

# AIAP Final Project

---

**24-1 AI Applications Programming**

**21900031 Jin Kwak**

School of Mechanical and Control Engineering



# 목차

---

1. Introduction about Welding Process
2. Data Processing
3. Analysis Model
4. Result
5. Conclusion



# 1. Introduction (1/3)

---

- **Introduction to Process**

- Welding is the process of fusing two/more parts(metal, glass or plastic) using pressure, heat or both.
- Requires practice and precision due to its technical/ labor-intensive barrier.

- **Problems of the process**

- It is difficult to detect welding defects instantaneously through visual testing.
- Destructive testing is not appropriate for this process
- Most of defects are caused from inconstant voltage/current
- Process parameters are dependent on each other → Difficult to analyze



# 1. Introduction (2/3)

---

- **Solution**
  - Independent process parameters(Voltage, Current) → Classify Normal/ Defect
- **Expectation**
  - Process Optimization using Process Data(Voltage, Current, Welding Speed)
  - Apply Machine Learning Algorithm to detect Defect
  - Influential Process Parameters and set boundaries for less defect



# 1. Introduction (3/3)

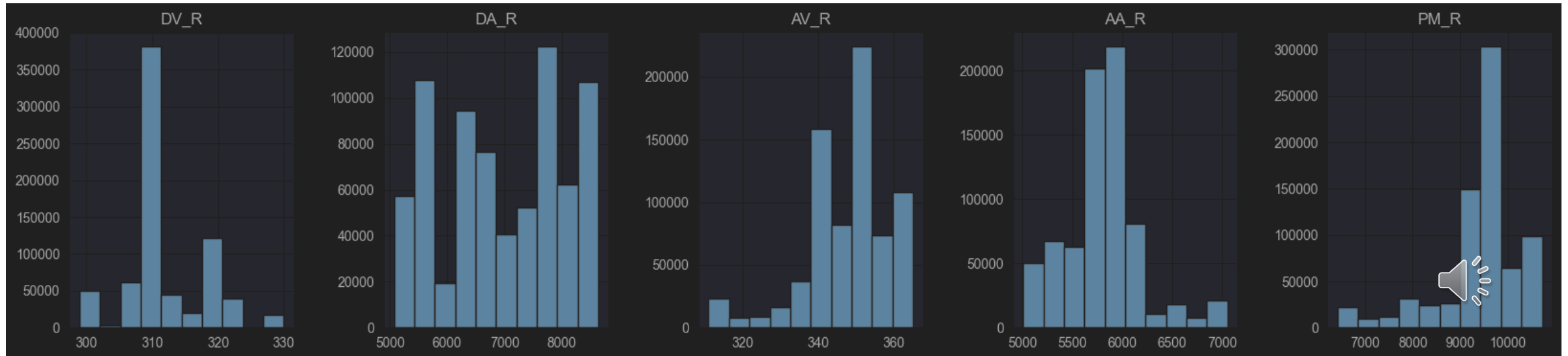
- Data Introduction
  - Category: Pipe Welding
  - Process: Welding(Submerged Arc Welding)
  - Data Collecting Equipment: PLC in Welding machine

속성	설명	구분
PIPE_NO	제품 번호 부여	독립 변수
DV_R	직류 전압 측정 값	독립 변수
DA_R	직류 전류 측정 값	독립 변수
AV_R	교류 전압 측정 값	독립 변수
AA_R	교류 전류 측정 값	독립 변수
PM_R	공정 용접 속도 측정 값	독립 변수
FIN_JGMT	FIN_JGMT = 1:양품/ FIN_JGMT=0:불량	종속 변수



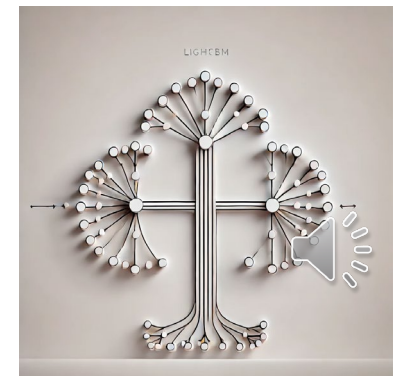
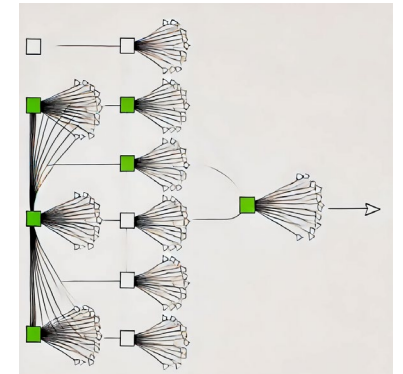
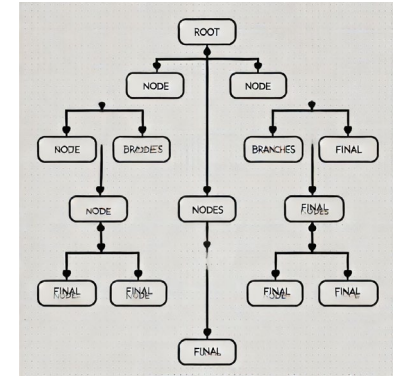
## 2. Data Preprocessing

- Remove Missing Value
- Replace Outliers
  - Replace values over Upper 2%, Lower 2% to Upper 2% value and Lower 2% value
- Data Improvement
- Secure Completeness/Accuracy of Data
- Improve on Model performance(Prevent overfitting/Improve Validity)



# 3. Analysis Model (1/2)

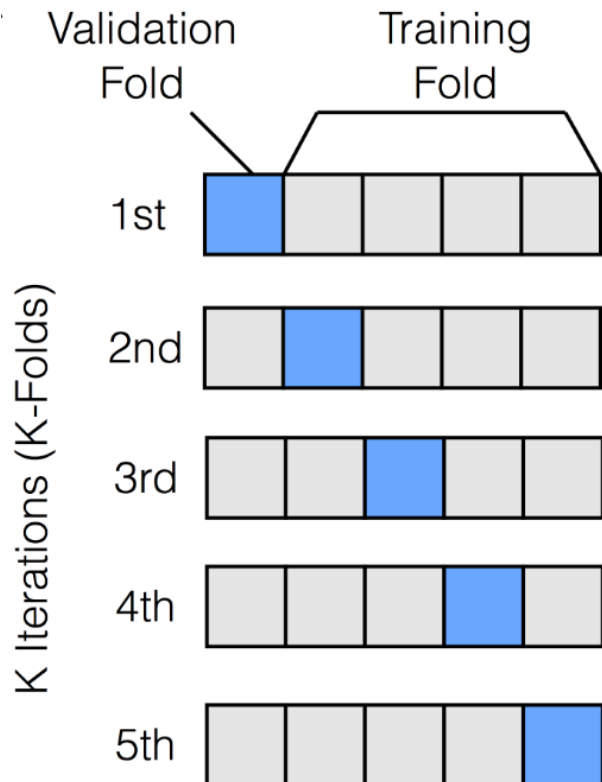
- **Decision Tree**
  - **Easy Analysis, Non-Linear Modeling**
  - **Overfitting/ Sensitive to small change in data**
  - **Algorithmic Bias due to imbalance in data**
- **XGBoost**
  - **L1/L2 Regularization, Early Stopping , Automatic Pruning → Prevent Overfitting**
  - **Parallel Processing, Efficient Memory Use, → High performance/efficiency**
  - **High Complexity**
- **LightGBM**
  - **Fast Learning/ High performance/ Memory Efficiency**
  - **High Complexity/ Large number of Parameters**



# 3. Analysis Model (1/2)

- **K-Fold Cross Validation**

- Evaluation of performance and capability of machine learning models.
- Crucial Role in building reliable and robust models.



## Decision Tree

KFold 1 교차검증 정확도 : 0.9581559420995013, 학습 데이터 크기:345280, 검증 데이터 크기:172641  
KFold 2 교차검증 정확도 : 0.9589029193697869, 학습 데이터 크기:345281, 검증 데이터 크기:172640  
KFold 3 교차검증 정확도 : 0.9582947173308619, 학습 데이터 크기:345281, 검증 데이터 크기:172640  
교차검증 중 가장 높은 정확도 : KFold 2, 0.9589029193697869

## XGBoost

KFold 1 교차검증 정확도 : 0.9493573369014313, 학습 데이터 크기:345280, 검증 데이터 크기:172641  
KFold 2 교차검증 정확도 : 0.9499478683966636, 학습 데이터 크기:345281, 검증 데이터 크기:172640  
KFold 3 교차검증 정확도 : 0.9488531047265987, 학습 데이터 크기:345281, 검증 데이터 크기:172640  
교차검증 중 가장 높은 정확도 : KFold 2, 0.9499478683966636

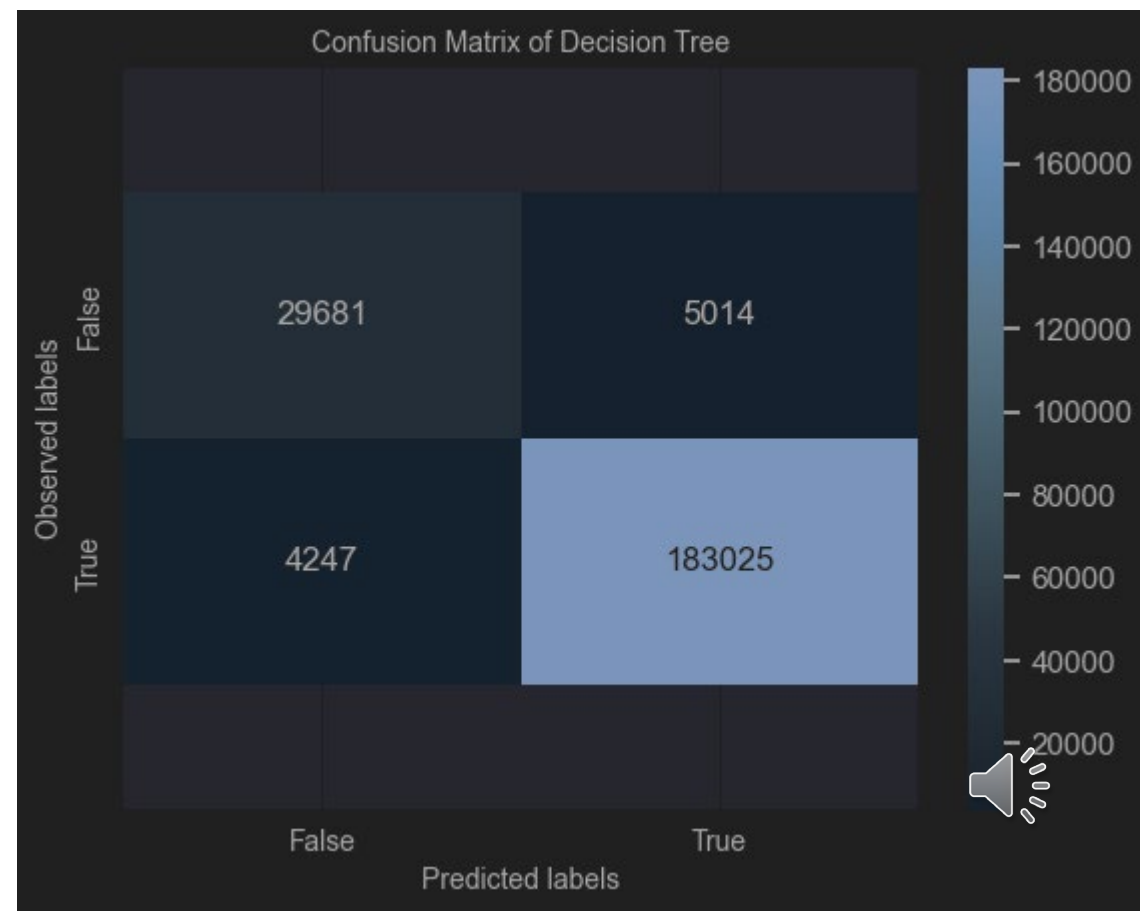
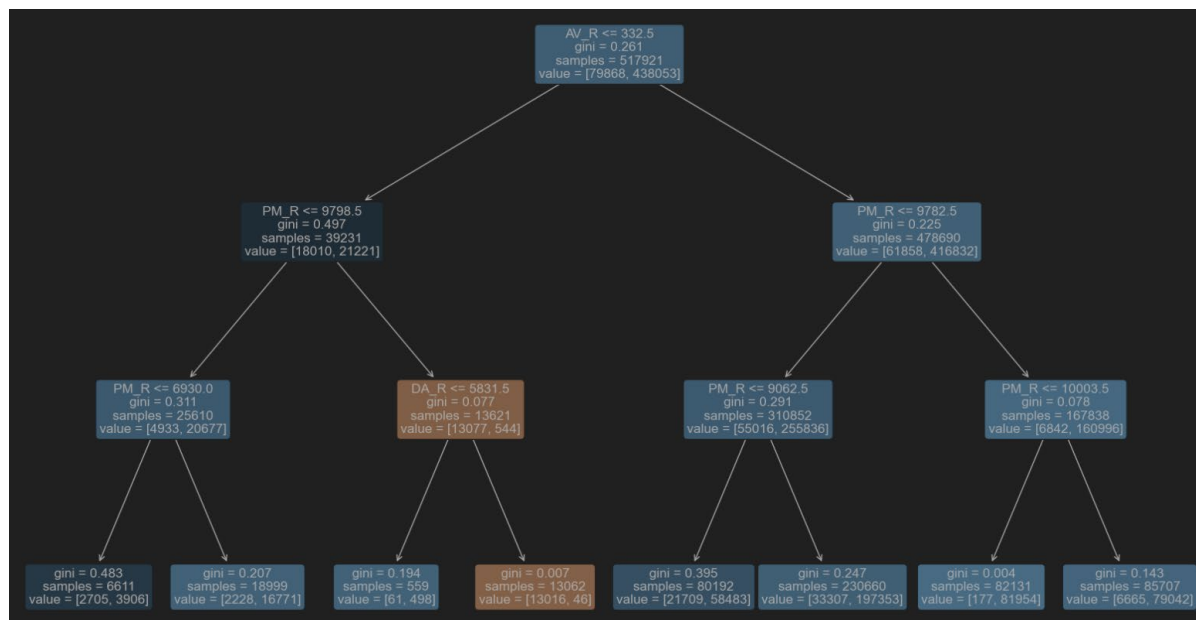
## Light GBM

KFold 1 교차검증 정확도 : 0.9474168940170643, 학습 데이터 크기:345280, 검증 데이터 크기:172641  
KFold 2 교차검증 정확도 : 0.9469010658016682, 학습 데이터 크기:345281, 검증 데이터 크기:172640  
KFold 3 교차검증 정확도 : 0.9472254402224282, 학습 데이터 크기:345281, 검증 데이터 크기:172640  
교차검증 중 가장 높은 정확도 : KFold 1, 0.9474168940170643



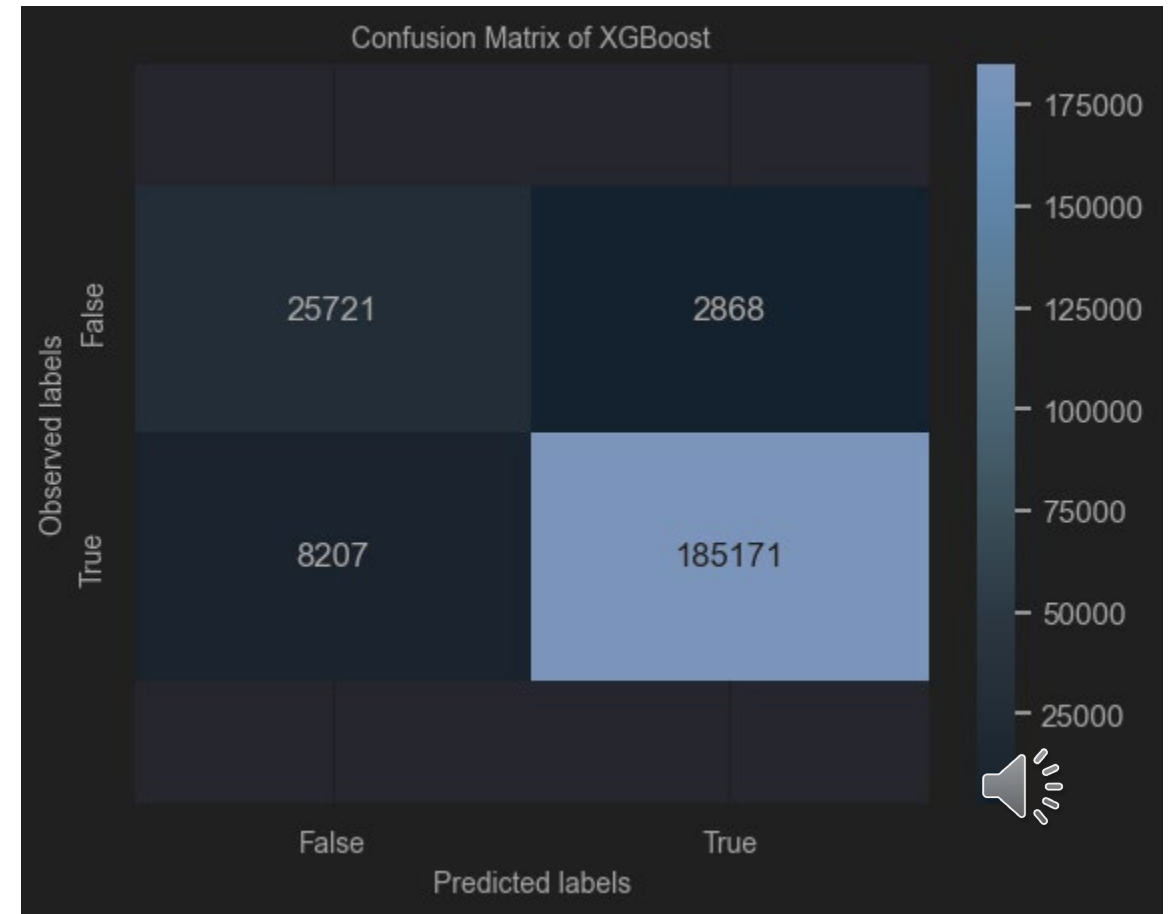
# 4. Result (1/4)

- Decision Tree
  - Accuracy: 0.9583
  - Precision: 0.9773
  - Recall: 0.9733
  - F1 Score: 0.9753



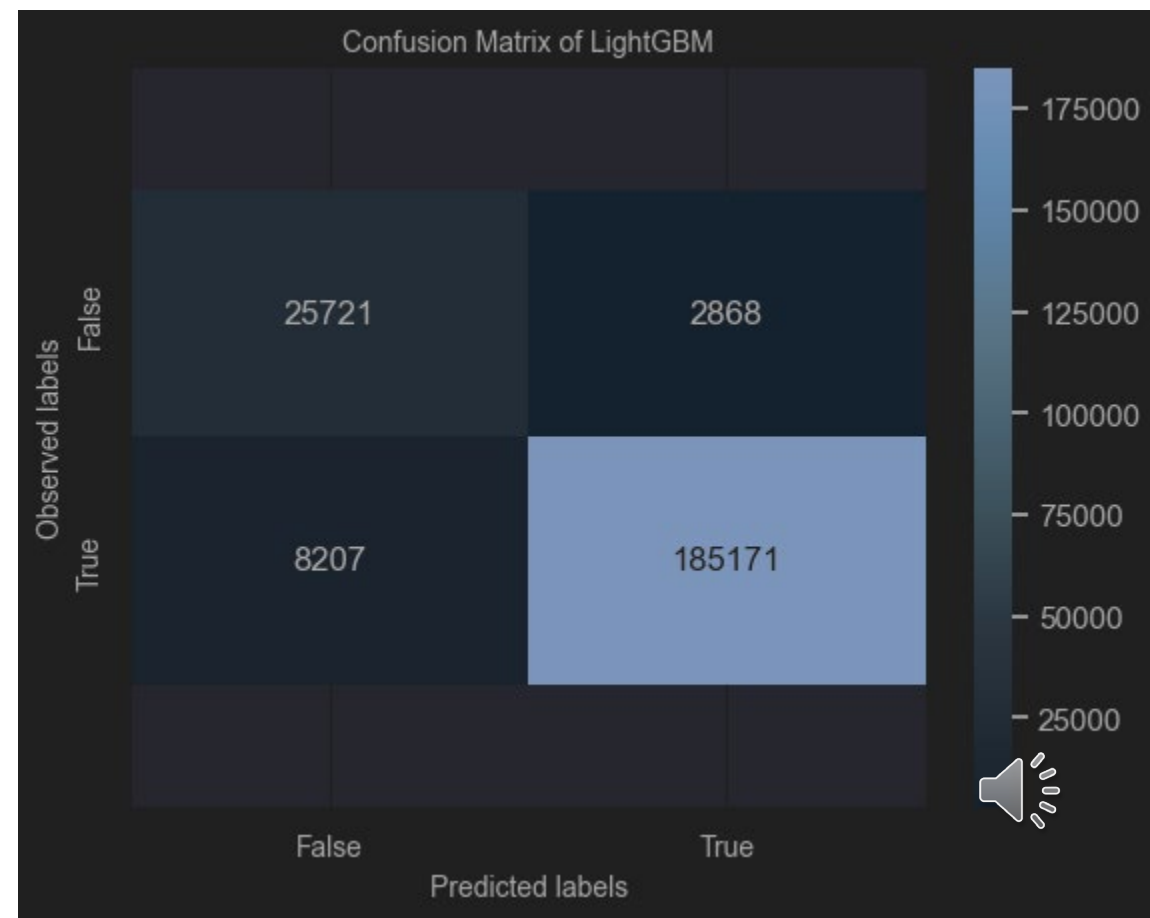
## 4. Result (2/4)

- **XGBoost**
  - Accuracy: 0.9501
  - Precision: 0.9576
  - Recall: 0.9847
  - F1 Score: 0.9710



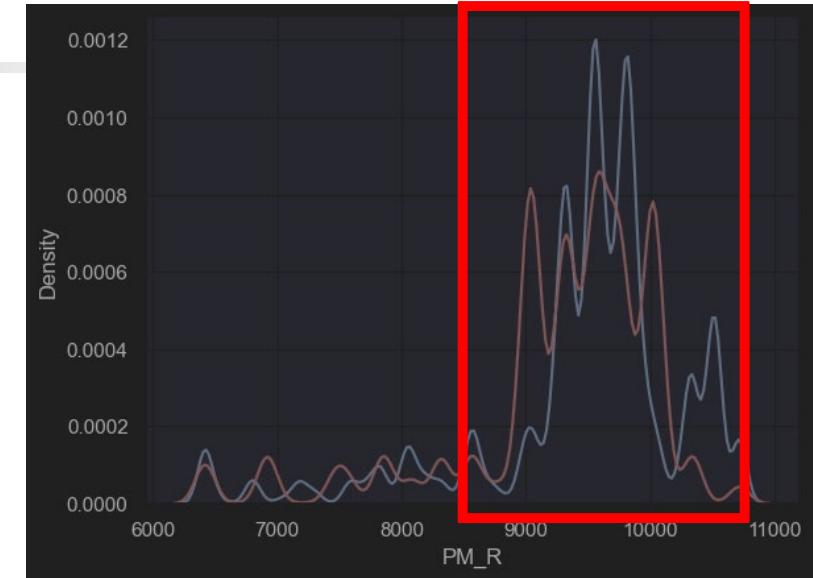
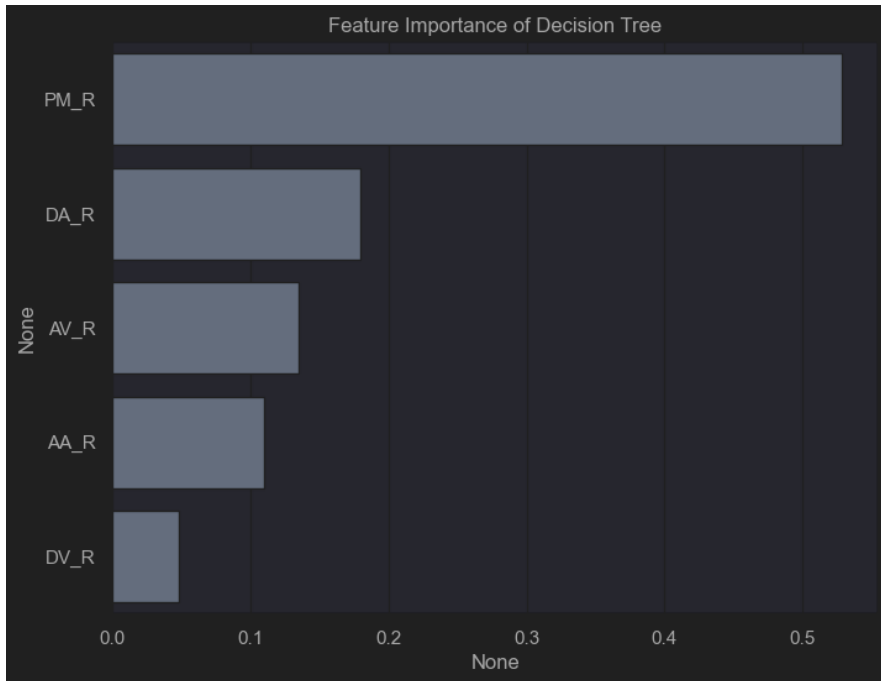
## 4. Result (3/4)

- LightGBM
  - Accuracy: 0.9478
  - Precision: 0.9520
  - Recall: 0.9883
  - F1 Score: 0.9697



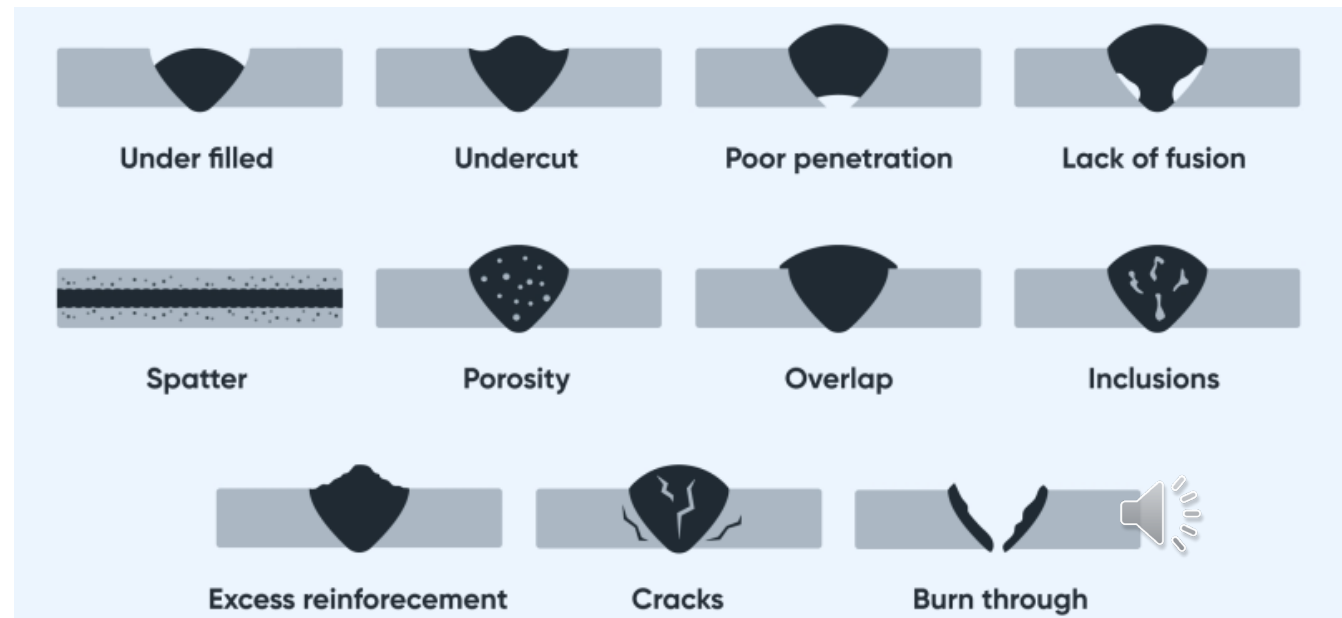
## 4. Result (4/4)

- The most important feature that influences welding process is Welding Speed
- The optimum value boundary for welding speed/ Alternating Voltage are determined where other variables are in average boundary



# 5. Conclusion

- Conservatively, hyperparameters can be adjusted to get high threshold for precision increase.
- Welding Process Optimization using **Decision Tree** enables Real Time Classifier
- Human labor for defect detection may be automated
- Improvement for this process
  - For each type of welding defects, the variables trend may vary → Make each defect classification model and go all over them.
- Difficulties
  - Data for each defects should be collected.



# THANK YOU

---

