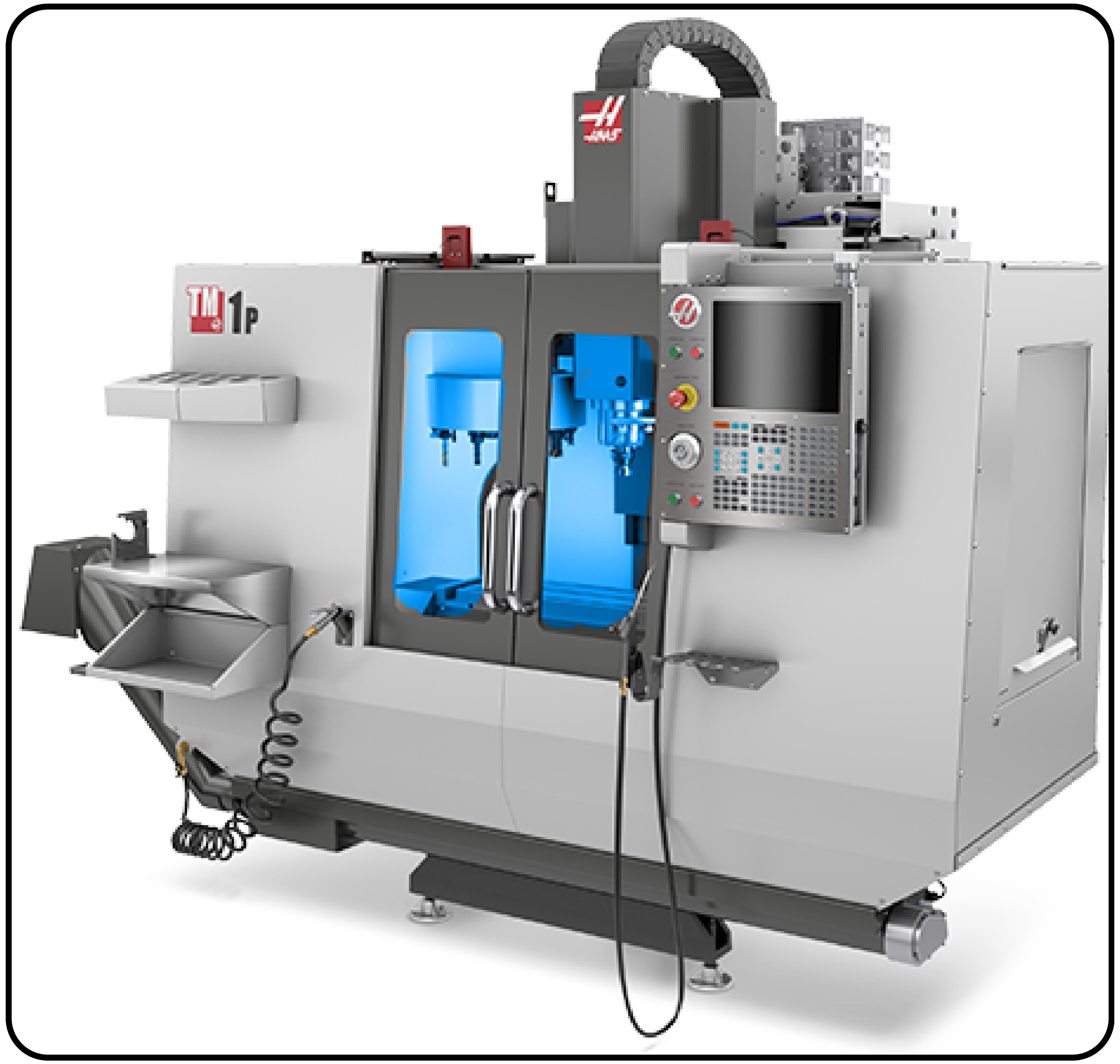
ME-217 Industrial Data Analytics

IDA MINI PROJECT

Final Report

Team
Varad S. Pendse
Tejas Bhavekar
Y. Akshay Kumar



OVERVIEW OF THE PROJECT

In machining processes, tool wear is a critical factor that affects product quality, production efficiency, and operational costs. Monitoring and predicting tool wear helps ensure **timely interventions, reducing downtime and avoiding excessive wear** that could lead to defects or machine damage.

The objective of this project is to develop a predictive model to estimate **flank wear (VB)** in cutting tools using data collected from various sensors. The system aims to accurately predict flank wear, enabling proactive maintenance to minimize tool failure and enhance manufacturing efficiency.

DATA DESCRIPTION

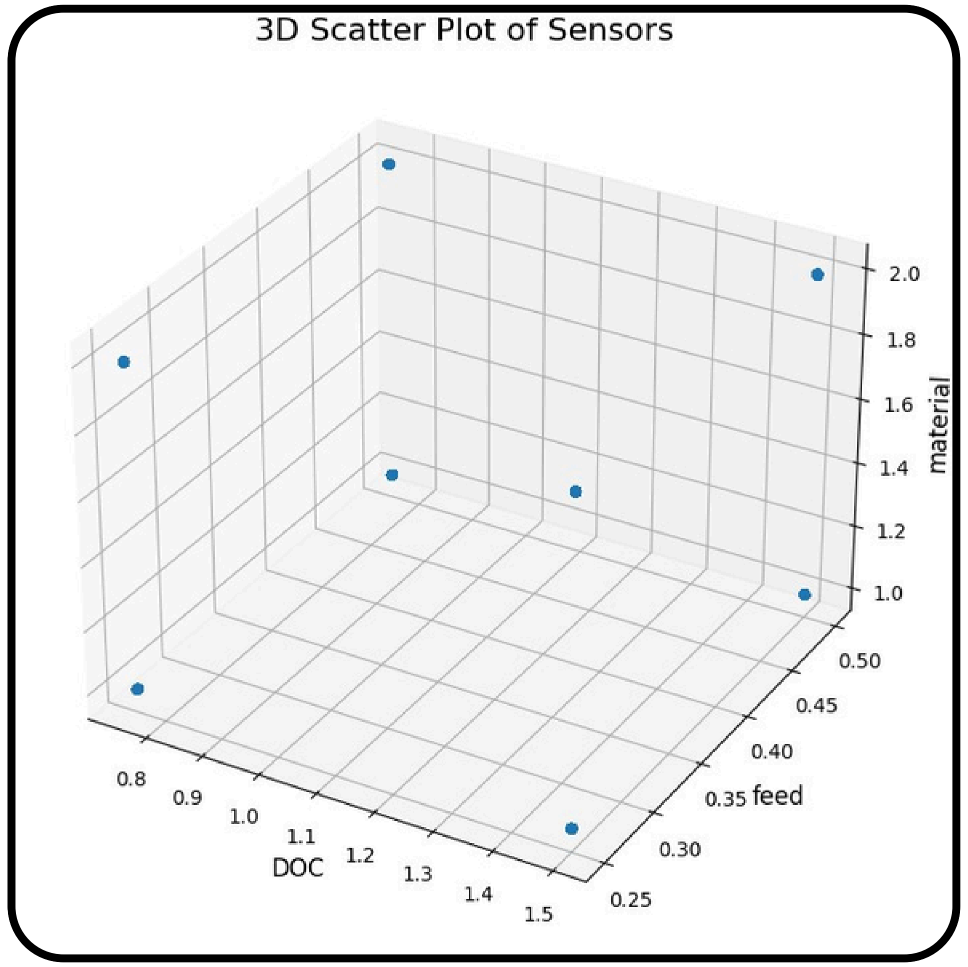
This dataset offers operational insights into a CNC milling machine operating under 16 different cases of varying conditions, specifically depth of cut (DOC), feed rate, and material type. These cases represent 8 unique parameter combinations, with each case tested over multiple runs, each run containing 9000 sensor readings. The available sensor data includes:

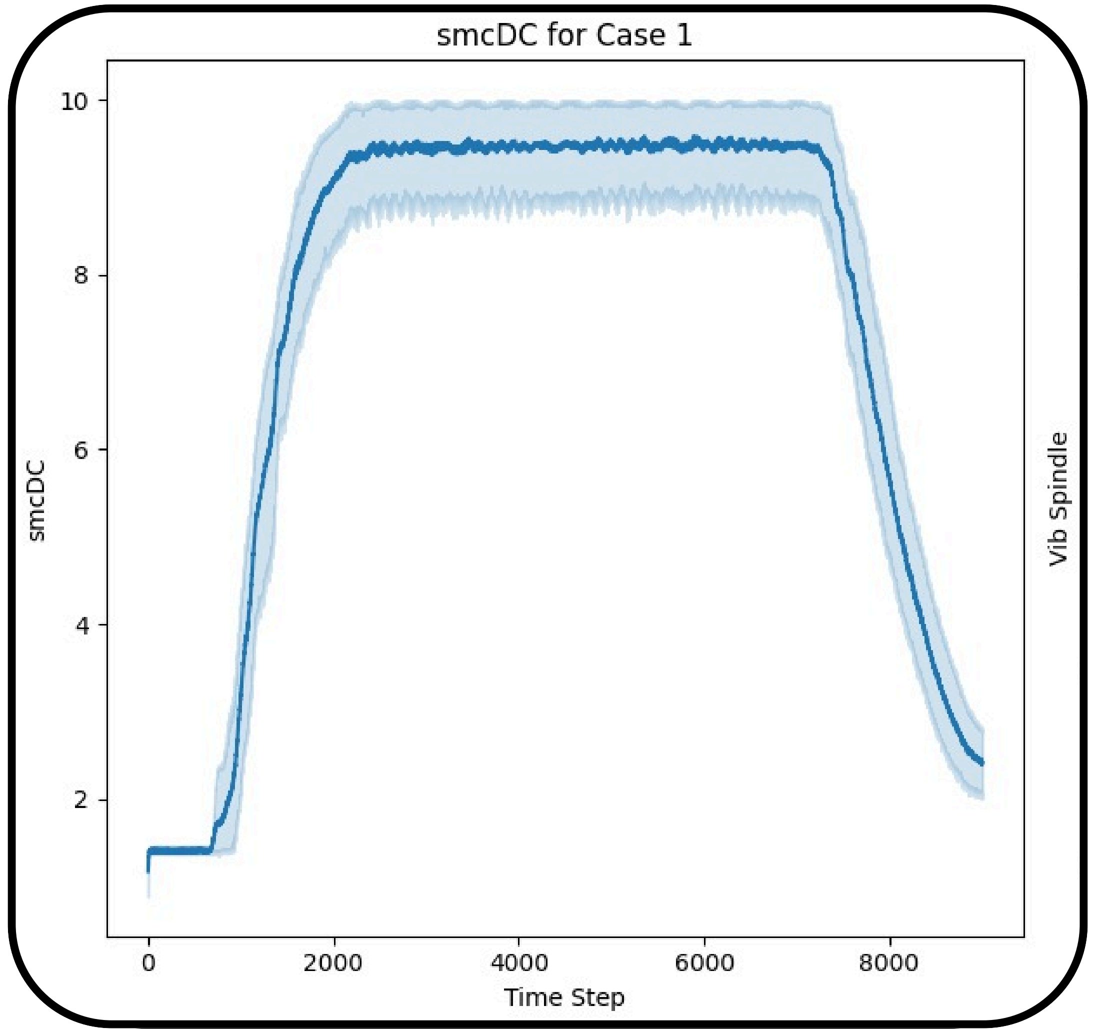
- **smcDC**: DC spindle motor current.
- **vib_spindle**: Vibration data from the spindle.

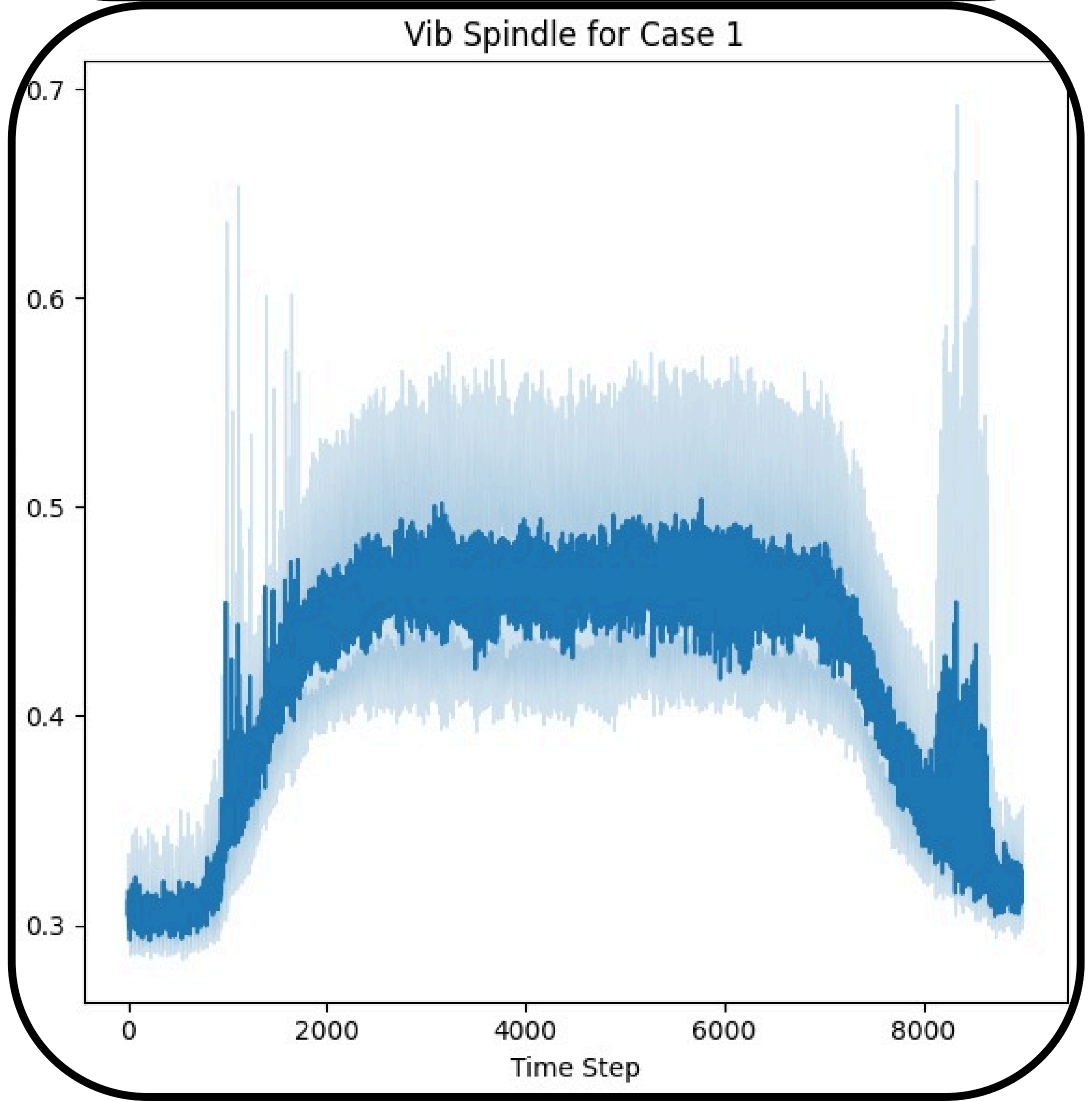
The primary goal is to identify optimal parameter settings to enhance runtime efficiency and improve outcomes by accurately predicting flank wear (VB). The dataset has been organized into:

- **Features**: Four operating conditions—Time, Depth of cut, Feed, and Material type along with the two sensor readings.
- **Label**: Flank wear (VB), which serves as the target variable for prediction.

Using this setup, we aim to develop a predictive model that can determine the optimal parameter configuration for minimized flank wear and improved machine performance.







SOLUTION

Data Preprocessing:

- The data is filtered based on specified conditions and standardized using StandardScaler. Signal processing techniques like rectification, low-pass filtering, and smoothing were applied to sensor data for noise reduction and feature enhancement.

Model Architecture:

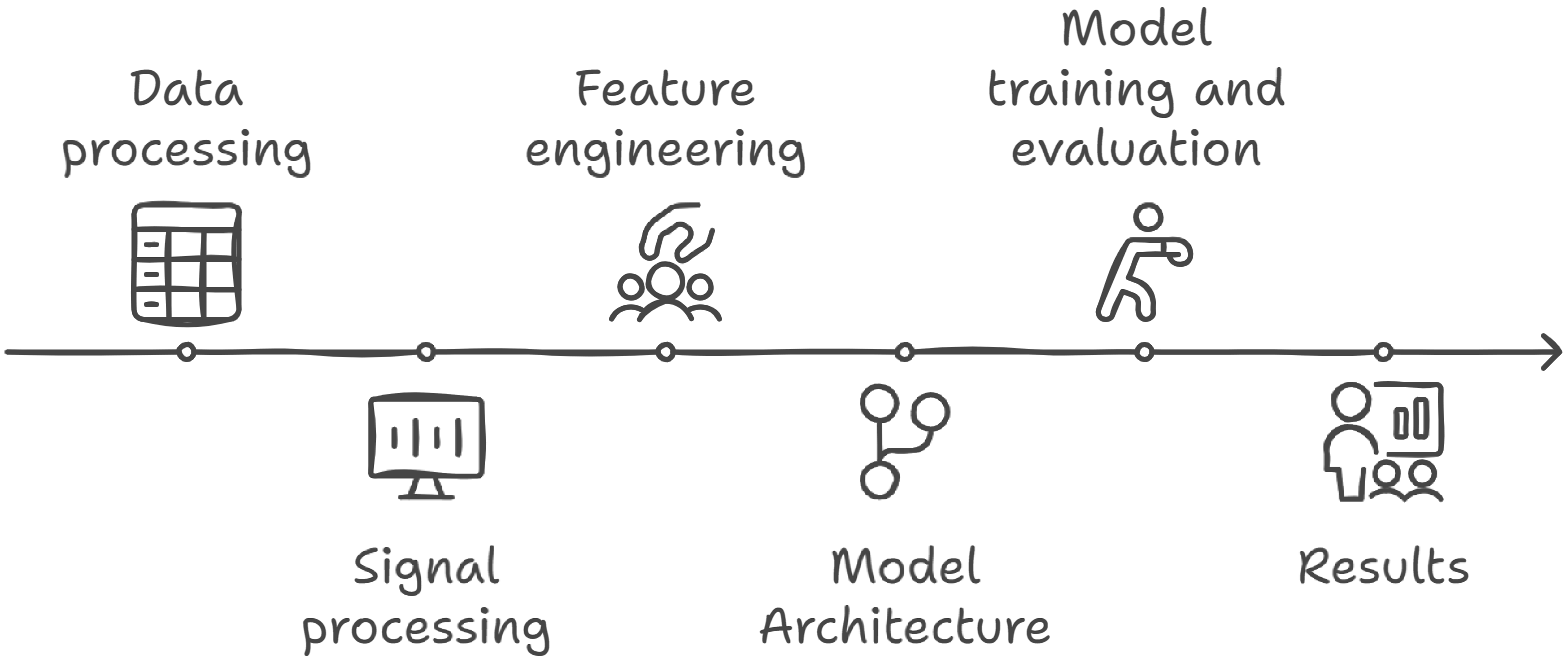
A multi-branch neural network was designed, incorporating:

- Operating Conditions are processed through dense layers.
- Sensor Data is processed through frequency-domain (FFT), time-domain features (mean, standard deviation, RMS), and Conv1D layers.
- Final Output: A regression model predicting flank wear (VB).

Model Evaluation:

- The model is evaluated using RMSE and visualized through a Predicted vs. Actual line plot, demonstrating the model's predictive accuracy.

Methodology



DATA PROCESSING

Data Selection

- The dataset consists of **sensor readings** (vib_spindle and smcDC) and **operating conditions** (e.g., time, DOC, feed, material).
- The goal is to predict flank wear (VB), which is influenced by sensor data and operating conditions.

Handling Missing Data

- Before applying standardization or any other processing steps, it's essential to ensure that there is no missing data in the dataset. Missing values can **skew** the model's training, so we employed **dropping all the missing values** from the dataset.

Data Filtering

- Filtered to keep **relevant features**: vib_spindle, smcDC, time, DOC, feed, material, and target VB.

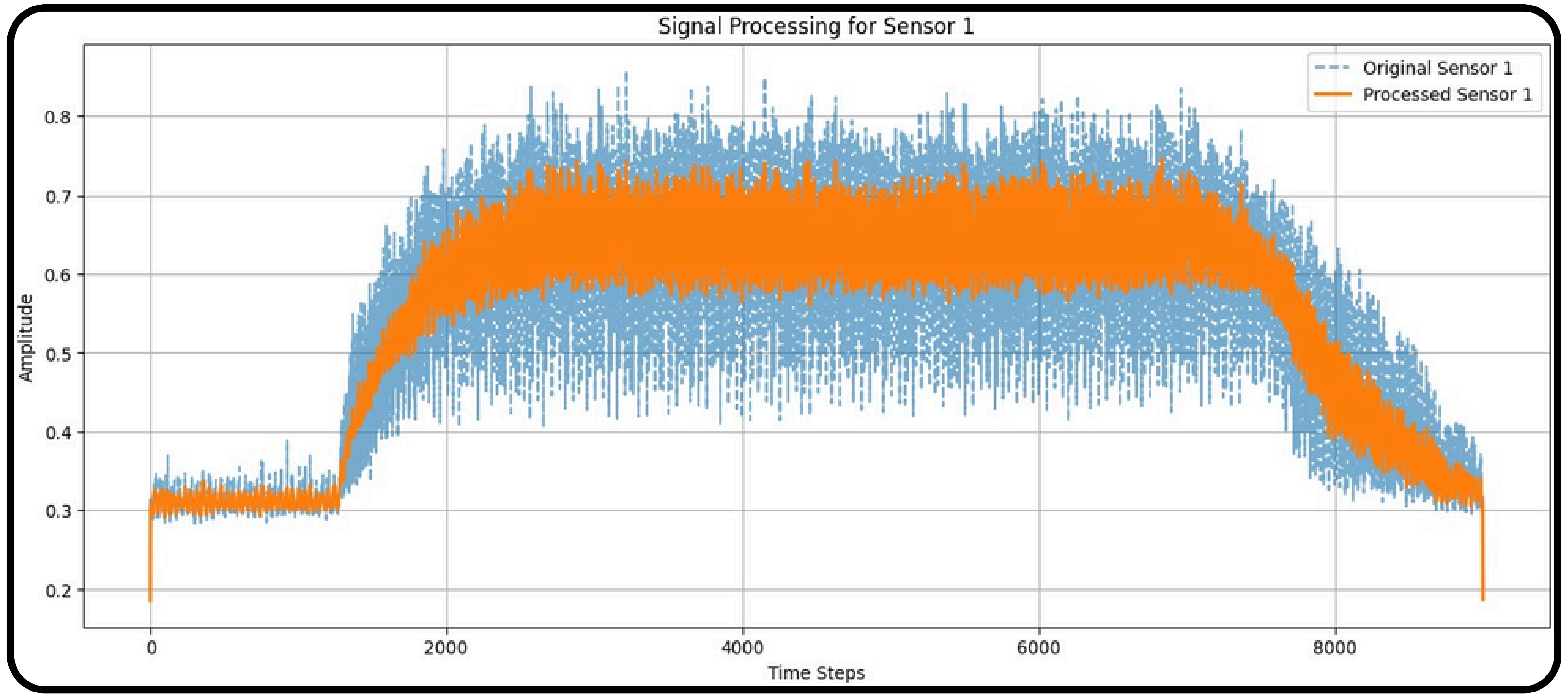
Standardization

- The sensor data columns are standardized using **StandardScaler** to ensure all features are on the same scale for model training.

SIGNAL PROCESSING

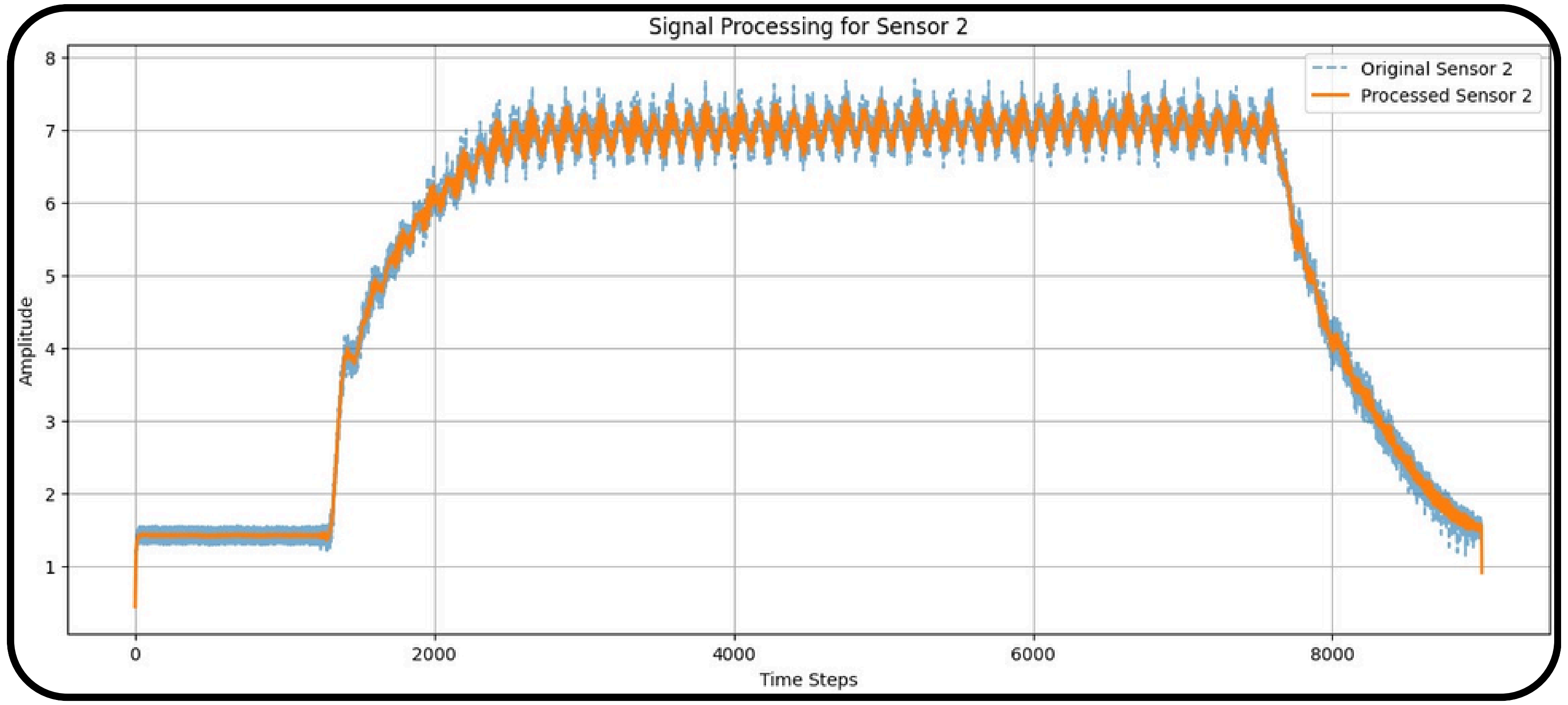
Methods Used

1. Rectification: Converts signal values to absolute, removing negative fluctuations for uniform analysis.
2. Low-Pass Filtering: Removes high-frequency noise using a Butterworth filter, retaining meaningful trends.
3. Smoothing: Applies a moving average, reducing minor fluctuations for a cleaner signal.



Advantages of Signal Processing

1. Enhances signal clarity by removing noise and irrelevant fluctuations.
2. Improves model performance by providing consistent, noise-free inputs.



FEATURE ENGINEERING

Operating Conditions:

- Extracted and processed operating conditions (e.g., DOC, feed, material), assumed constant per case, with higher importance (weighted 1.5x for Model 3).

Frequency-Domain Features:

- Applied Fast Fourier Transform (FFT) to sensor data to capture periodic signals indicative of machine wear patterns.

Time-Domain Features:

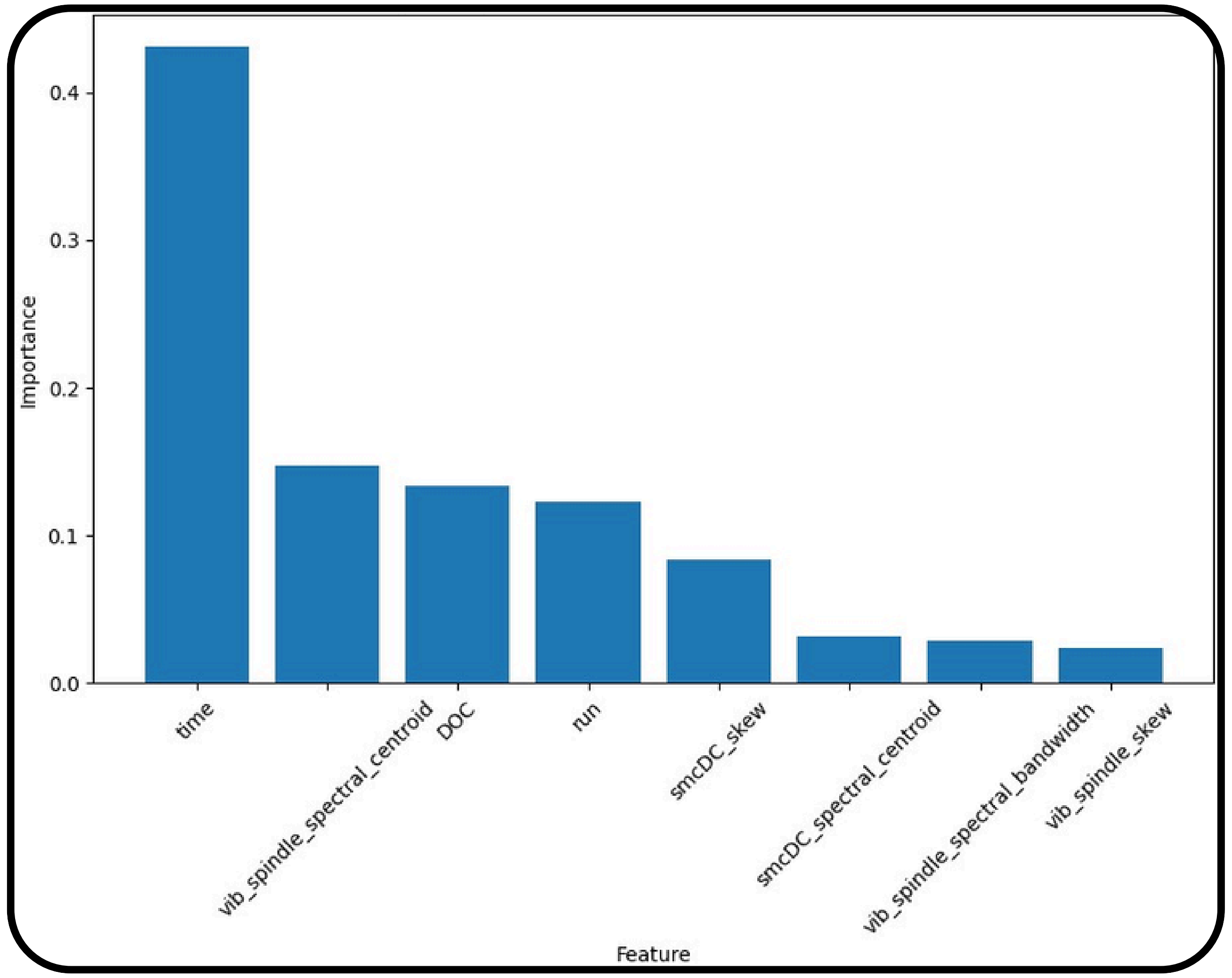
- Calculated key statistics from sensor data: mean, standard deviation, and RMS to capture time-dependent patterns.

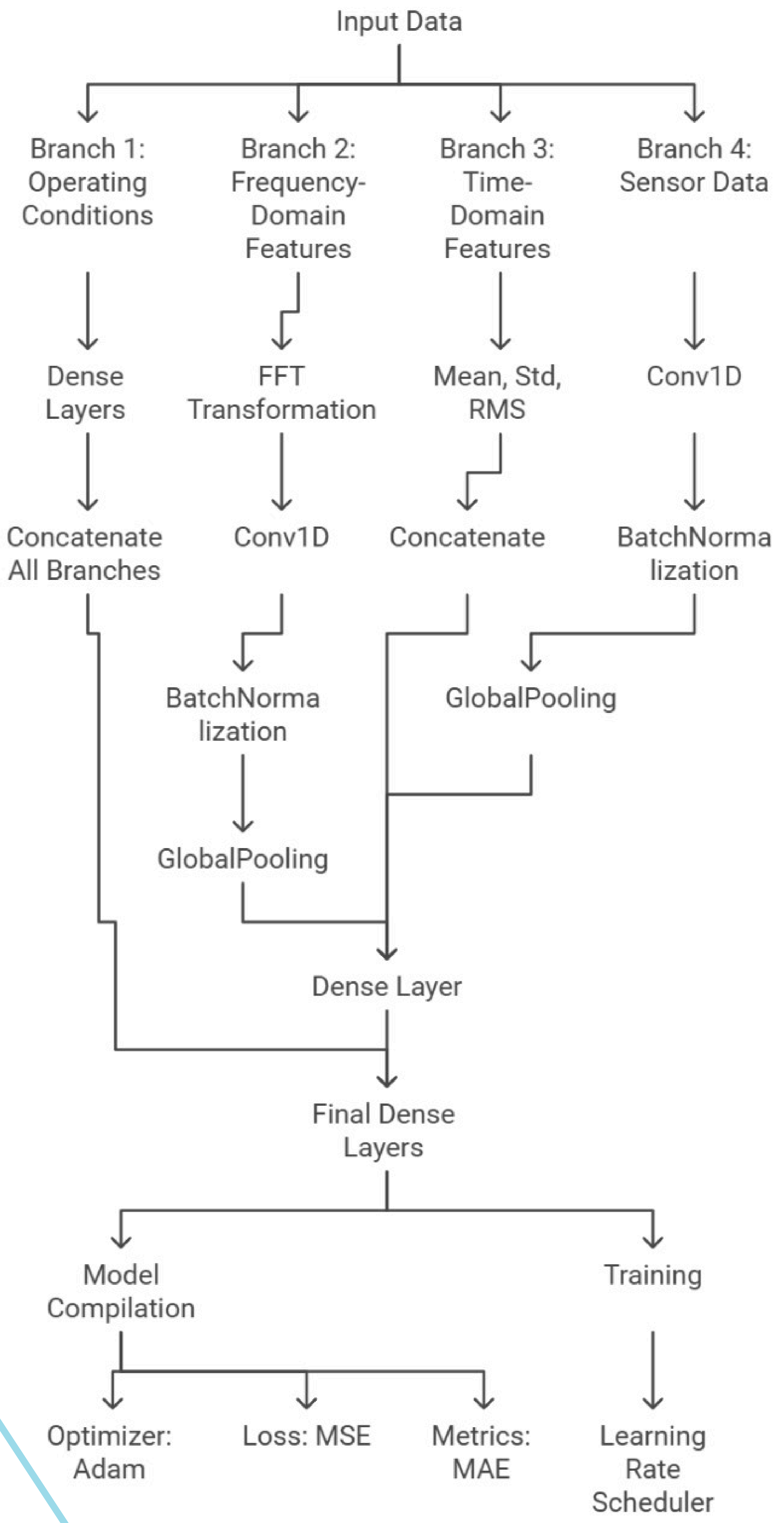
Feature Extraction (Model 3):

- Combined Conv1D layers for temporal patterns with dense layers for operating conditions, followed by concatenation of all feature outputs.

Dimensionality Reduction:

- Used Recursive Feature Elimination (RFE) to select the top 10 most important features using a RandomForestRegressor.





MODELLING IDEATION

MODEL 1

This model uses a multi-branch architecture to combine operating conditions and sensor data for enhanced predictive maintenance.

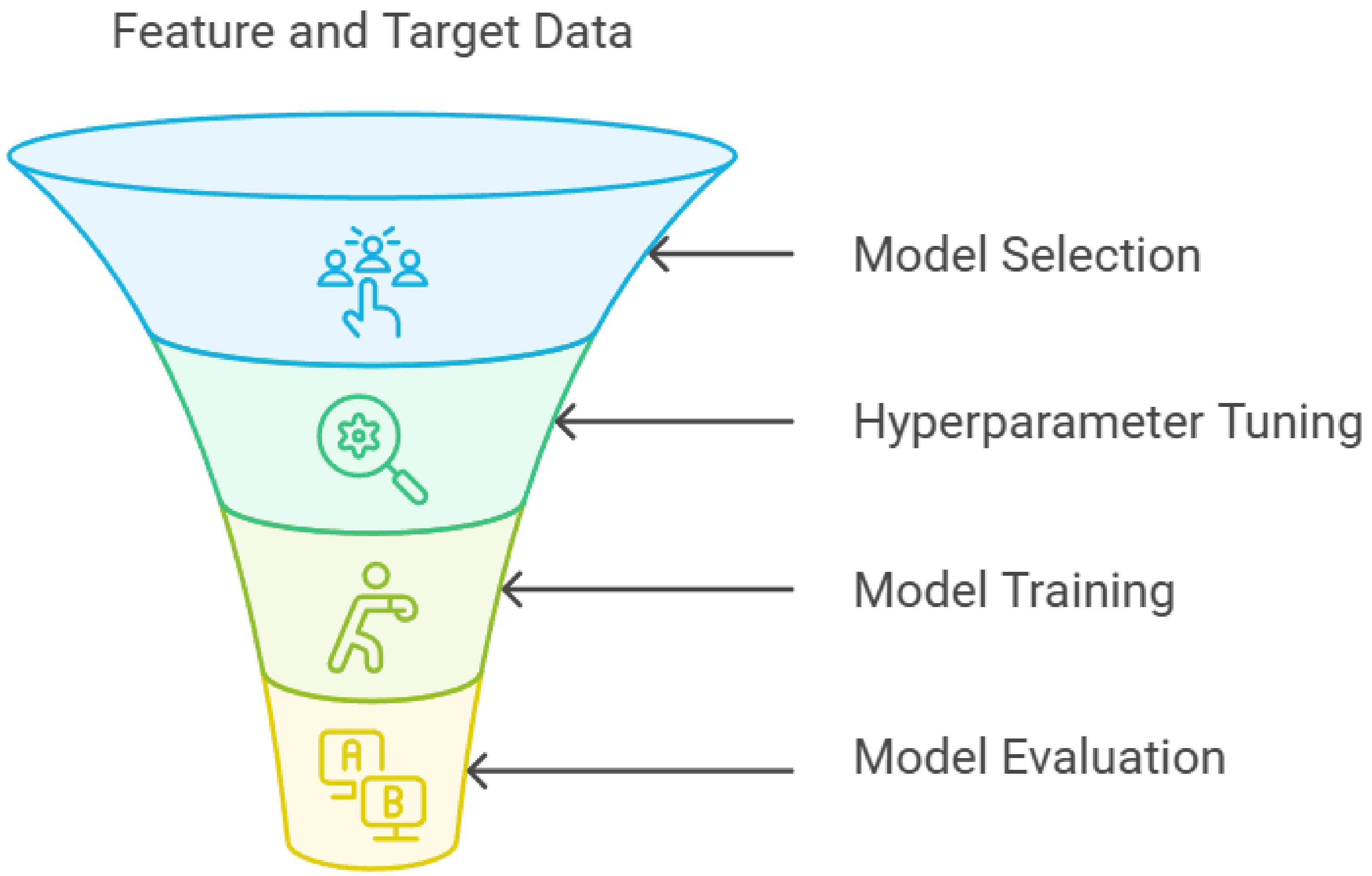
- Multi-Branch Architecture:
  - Operating conditions are processed using dense layers.
  - Frequency-domain features: FFT is applied to sensor data to detect periodic wear patterns.
  - Time-domain features: Mean, standard deviation, and RMS are extracted to capture time-dependent patterns.
  - Sensor data: Conv1D layers detect temporal patterns.

- Innovation:
  - Combines both frequency and time-domain features, alongside operating conditions, for a comprehensive model.
  - The multi-branch design ensures each feature type contributes optimally, improving failure prediction accuracy.

This approach improves sensitivity to machine faults by processing diverse feature domains independently and combining them effectively

MODELLING IDEATION

MODEL 2



This model leverages XGBoost Regressor, a gradient-boosted decision tree algorithm, for predicting flank wear (VB) based on machining process parameters and sensor data.

1. Input:
   - Features (machining parameters, sensor data) and target variable (flank wear).
2. Model Type:
   - XGBoost Regressor: Boosted decision trees to predict regression outcomes.
3. Hyperparameter Optimization:
   - Optimized using RandomizedSearchCV with 200 combinations, selecting key parameters like n_estimators, learning_rate, and max_depth.
4. Training:
   - 5-fold cross-validation with negative mean squared error as the scoring metric.
5. Evaluation:
   - Evaluated using RMSE (Root Mean Squared Error) and R^2 Score to assess model accuracy and explanatory power.

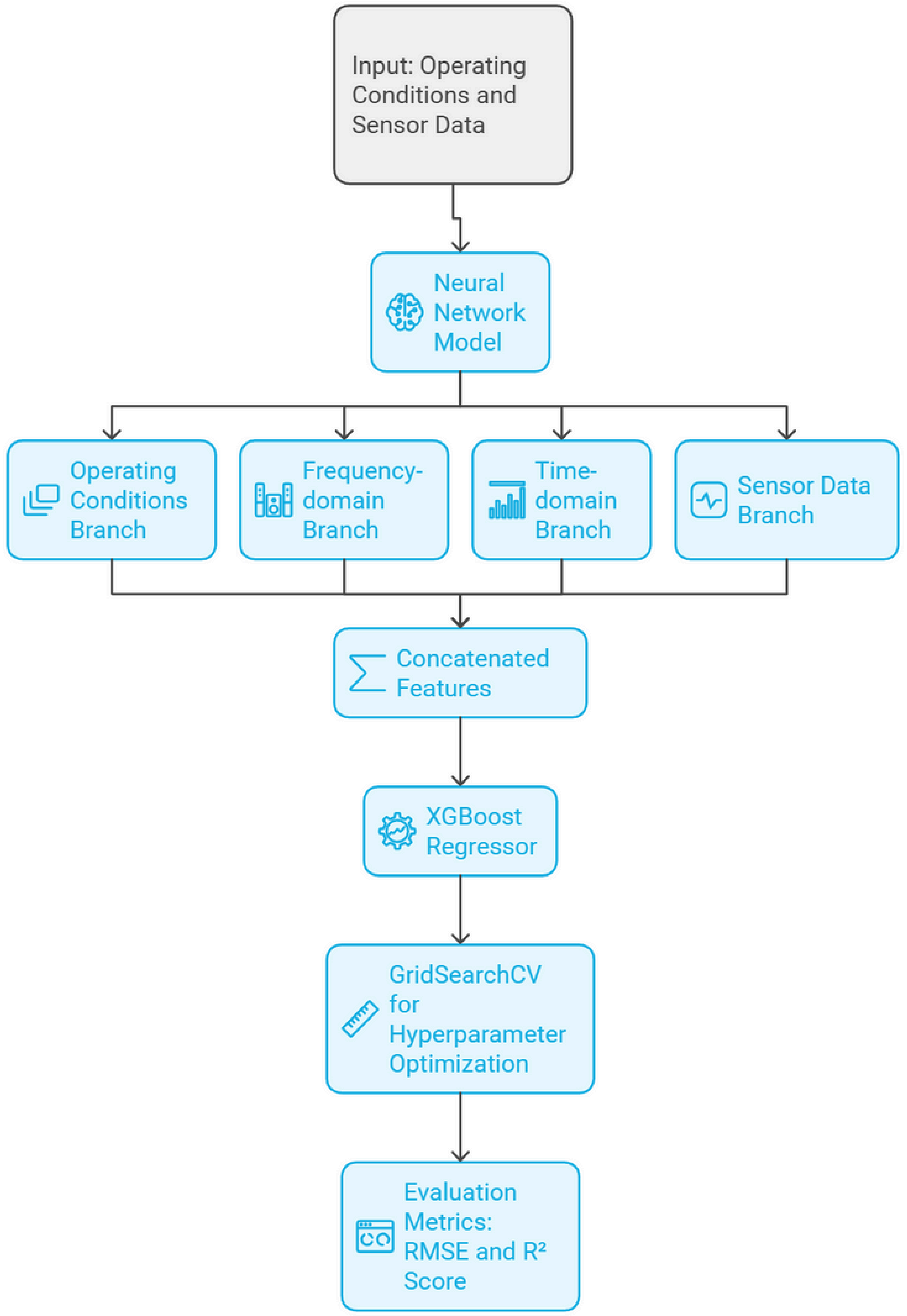
Innovation: The hyperparameter optimization via RandomizedSearchCV ensures a fine-tuned model, enhancing predictive accuracy for flank wear.

MODELLING IDEATION

MODEL 3

This approach combines deep learning for feature extraction with XGBoost for regression to predict flank wear (VB).

1. Feature Extraction (Neural Network):
   - Inputs: Operating conditions and sensor data.
   - Architecture:
     - Operating Conditions: Dense layers with higher weightage.
     - Frequency-domain: FFT for periodic pattern extraction.
     - Time-domain: Features like mean, standard deviation, RMS.
     - Sensor Data: Conv1D layers for temporal patterns.
   - Output: Concatenated features for regression.
2. XGBoost Regressor:
   - Features from the neural network are input into XGBoost Regressor.
   - GridSearchCV: Hyperparameter optimization via cross-validation.
   - Metrics: RMSE and R^2 Score to evaluate performance.
3. Novelty:
   - Hybrid Approach: Combines deep learning for feature extraction with XGBoost for enhanced accuracy.
   - Comprehensive Features: Integrates frequency-domain, time-domain, and sensor features for robust predictions.



MODEL TRAINING AND EVALUATION

Optimizer & Loss Function:

- Adam optimizer with a learning rate of 0.00075 and Mean Squared Error (MSE) loss function were used for models involving neural networks

Learning Rate Scheduler:

- ReduceLROnPlateau callback was used to decrease the learning rate when validation loss stagnated, allowing the model to refine its learning.

Training Process:

- Each of the models was trained for 100 epochs, with a batch size of 16, and the training was monitored using callbacks to optimize learning.

RMSE Calculation:

- Root Mean Squared Error (RMSE) was calculated to evaluate model performance and its ability to predict flank wear.

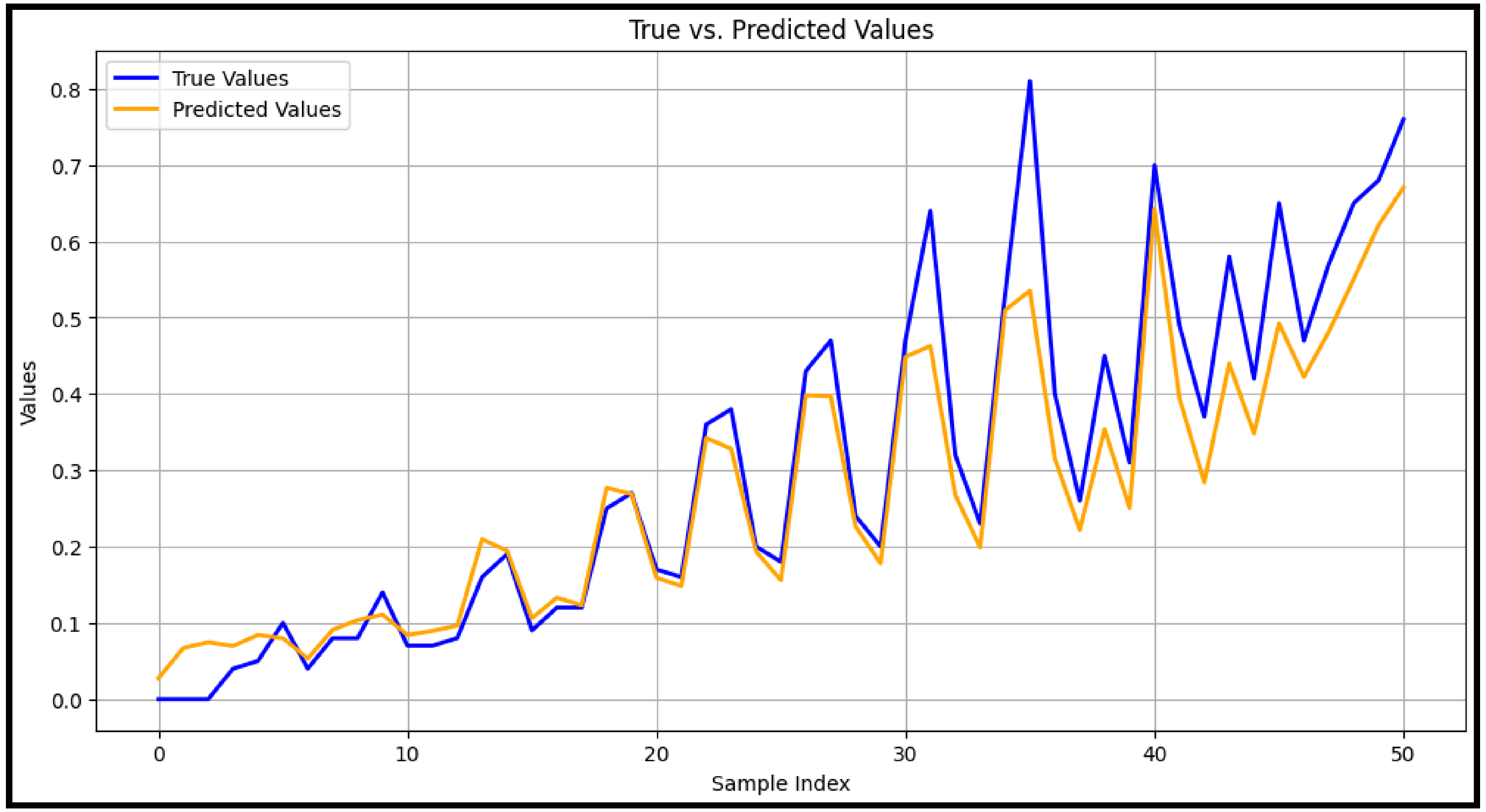
Predicted vs. Actual Comparison:

- A comparative line plot was used to visualize how close the predicted values are to the actual values, showing model accuracy. (shown on the next page).

RESULTS

Root Mean Squared Error (RMSE): Value: 0.0719 (using Model 1)

- The RMSE value of 0.0719 indicates that the model's predictions are generally close to the actual values, suggesting a good fit overall. The plot further highlights that the model performs consistently across most of the test data.
- Model 2 gave an RMSE of 0.10 and Model 3 gave an RMSE of 0.0751.



FUTURE WORK

- In the future, the model can be enhanced to better adapt to time-dependent and condition-specific patterns, improving its predictive accuracy for flank wear.
- With further tuning and the inclusion of additional features, it can effectively handle the variability in machining processes, ensuring more reliable predictions and more efficient maintenance schedules. This will contribute to reducing tool failure, optimizing operational efficiency, and ultimately lowering operational costs.

THANK YOU