vol. 14, Issue. 2, 2001, pp. 112-133
- [5] Tianyu Yu, Zhixiong Li, Dazhong Wu. "Predictive Modeling of Material Removal Rate in Chemical Mechanical Planarization with Physics-informed Machine Learning", *Wear*, Vol. 426-427, 2019. pp. 1430-1438.
- [6] N.Y. Nguyen, Z.W. Zhong, Yebing Tian, "An analytical investigation of pad wear caused by the conditioner in fixed abrasive chemical-mechanical polishing", *International Journal of Advanced Manufacturing Technology*, Vol. 77, Issue. 5-8, 2015, pp. 897-905.
- [7] Josue Obregon, Aekyung Kim, Jae-Yoon Jung, "RuleCOSI: Combination and simplification of production rules from boosted trees for imbalanced classification", *Expert Systems With Applications*, Vol. 126, 2019, pp. 64-82.

## 사 사

이 논문은 과학기술정보통신부의 재원으로 한국연구재단(No. 2019R1F1A1064125)과, 산업통상자원부의 재원으로 한국산업기술진흥원(스마트공장 운영설계 전문인력 양성사업, N0002429)의 지원을 받아 수행된 연구임.

## 저자 소개



### 안 정 언(Jeong-Eon Ahn)

·2019년 : 경희대학교 산업경영  
공학부 (공학사)

·2020년~현재 : 경희대학교  
산업경영공학과 산업지능연구실  
석사과정

·2020년~현재 : 경희대학교

스마트제조지능 (SMI) 융합 과정

·관심분야 : 빅데이터, 제조데이터 사이언스,  
반도체 고영 품질



### 정 재 윤(Jae-Yoon Jung)

·1999년 : 서울대학교 산업공학과  
(공학사)

·2001년 : 서울대학교 산업공학과  
(석사)

·2005년 : 서울대학교 산업공학과  
(박사)

·2005년~2006년 : 아인트호벤공대 정보시스템학과  
박사후연구원

·2006년~2007년 : 서울시 유비쿼터스컴퓨팅 원천기술  
지원센터 선임연구원

·2007년~현재 : 경희대학교 산업경영공학과 조교수,  
부교수, 교수

·관심분야 : 빅데이터, 프로세스마이닝,  
제조데이터 사이언스