FIRE WEATHER INDEX (FWI) PREDICTION MODEL

Milestone 3:

Evaluation and Optimization

Submitted by:

Himanshu Ramole

Infosys Springboard Mentor:

Praveen

Date: 08-01-2026

1. Introduction to Model Evaluation

After training and optimizing the Ridge Regression model in Milestone 2, the next critical step is to **evaluate its performance on data**.
Model evaluation helps determine how well the trained model generalizes beyond the training dataset and whether it can be reliably used for prediction.

In this milestone, the model was evaluated using:

- Mean Absolute Error (MAE)
- Root Mean Squared Error (RMSE)
- R^2 Score (Coefficient of Determination)
- Predicted vs Actual value visualization

The evaluation was primarily performed on the **test dataset**, as test performance reflects real-world predictive capability.

2. Evaluation Metrics Used

2.1 Mean Absolute Error (MAE)

Definition:
Mean Absolute Error measures the average magnitude of errors between predicted and actual values, without considering their direction.

Formula:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i |$$

Where:

- y_i= actual value
- $\hat{y}_i$= predicted value
- n= number of observations

Interpretation:

- Lower MAE indicates better model performance
- MAE treats all errors equally
- It is easy to interpret because it is in the same unit as the target variable (FWI)

Why MAE was used:

- Provides a clear understanding of average prediction error
- Less sensitive to outliers compared to RMSE

2.2 Root Mean Squared Error (RMSE)

Definition:
RMSE measures the square root of the average squared differences between predicted and actual values. It penalizes larger errors more strongly than MAE.

Formula:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Interpretation:

- Lower RMSE indicates better predictive accuracy
- Larger errors have a higher impact due to squaring
- Useful when large prediction errors are undesirable

Why RMSE was used:

- Penalizes large deviations
- Provides insight into error severity
- Commonly used in regression evaluation

2.3 R^2 Score (Coefficient of Determination)

Definition:
R^2 score measures the proportion of variance in the target variable that is explained by the model.

Formula:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

Where:

- $\bar{y}$= mean of actual values

Interpretation:

- $R^2 = 1$ → perfect prediction
- $R^2 = 0$ → model explains no variance
- $R^2 < 0$ → model performs worse than mean prediction

Why R^2 is the primary metric:

- Indicates overall explanatory power of the model
- Helps assess how well the model captures relationships between features and FWI
- Mentor specifically emphasized the importance of R^2 score

In this project, **test R^2 score** was considered the main evaluation metric to assess generalization.

3. Predicted vs Actual Values Plot

Purpose of the Visualization

The predicted vs actual plot visually compares model predictions against true values.

- X-axis: Actual FWI values
- Y-axis: Predicted FWI values

A diagonal reference line represents ideal predictions.

Interpretation:

- Points close to the diagonal indicate accurate predictions
- Wide scatter suggests poor model performance
- Patterns may reveal bias or underfitting

This plot provides an intuitive understanding of model behavior beyond numerical metrics.

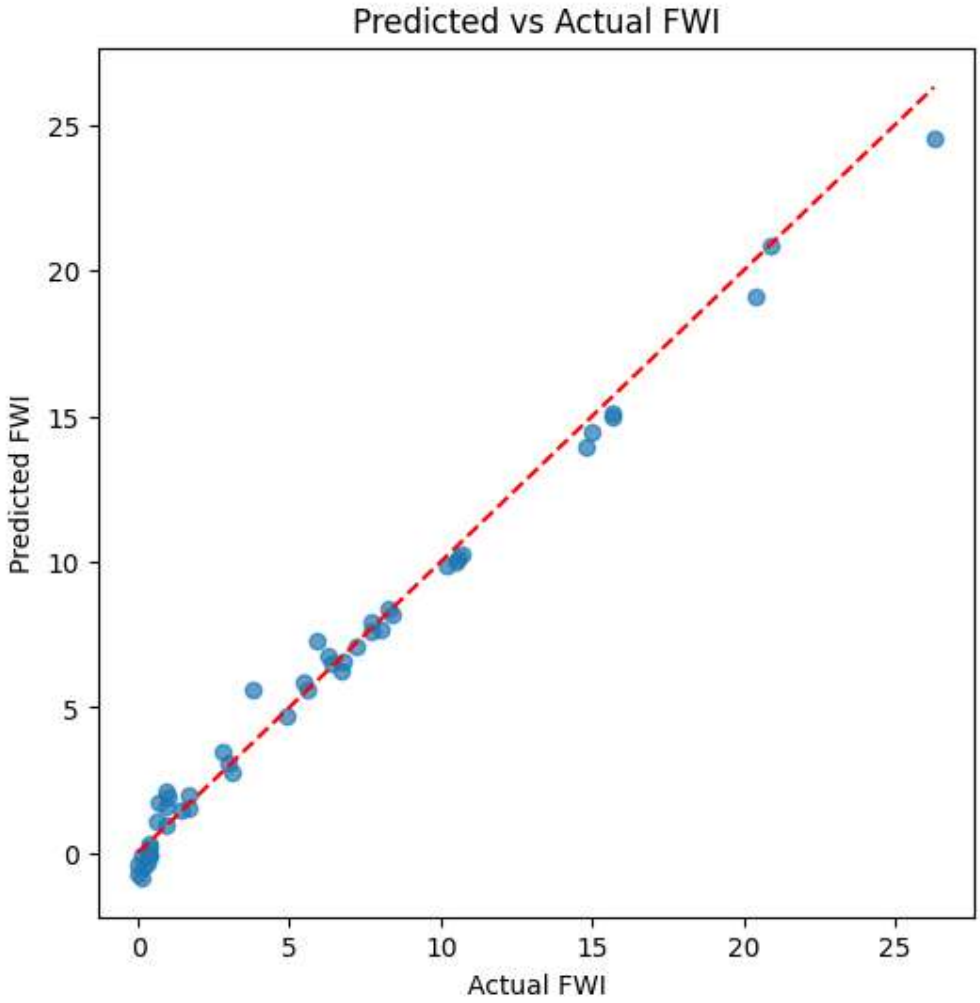

Fig. Predicted vs Actual FWI

4. Model Optimization (Alpha Tuning)

Ridge Regression includes a regularization parameter **alpha (α)**, which controls the strength of L2 regularization.

- Small alpha → model behaves like linear regression
- Large alpha → stronger regularization, reduced variance

To identify the optimal alpha value, **GridSearchCV** was used during model training.

Why GridSearchCV was applied:

- Performs cross-validation
- Evaluates multiple alpha values
- Selects the value that maximizes R^2 score
- Ensures balanced bias–variance tradeoff

The final model used in this milestone corresponds to the **tuned alpha value**, ensuring improved generalization.

5. Conclusion

In Milestone 3, the Ridge Regression model was thoroughly evaluated using standard regression metrics and visualization techniques. The use of MAE and RMSE quantified prediction errors, while the R^2 score measured the model's explanatory strength. The predicted vs actual plot visually confirmed the model's effectiveness, and prior alpha tuning ensured optimal regularization.

Overall, the evaluation results indicate that the model generalizes well and is suitable for further optimization and deployment in subsequent milestones.