Prediction of Martensite Start Temperature of High-Strength Low-Alloy (HSLA) Steels with Given Different Compositions

REPORT SUBMITTED FOR

MSE497 – UNDER GRADUATE PROJECT II

By

Prashant Gupta – 220803
Mohit Jangid – 220664
Shikhar Chaudhary – 221008
Vibhanshu Choudhary – 221189

Under the guidance of

Prof. Amarendra Kumar Singh
Prof. Soumya Sridar

Indian Institute of Technology Kanpur

Acknowledgment

We would like to sincerely thank Professor Amarendra Kumar Singh and Professor Soumya Sridar for giving us the opportunity to work under their guidance on this project. Their consistent support and valuable feedback helped us maintain clarity and direction throughout the course of our work.

We are especially grateful for the time they dedicated to discussing our progress and assisting us in navigating challenges during the project. Their guidance played a key role in shaping our approach and kept us motivated during difficult phases.

Thank you once again for your unwavering support and encouragement throughout this journey.

Sincerely,
Prashant Gupta (220803)
Mohit Jangid (220664)
Shikhar Chaudhary (221008)
Vibhanshu Choudhary (221189)

Abstract

This report explores a data-driven framework for predicting the Martensite Start Temperature (MsT) of High-Strength Low-Alloy (HSLA) steels based on chemical composition. A curated dataset of experimentally validated steel compositions and corresponding MsT values was used to train various machine learning models. Emphasis was placed on meaningful data preprocessing, including outlier removal and normalization. Feature importance analysis and model comparison provided insights into compositional influences and the predictive capabilities of different algorithms.

May 3, 2025

Contents

1 Introduction

1.1 Property and Material of Interest

1.1.1 High-Strength Low-Alloy (HSLA) Steels

High-Strength Low-Alloy (HSLA) steels are a class of advanced engineering materials designed to provide superior mechanical properties—such as high strength, good toughness, and excellent weldability—while maintaining relatively low carbon and alloying element content.Its common application includes in automative chassis and body structures, pipelines, pressure vessels, heavy machinery and shipbuilding.

1.1.2 Martensite and Martensitic Transformation

Martensite is a hard, brittle phase formed in steels through a diffusionless, shear-type transformation from austenite (γ-Fe) upon rapid cooling (quenching). Unlike pearlite or bainite, which form via diffusion-controlled mechanisms, martensite forms almost instantaneously at temperatures below the Martensite Start (Ms) Temperature.

Key Features of Martensite:

- **High Hardness & Low Ductility:** Due to its body-centered tetragonal (BCT) crystal structure, martensite is extremely hard but brittle.
- **Transformation Mechanism:** The shift from FCC austenite to BCT martensite occurs via a displacive (shear) transformation, leading to a significant increase in strength.
- **Dependence on Cooling Rate:** Martensite only forms if the cooling rate exceeds the critical rate required to suppress diffusion-based transformations (e.g., pearlite or bainite).

1.2 Martensite Start (Ms) Temperature

1.2.1 Definition of Ms Temperature

The Martensite Start (Ms) Temperature is the critical temperature at which austenite begins transforming into martensite upon rapid cooling. Below this temperature, the martensitic transformation continues until the Martensite Finish (Mf) temperature is reached.

1.2.2 Importance of Knowing Ms Temperature

Understanding and predicting the Ms temperature is crucial for several reasons.

a) Phase Transformation Control

- The Ms temperature determines the onset of martensitic transformation, which directly influences the final microstructure and mechanical properties.
- If the Ms temperature is too low, the retained austenite may remain, reducing hardness and dimensional stability.

b) Heat Treatment Optimization

- In quenching processes, knowing the Ms temperature helps in selecting appropriate cooling rates to avoid cracking or distortion.
- Tempering treatments can be adjusted on the basis of the Ms temperature to achieve the desired ductility-toughness balance.

1.2.3 Understanding Ms Temperature Through CCT Diagrams

A Continuous Cooling Transformation (CCT) Diagram illustrates the phase transformations in steel under different cooling rates. The Ms temperature appears as a horizontal line, indicating the start of martensite formation.

Examples:

- **Slow Cooling:** Forms pearlite or bainite before reaching Ms.
- **Fast Cooling (Quenching):** Suppresses diffusion-based transformations, leading to martensite formation below Ms.

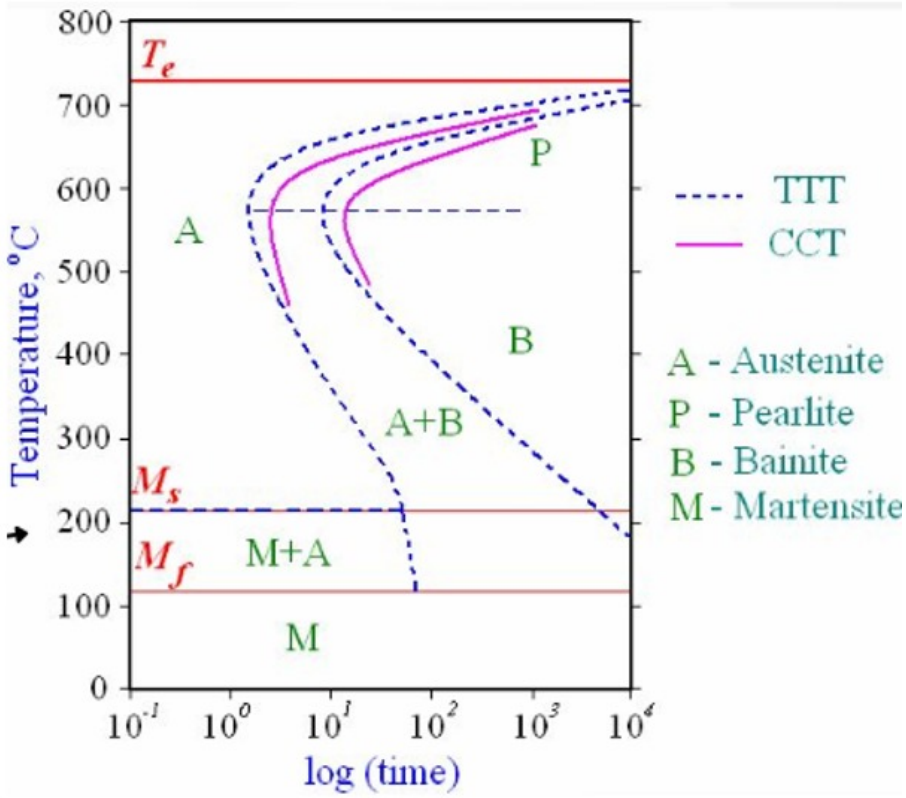


Figure 1: CCT and TTT diagrams showing transformation products and Ms/Mf temperatures.

2 Data Preprocessing and Data Curation

The accuracy and reliability of any prediction model depend heavily on the quality of the input data. In our study, which focuses on predicting the martensite start temperature (Ms) of high-strength low-alloy (HSLA) steels using data-driven methods, proper data cleaning and preparation were crucial to ensure trustworthy and useful results

2.1 1. Data Source and Theoretical Understanding

The dataset employed in this work was obtained from the MsT Predictor, developed and maintained by Northwestern University, and is available at: `http://info.eecs.northwestern.edu/MsTpredictor`

The data comprises a curated collection of steel compositions and their corresponding MsT values, and draws upon an extensive body of experimental metallurgical studies. The theoretical and experimental basis for this database is discussed in the IEEE publication:

"A Machine Learning Framework for Martensite Start Temperature Prediction in Steels" (IEEE Xplore).

Experimental Methodology for MsT Measurement

The martensite start temperature (MsT) is typically measured using a quenching dilatometer, which tracks the change in length or volume of a steel sample as it cools from the austenitic phase. The transformation from austenite to martensite leads to a measurable volume expansion, and the temperature at which this change begins is recorded as the MsT.

This technique has been widely used in metallurgy for decades. One significant compilation of such data is by Sourmail et al., who reported MsT values (in Kelvin) along with steel compositions (in weight percent) for over 1091 steels. The dataset used in this project is based on such experimentally obtained values, later processed and formatted for machine learning applications.

2.2 2. Exploratory Data Analysis and Outlier Elimination

Carbon (C)	Manganese (Mn)	Silicon (Si)
Chromium (Cr)	Nickel (Ni)	Molybdenum (Mo)
Vanadium (V)	Cobalt (Co)	Aluminum (Al)
Tungsten (W)	Copper (Cu)	Niobium (Nb)
Nitrogen (N)	Boron (B)	Titanium (Ti)

Table 1: Alloying elements used in the model

Upon acquisition of the raw dataset, an analysis was performed to examine the distribution and variability of the alloying elements.

A prominent observation was the presence of multiple entries with anomalously high Nickel concentrations, significantly deviating from the compositional norms of conventional HSLA steels. While Nickel does play a critical role in stabilizing austenite and

reducing MsT, concentrations exceeding ∼5 wt% are uncommon in practical alloy design for HSLA applications.

This curation step ensured that the retained dataset remains metallurgically meaningful and statistically robust.

2.3 3. Data Normalization

To prepare the data for model training, **StandardScaler** from the `scikit-learn` library was applied to all features to achieve z-score normalization. This step was essential for enabling efficient optimization.

3 Modeling Framework

This section outlines the step-by-step methodology adopted for building predictive models for estimating the Martensite Start Temperature (MsT) of steels from their alloying composition. Our modeling pipeline incorporated a wide variety of machine learning techniques ranging from linear models to complex ensemble learning and hybrid deep learning approaches.

3.1 Models Explored

A variety of regression models were explored to capture the complex relationships between the alloy compositions and martensite start temperature (Ms). These include:

- **Linear Models:**
  - Linear Regression – used as a baseline model.
- **Tree-based Models:**
  - Decision Tree Regressor – simple and interpretable tree model.
  - Random Forest Regressor – ensemble of decision trees to reduce variance.
  - Gradient Boosting Regressor (GBR) – boosted trees for improved performance.
  - XGBoost Regressor – optimized implementation of gradient boosting, tuned using both Grid Search CV and Optuna for hyperparameter optimization.
- **Kernel-based Models:**
  - Support Vector Regressor (SVR) with RBF kernel – effective in high-dimensional spaces.
- **Neural Network Model:**
  - Artificial Neural Network (ANN) – implemented using `MLPRegressor`, capable of capturing complex nonlinear patterns.
- **Hybrid Models (Ensemble and Stacking):**
  - Weighted average and stacking of XGBoost and ANN.
  - Full stacked ensemble combining XGBoost, MLP, GBR, and Extra Trees with a Linear Regression meta-learner for final prediction.

3.2 Ensemble Learning

Ensemble learning is a machine learning paradigm where multiple models, often referred to as *base learners* or *weak learners*, are combined to solve a particular problem and improve performance compared to individual models. The core idea is that a group of diverse models can provide a more robust and accurate prediction than any single model alone, as they can capture different patterns in the data.

There are several common ensemble methods:

- **Bagging (Bootstrap Aggregating):** Reduces variance by training multiple models on random subsets of the data (e.g., Random Forest).
- **Boosting:** Reduces bias by sequentially training models where each model corrects the errors of its predecessor (e.g., Gradient Boosting, XGBoost).
- **Stacking:** Combines multiple models using a meta-model that learns how to best integrate their predictions.

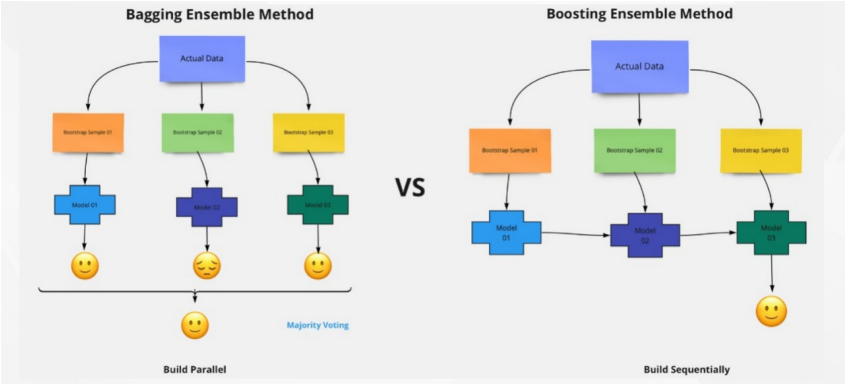


Figure 2: Illustration of common ensemble learning methods

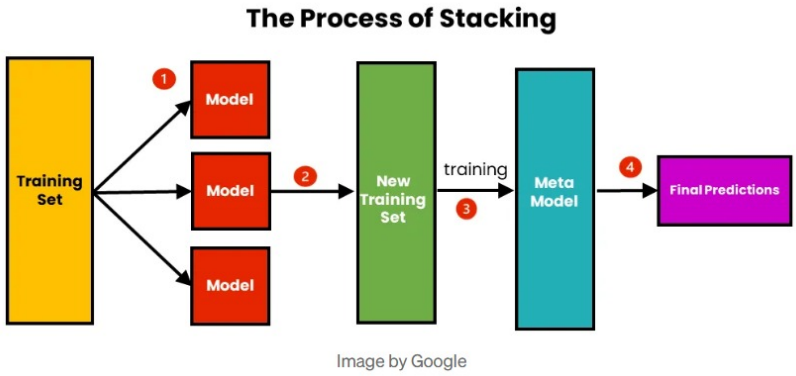


Figure 3: Illustration of common ensemble learning methods

3.3 Hyperparameter Optimization

To maximize model accuracy, various hyperparameter tuning techniques were applied:

- **Grid Search CV:** Used for exhaustive tuning of XGBoost, Random Forest, and Gradient Boosting models.
- **Optuna (Bayesian Optimization using TPE):**
  - Improved XGBoost accuracy by fine-tuning:
    * `learning_rate`
    * `gamma`
    * `max_depth`
    * `colsample_bytree`, etc.

3.4 Model Evaluation

All models were trained using an **80-20 train-test split**. The performance was evaluated using the following metrics:

- **R^2 Score:** Measures the proportion of variance in the dependent variable that is predictable from the independent variables.
- **Mean Absolute Error (MAE):** Represents the average absolute difference between predicted and actual values.

3.5 Final Model: Stacked Ensemble Approach

The final predictive model for estimating the martensite start temperature (Ms) is built using a **stacked ensemble learning technique**. This method integrates the predictions of multiple diverse models to enhance overall performance, reduce variance, and better capture complex patterns in the data.

3.5.1 What is Stacking?

Stacking (Stacked Generalization) is a type of ensemble learning where multiple base models (called *level-0 learners*) are trained on the same dataset. Then, another model (called the *meta-learner* or *level-1 model*) is trained to combine their predictions in an optimal way.

- Each base model independently learns from the training data.
- Their predictions (on unseen data or via cross-validation) become inputs for the meta-learner.
- The meta-learner captures relationships among the base models' outputs and makes the final prediction.

3.5.2 Why Stacking?

- Different types of models capture different aspects of the data.
- Combining tree-based models (good at capturing structure and interactions) with neural networks (good at non-linear patterns) provides complementarity.
- The meta-learner smooths out biases from the individual models and leads to improved generalization.

3.5.3 Base Models (Level-0 Learners)

- **XGBoost Regressor:** A powerful boosting algorithm that builds trees sequentially, where each new tree corrects errors from the previous ones. Excellent for handling structured/tabular data and capturing complex feature interactions.
- **MLP Regressor (Artificial Neural Network):** A multi-layer feedforward neural network that learns non-linear mappings from inputs to outputs. Ideal for modeling highly complex and abstract relationships in data.
- **Gradient Boosting Regressor (GBR):** A boosting model often more robust to overfitting than XGBoost for smaller datasets. Focuses on minimizing loss functions using stage-wise additive models.
- **ExtraTrees Regressor:** An ensemble of unpruned decision trees built on randomly selected features and splits. Offers low variance and complements boosting models with its different learning paradigm.

3.5.4 Meta-Learner (Level-1 Model)

A **Linear Regression** model is used as the meta-learner. It learns to weigh the outputs of the base models and combine them linearly to make the final prediction. Chosen for its simplicity and interpretability, ensuring the stacking doesn't overfit at the top level.

3.5.5 Benefits of the Stacked Ensemble

- **Robust Performance**: Combines multiple learning biases, reducing overfitting.
- **Improved Accuracy**: Aggregates the strengths of individual models for better predictive power.
- **Generalization**: Works well on unseen data by minimizing both bias and variance.
- **Flexibility**: Different models can be optimized independently and later combined without retraining the whole pipeline.

3.5.6 Why It Is the Final Model

After comparing many individual models and hybrid combinations, this stacked model was selected based on:

- High R^2 score (explains the most variance in Ms),
- Low Mean Absolute Error (minimal average deviation from true values),

making it both accurate and stable.

4 Results and Discussion

4.1 Initial Results and Feature Importance Interpretation

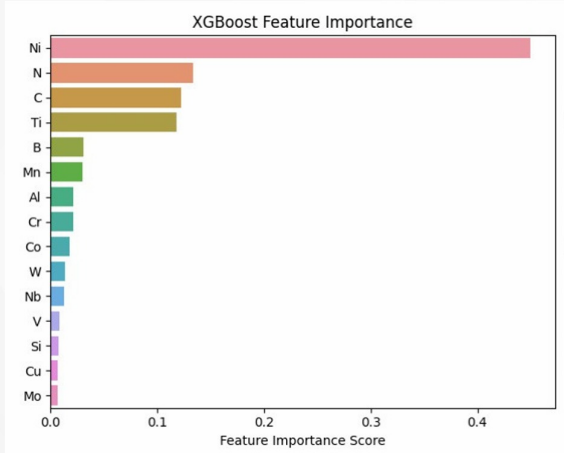
The initial performance of the predictive models is summarized in Table 2. The models were evaluated using Mean Absolute Error (MAE) and the Coefficient of Determination (R^2) to assess accuracy and generalization performance.

Table 2: Initial Model Performance Metrics

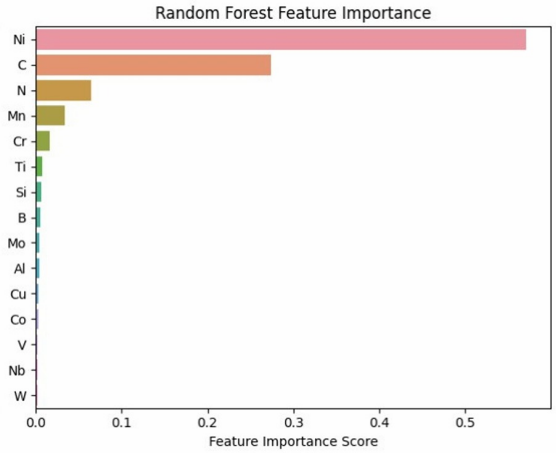
Sr. No	**Model**	**MAE (K)**	R^2 **Score**
1	XGBoost	19.0677	0.9354
2	ANN	20.1425	0.9339
3	XGBoost + ANN	19.1912	0.9342
4	Random Forest	22.5300	0.9013
5	Decision Trees	28.8832	0.8830
6	Gradient Boosting	21.2091	0.9035
7	Linear Regression	33.8922	0.8242
8	SVM	31.7009	0.7861

4.2 Issue with Feature Importance Interpretation

Among the trained models, **XGBoost** and **XGBoost + ANN** achieved the highest R^2 scores, indicating strong predictive performance. However, an analysis of their feature importance rankings revealed a common and concerning trend: both models identified **Nickel (Ni)** as the most influential feature in determining the martensite start temperature (MsT).



(a) XGBoost Feature Importance



(b) Random Forest Feature Importance

This result contradicts established metallurgical understanding. While Nickel is known to lower the MsT due to its austenite-stabilizing nature, its effect is generally less pronounced compared to elements like **Carbon**, which has a significantly stronger and nonlinear influence on the transformation.

The overemphasis on Nickel suggested the presence of bias within the dataset—likely caused by outlier compositions with abnormally high Nickel content that did not reflect typical HSLA steel chemistry. To resolve this issue, the dataset was refined by identifying and removing these outliers.

4.3 Model Performance After Dataset Refinement

Following the removal of high-Nickel outliers and refinement of the dataset to better reflect realistic HSLA steel chemistries, the models were retrained. The revised results are shown in Table 3.

Table 3: Model Performance After Dataset Refinement

Model	R^2 **Score**	**MAE (K)**
Linear Regression	0.8560	26.27
Decision Tree	0.7803	26.71
Support Vector Regressor (SVR)	0.8799	19.58
Random Forest	0.8558	20.00
Gradient Boosting	**0.8968**	**18.18**
XGBoost (GridSearchCV)	0.8799	18.89
XGBoost + Optuna	0.8750	19.15
MLP (ANN)	0.8408	21.31
XGBoost + ANN (Stacked)	**0.8965**	**18.08**
XGBoost + ANN (Weighted Avg)	0.8954	18.28
Stacked Regressor (XGB + MLP + GBR + ET)	**0.9018**	**17.06**

To further understand the improved prediction quality, feature importance plots were analyzed for the top 2 performing models. These plots, shown below, now correctly indicate **Carbon (C)** as the most influential feature—aligning with metallurgical expectations.

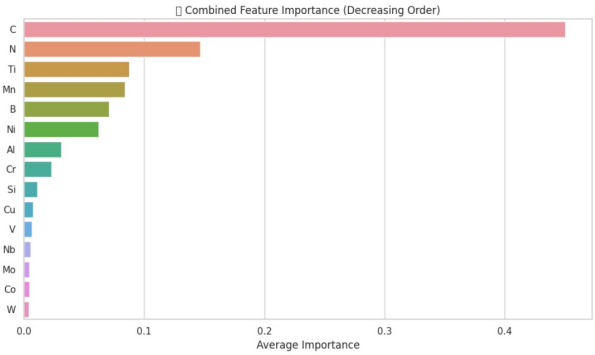


Figure 5: Stacked Regressor (Final)

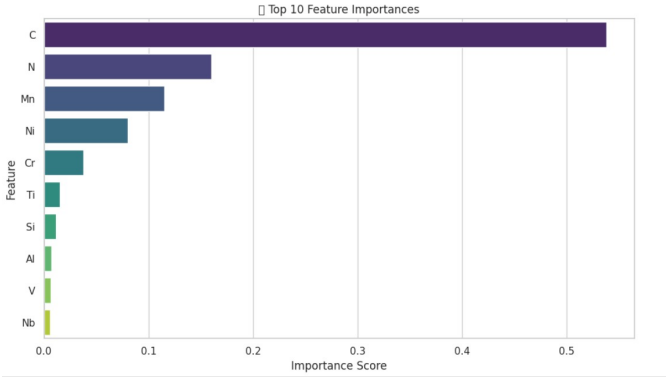


Figure 6: Gradient Boosting

These visualizations confirm that post-refinement, the models not only improved in numerical accuracy but also became more aligned with physical understanding, highlighting the importance of domain-aware data curation in machine learning applications within materials science.

4.4 Comparative Evaluation and Observations

A comparative analysis of model performance before and after dataset refinement revealed key insights. Although the removal of high-Nickel outliers slightly reduced the R^2 scores for some models, the overall integrity and interpretability of the models improved significantly.

Most notably, the feature importance distributions post-refinement were far more aligned with metallurgical understanding. Unlike the initial results—where **Nickel (Ni)** was incorrectly ranked as the most influential factor—the refined models correctly identified **Carbon (C)** as the dominant element influencing the martensite start temperature (MsT), consistent with domain knowledge.

5 Conclusion

In this study, we developed and evaluated multiple regression models to predict the martensite start temperature (Ms) based on material composition. Among all models tested, the final stacked ensemble — combining XGBoost, Multi-Layer Perceptron (MLP), Gradient Boosting Regressor (GBR), and Extra Trees Regressor — emerged as the most effective approach.

This ensemble leveraged the strengths of diverse learning paradigms, including decision trees, gradient boosting, and neural networks. A linear regression meta-learner was used to combine the predictions of these base models, resulting in a balanced and high-performing predictive framework. The stacked model achieved strong metrics across all evaluation scores, notably in terms of R^2 score - **0.9018** and MAE - **17.06K**, indicating both accuracy and generalization capability.

References

[1] IEEE Xplore. *Stacked Generalization-based Ensemble Deep Learning for Solar Irradiance Forecasting*. Available: `https://ieeexplore.ieee.org/abstract/document/8964188/references#references`

[2] S. van der Zwaag, W. Xu. *Prediction of Martensite Start Temperature using Artificial Neural Networks*. Available: `https://www.researchgate.net/profile/S-Zwaag/publication/27348305_Prediction_of_martensite_start_temperature_using_artificial_neural_networks/links/0c960533d2ac533c59000000/Prediction-of-martensite-start-temperature-using-artificial-neural-networks.pdf`

[3] Northwestern University. *MsTpredictor - Martensite Start Temperature Prediction Tool*. Available: `http://info.eecs.northwestern.edu/MsTpredictor`

6 Appendix

Add tables, graphs, or additional code here.

- Sourmail et al., IEEE Xplore, "A Machine Learning Framework for Martensite Start Temperature Prediction in Steels".

- XGBoost Documentation – https://xgboost.readthedocs.io/
- Code File -https://www.kaggle.com/code/gprashant2203/ugp-de-practise-final/notebook
- Raw Data -https://view.officeapps.live.com/op/view.aspx?src=http%3A%2F%2Finfo.eecs.northwestern.edu%2FMsTpredictor%2Fstatic%2FMsT_Database_2019.xlsx%3Fv%3D4e9197c4e8663c689f5050c58d187651&wdOrigin=BROWSELINK
- Refined dataset -'https://docs.google.com/spreadsheets/d/1IdARML8SvKMpYTmjd_m3n7RgKAlvk9YFlVUm83Yvdb0/edit?gid=0#gid=0