

ZERO SHOT

With reference to Atomic Data Services from the RAG system as Context, please suggest suitable services to address each of the steps below.

Given

Datasource 1: CSV file with 31 columns, including a row ID, timestamp, 28 temperature values, and X axis displacement values from a sensorized spindle on a working center machine.

I need services to perform the following steps.

STEP 1: A Collect service for reading the content of Datasource 1 by excluding the row ID column.

STEP 2: A Monitor service for building a multiple linear regression model on the dataset obtained from STEP 1. Consider the 28 temperature features as input for the model and the displacements feature as target. Produce as output only the two columns regarding the real and the predicted displacements.

Based on the context provided, here are the services to address each of the steps:

STEP 1: A Collect service for reading the content of Datasource 1 by excluding the row ID column.

- There is no direct service mentioned in the context that fits this step exactly. However, you can use any Collect or Retrieve data service that supports filtering and extracting specific columns from a CSV file. For example, you could create a custom Python script using libraries like Pandas or Scikit-Learn to read the CSV file and exclude the row ID column before returning the result.

STEP 2: A Monitor service for building a multiple linear regression model on the dataset obtained from STEP 1. Consider the 28 temperature features as input for the model and the displacements feature as target. Produce as output only the two columns regarding the real and the predicted displacements.

- The context includes a Multiple Linear Regression Service named `"/multiple_linear_regression"`. This service accepts a Pandas DataFrame, list of independent variables (the 28 temperature features), dependent variable (displacements feature), and produces as output only the real and predicted displacements in its response.