# Report on Welding Process Optimization Using AI

**Author: Jin Kwak/21900031**

**Date: 2024.06.14**

[YOUTUBE LINK](#)

# Introduction

## In the beginning

Welding is a important process in manufacturing industries. The process involves the fusion of two or more parts (typically metals/ glass/ plastic ...etc) through the application of heat, pressure, or both. Due to the technical and labor-intensive(depends on worker's ) nature of welding, ensuring the quality of welds is paramount. Studies have shown that welding defects often arise from variations in voltage, current, welding speed during the welding process. These defects are challenging to detect through visual inspection alone, as they may not be immediately apparent on the surface.

Visual inspection(Visual Testing) and destructive testing methods are traditionally used to identify welding defects, but these methods have limitations. Visual inspection is not always reliable and it cannot be tested instantaneously, whereas destructive testing is impractical for every welded part. Moreover, the process parameters in welding are interdependent with each other, making it difficult to analyze and optimize the process.
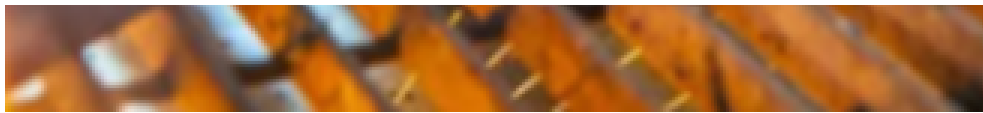
Figure 1. Welding Process

## Research Motivation

The motivation for this research is to leverage artificial intelligence (AI) to optimize the welding process. By analyzing process data such as voltage, current, and welding speed, we aim to develop a machine learning model that can classify welds as normal or defective in real-time. This model will help in reducing the reliance on manual inspections, improving the overall efficiency of the welding process, and ensuring higher quality in the final products.

# Methodology

## Data Collection

The data used in this study were collected from a submerged arc welding process used in pipe welding. The data attributes include:

| 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 |

## Data Preprocessing

1. **Remove Missing Values**: Ensured the dataset's completeness by eliminating any records with missing values.

2. **Remove Outliers**: Identified and replaced outliers. Values above the upper 2% and below the lower 2% were replaced with the respective boundary values.

3. **Data Normalization**: Applied Min-Max scaling to normalize the data, bringing all values into the range [0, 1].
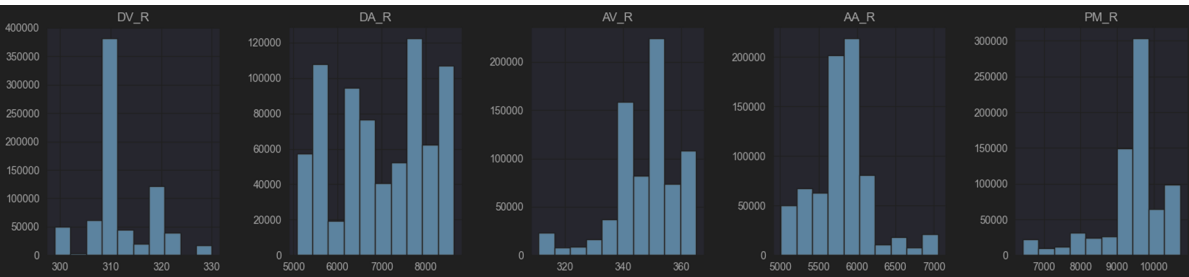

Figure 2. Data Preprocessing

## Model Development

Three machine learning algorithms were employed to analyze and classify the welding data:
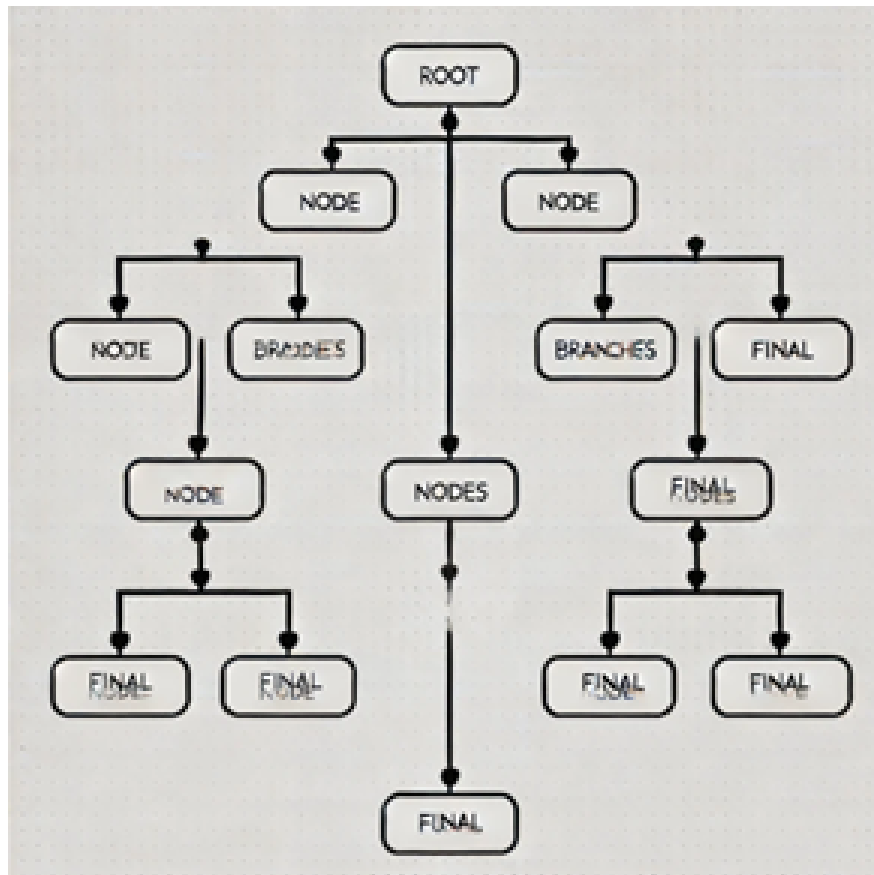
# 1. Decision Tree:



Figure 3. Decision Tree Visualization

- Easy to analyze
- Easy interpretation and visualization
- Non-linear classification
- May overfit to the data if the layer is deep
- Sensitive to small changes in data
- Algorithmic Bias due to imbalance in data
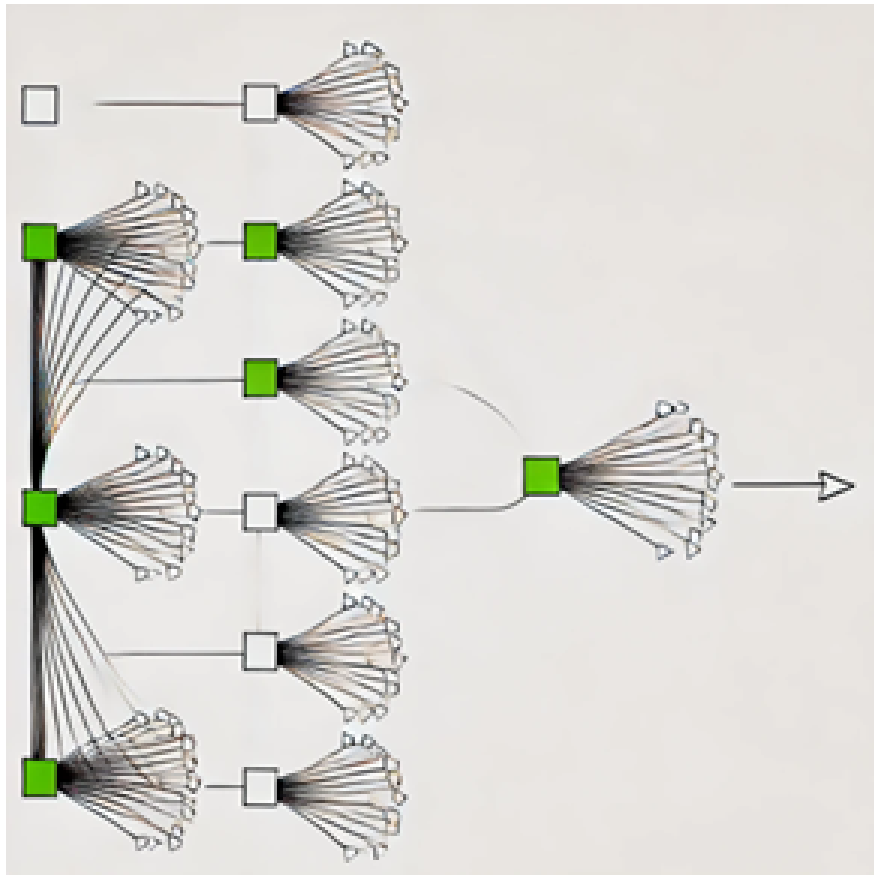
## 2. XGBoost:



Figure 4. XGBoost Visualization

- L1, L2 Regularization prevent overfitting
- Early Stopping Method to prevent overfitting and time-saving
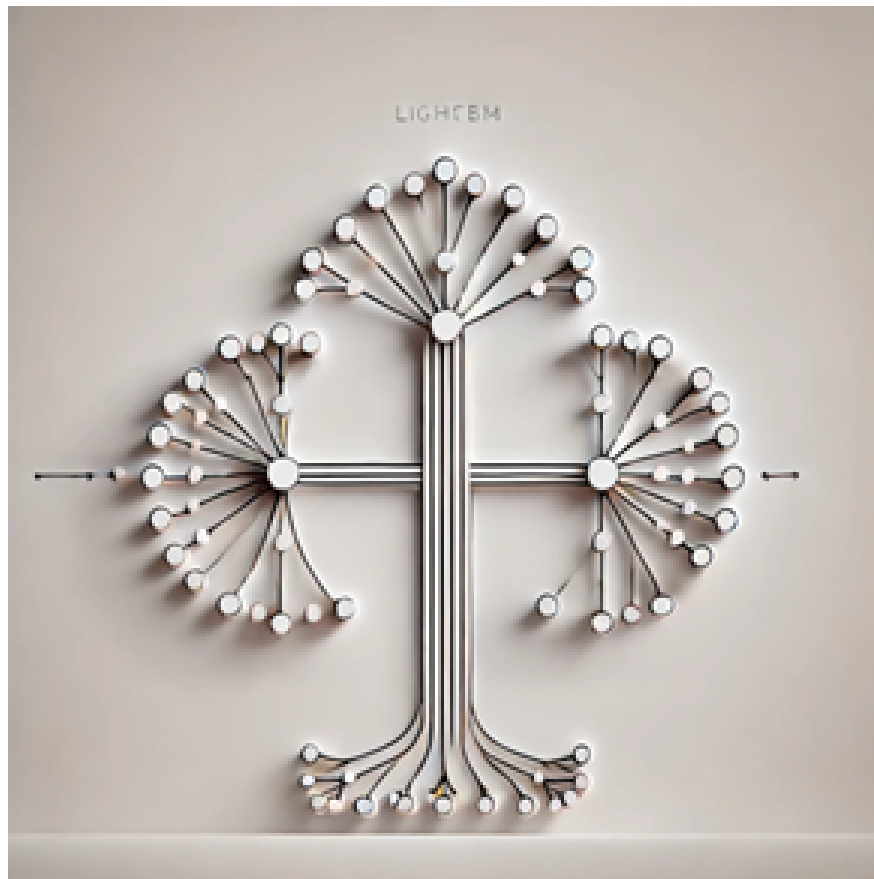- Automatic Pruning method to prevent overfitting

### 3. LightGBM:



Figure 5. LightGBM Visualization

- Fast Learning
- High Performance
- Memory Efficiency
- High complexity
- Takes time for balancing

## Model Validation

K-Fold cross-validation was used to evaluate the models' performance. This method involves splitting the data into k subsets (folds), training the model on k-1 folds, and validating it on the remaining fold. This process is repeated k times, with each fold used exactly once as the validation data. The final performance metrics are averaged over all k iterations.
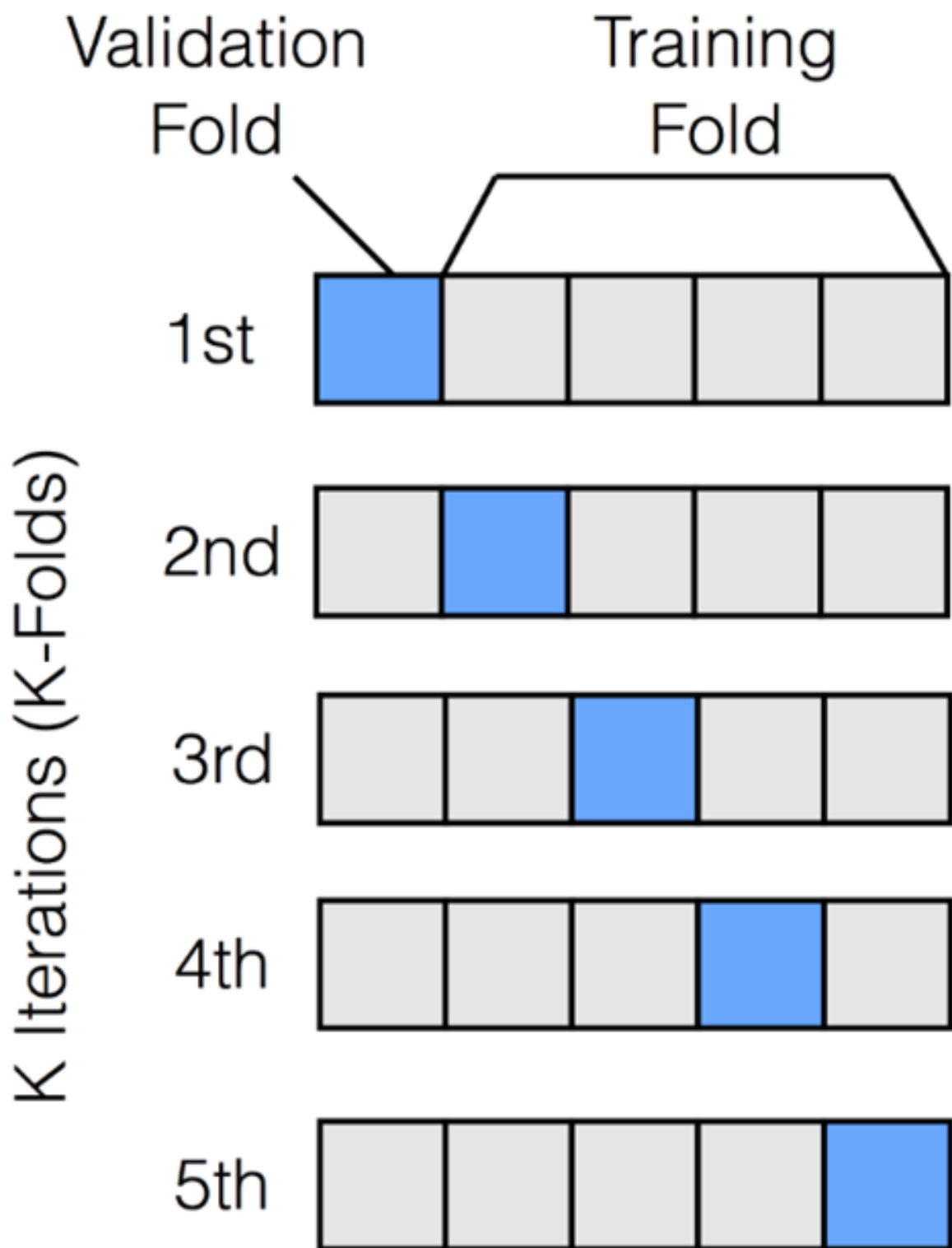
Figure 6. K-Fold Cross Validation Visualization

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

Figure 7. K Fold Cross Validation

# Results

**Decision Tree:**

- **Accuracy**: 95.83%

- **Precision**: 97.73%

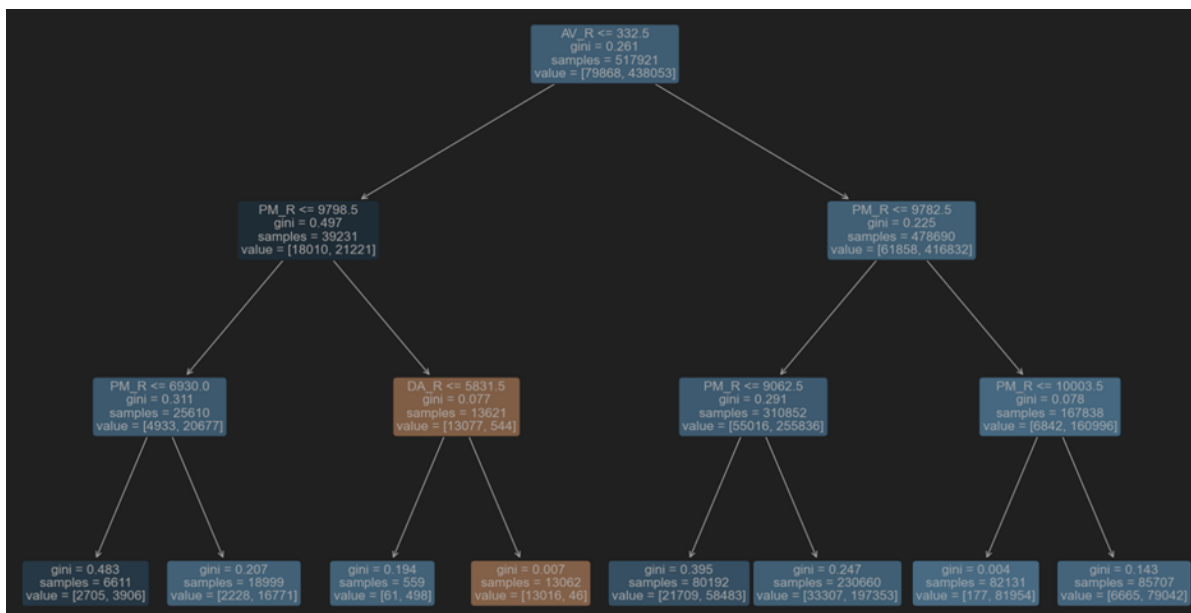- **Recall**: 97.33%

- **F1 Score**: 97.53%
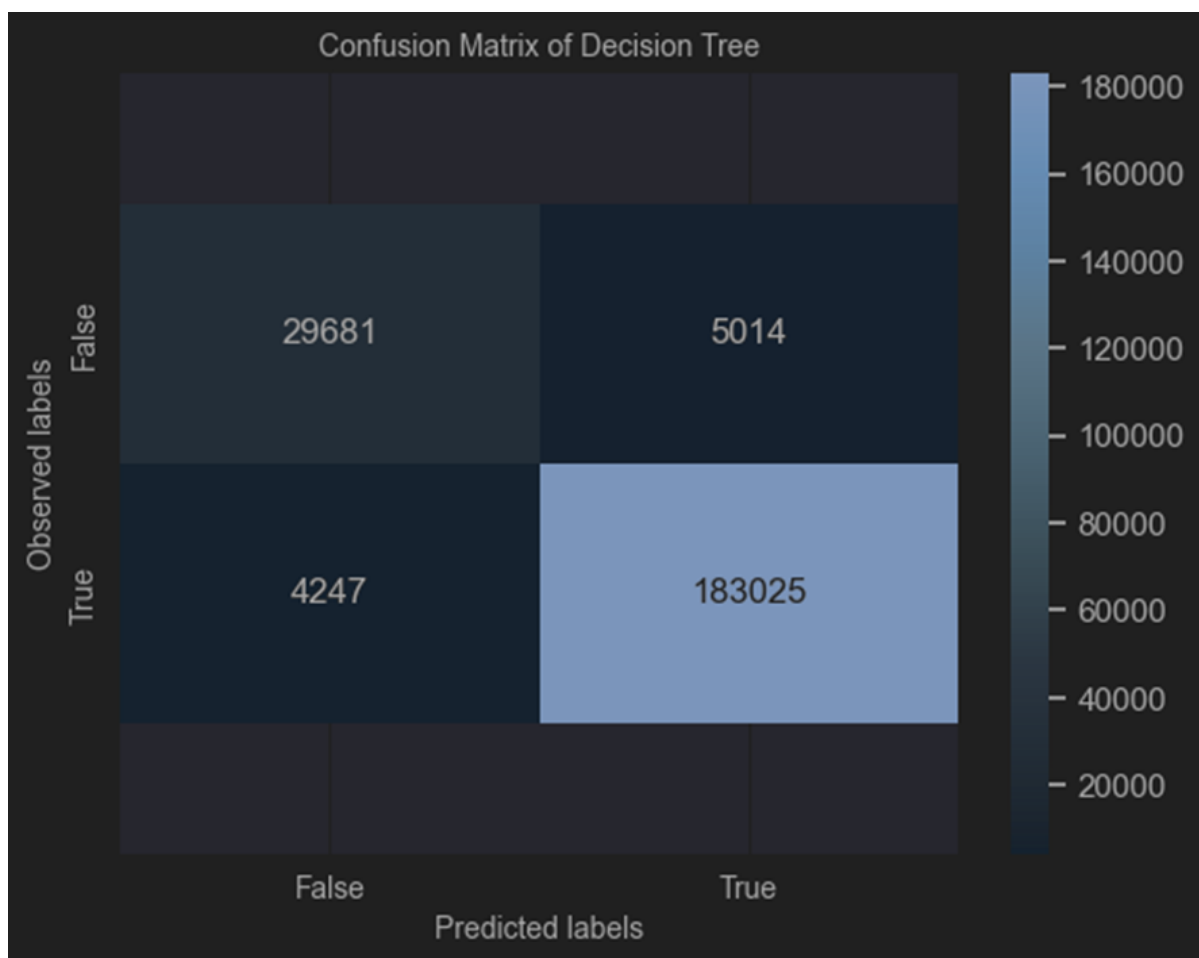


Figure 8. Decision Tree

Figure 9. Confusion Matrix of Decision Tree

**XGBoost:**

- **Accuracy**: 95.01%

- **Precision**: 95.76%

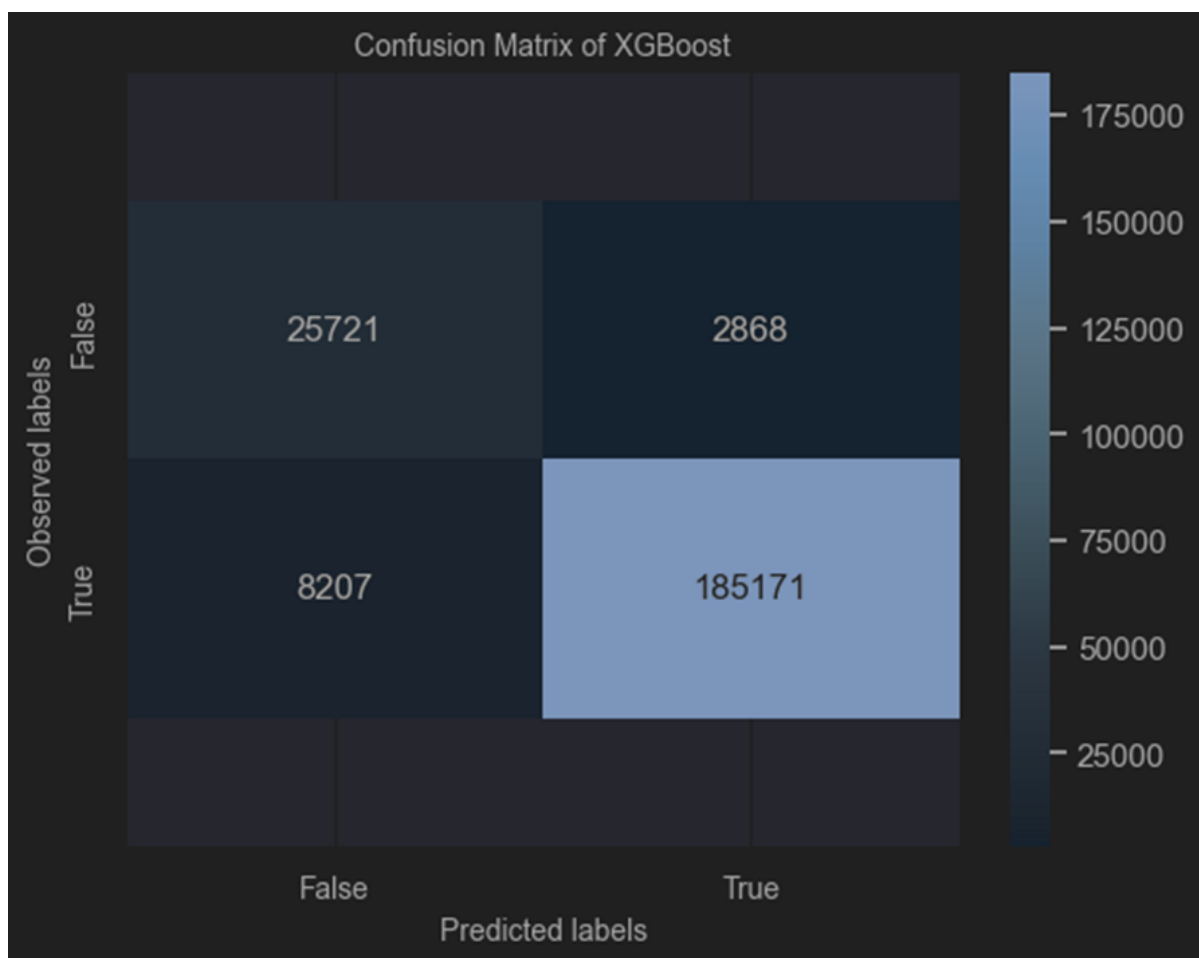- **Recall**: 98.47%

- **F1 Score**: 97.10%

Figure 10. Confusion Matrix of XGBoost

**LightGBM:**

- **Accuracy**: 94.78%

- **Precision**: 95.20%
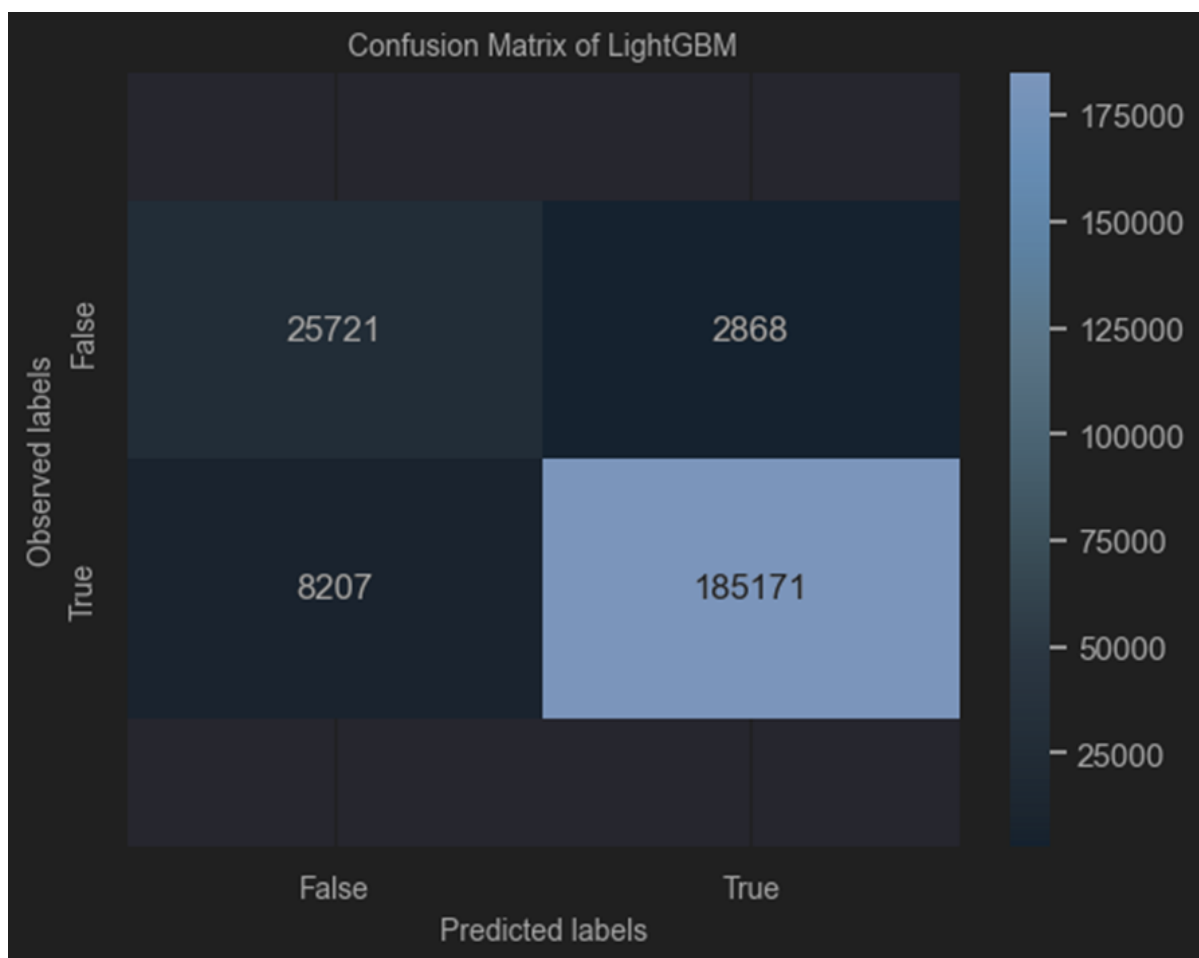
- **Recall**: 98.83%

- **F1 Score**: 96.97%

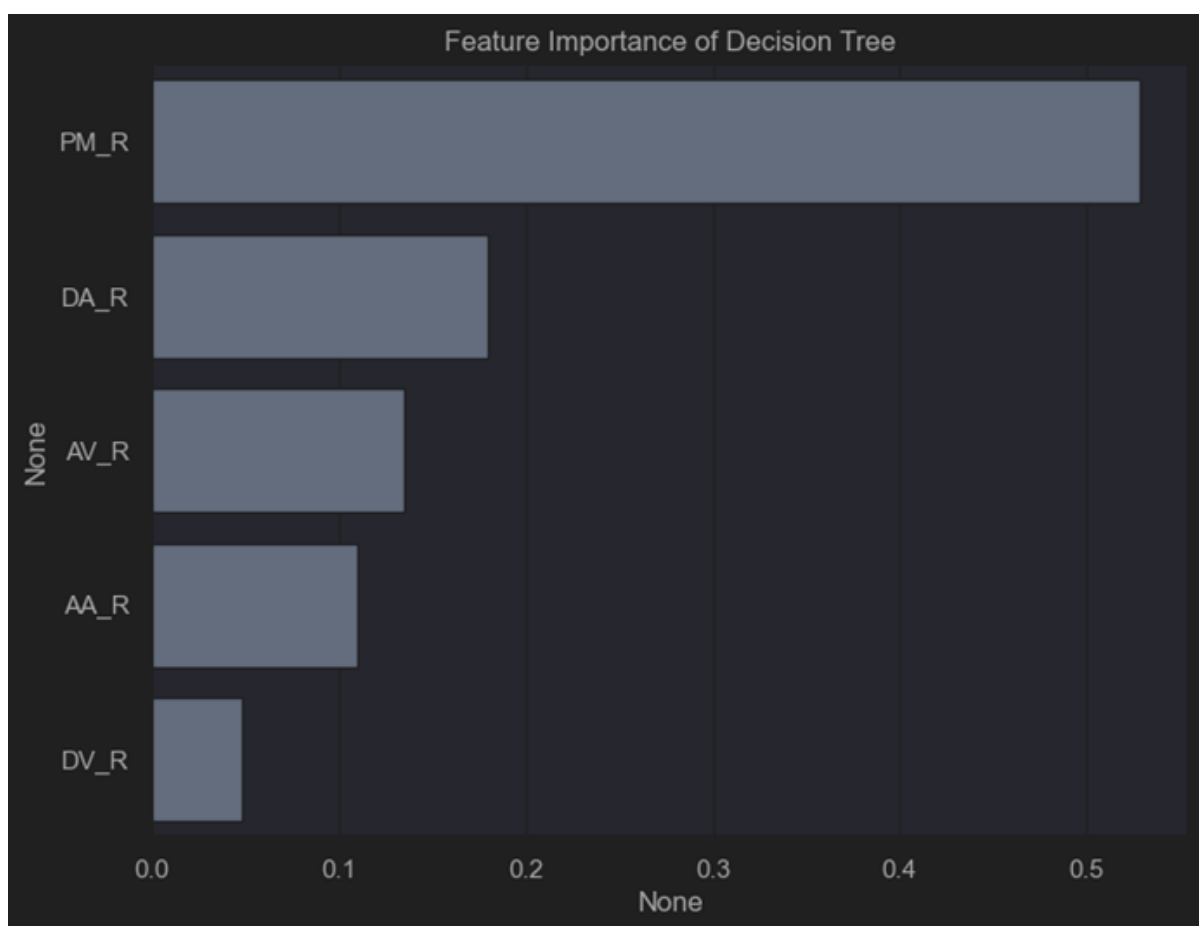Figure 11. Confusion Matrix of LightGBM


Figure 12. Feature Importance

In Figure 12, the most significant feature influencing the welding process was identified as welding speed, followed by alternating voltage. These findings suggest that controlling these parameters more precisely could reduce the occurrence of defects.

The optimum value boundary for welding speed/ Alternating Voltage are determined where other variables are in average boundary.



Figure 13. Welding Speed and Alternating Voltage with respect to density

# Discussion

**Performance Analysis:**

The Decision Tree model showed the highest accuracy and interpretability, making it a suitable choice for real-time defect detection. However, XGBoost and LightGBM also performed well, offering high accuracy with advantages in computation speed and memory efficiency. These models are particularly useful when handling large datasets, as in this study.

**Implications for Welding Process Optimization:**

By integrating AI models into the welding process, real-time classification of weld quality becomes feasible. This integration can significantly reduce the need for manual inspections or destructive testing, thereby saving time and resources. Moreover, understanding the key parameters that influence weld quality allows for better control and optimization of the welding process.

**Challenges and Future Work:**

While the models performed well, there are challenges in collecting diverse and comprehensive datasets that cover various welding conditions and defect types. Future work should focus on expanding the dataset and refining the models to improve their robustness and accuracy. Additionally, exploring other machine learning algorithms and techniques could provide further improvements.
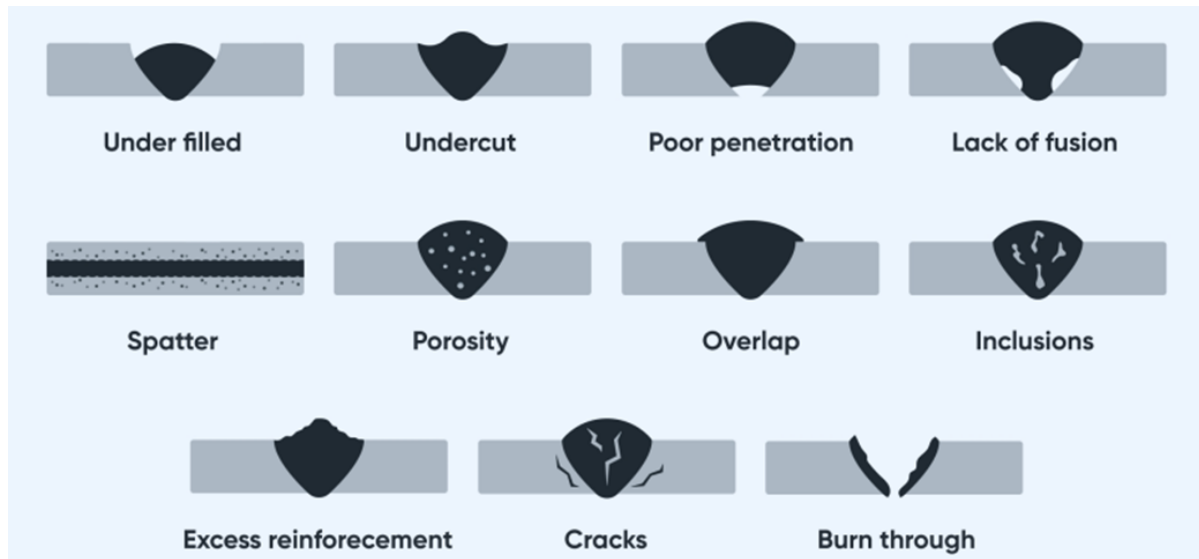


Figure 14. Defect types

# Conclusion

This study successfully demonstrated the application of AI in optimizing the welding process. By analyzing welding parameters using machine learning models, we achieved high accuracy in defect detection. The Decision Tree model, with its high interpretability and accuracy, was particularly effective for real-time defect detection.

Future efforts should aim at collecting more diverse data to cover different types of welding defects and refining the models for even higher accuracy. Implementing these AI models in the welding process can significantly enhance production quality and efficiency, leading to substantial cost savings and improved product reliability.

# References

- KAMP, Korea AI Manufacturing Platform. (2022). 용접 공정최적화 AI 데이터셋, 스마트제조혁신 추진단(㈜인터엑스), www.kamp-ai.kr