

MILESTONE 2

This project uses the Bejaia Region Fire Weather dataset to predict the Fire Weather Index (FWI) using linear and ridge regression.

1. Data and Preprocessing

- Loaded `Bejaia-Region-Dataset_Cleaned.csv` using pandas.
- Selected numeric features as inputs: Temperature, RH, Ws, Rain, DMC, DC, ISI, BUI, FFMC.
- Set FWI as the target variable.
- Split data into train and test sets using `train_test_split` with 80% train and 20% test.

2. Linear Regression Model

- Trained a `LinearRegression` model on the training data.
- Evaluated on the test set using:
 - Mean Absolute Error (MAE): ~0.525
 - Mean Squared Error (MSE): ~0.702
 - R² Score: ~0.985
- These metrics show that the basic linear model already fits the FWI data very well.

3. Ridge Regression and Alpha Tuning

- Trained a Ridge model (`Ridge(alpha=1.0)`) and compared metrics with linear regression.
- Then tested multiple alpha values: 0.1, 0.2, 1.0, 2.0, 5.0, 10.0.
- For each alpha, computed train and test:
 - MSE, MAE, and R².

- Plotted:
 - Alpha vs Train/Test MSE
 - Alpha vs Train/Test MAE
 - Alpha vs Train/Test R^2
- Observations:
 - Train and test curves are close for all alphas → no strong overfitting or underfitting.
 - Test MSE and R^2 are best around $\alpha = 0.1$, so this is taken as the optimal regularization strength.

4. Model Saving with Pickle

- After choosing the best Ridge model, the trained model is saved using Python's `pickle` module:
- The saved model can be loaded later with `pickle.load` and used directly for FWI predictions without retraining.

This completes data loading, basic linear modeling, Ridge regularization with alpha tuning, metric/graph analysis, and model persistence up to the current task.