

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.

Based on the context provided, there isn't an exact match for a service that reads a CSV file and excludes a specific column (row ID in this case) out-of-the-box. However, you can achieve this functionality by using multiple services. Here is a suggested approach:

1. Use a Collect service to read the content of Datasource 1:
 - Service name: `srv_collect_datasource1`
 - From both JSON files (srv8 and srv3), there isn't an exact match for a Collect service that reads CSV files. You can use a generic Collect service to read the content of Datasource 1.
2. Use a Transform service to filter out the row ID column:
 - Service name: `srv_filter_datasource1`
 - The given JSON files do not have an exact match for a Transform service that excludes a specific column, but you can achieve this by writing a custom transformation script. You will need to use a programming language like Python or R to read the dataframe from the Collect service output and drop the row ID column before passing it on as input to the next service.

So, the suggested sequence of services would be:

1. `srv_collect_datasource1` (Collect)
2. `srv_filter_datasource1` (Transform)

Look better in the provided context, because the suggested services are not relevant for my request. I want a service specific for reading a CSV file.

I apologize for the misunderstanding.

By looking better into the context I found the following service which seems apt to read a CSV file: "srv2_read_csv". You can send a GET request to the route "read_csv" with a path to read the desired file.

STEP 2: A Dispatch service for selecting a subset of the 28 temperature features on the dataset obtained from STEP 1 by applying Factor

Analysis. Produce as output only the columns regarding the temperature features selected by the algorithm and displacements.

Based on the context provided, the "srv8_factor_analysis" service described in the JSON file ``/content/data/srv8.json`` can be used for STEP 2. This service accepts a Pandas DataFrame and allows for specifying the variables upon which the Factor Analysis has to be performed.

STEP 3: A Monitor service for building a multiple linear regression model on the dataset obtained from STEP 2. Consider the temperature features selected in STEP 2 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, there is a service called ``srv12`` that seems to fit the requirements of STEP 3. Here are the details:

Service name: ``srv12_multiple_linear_regression`` (Monitor)

In the first JSON file, the service's title is "srv12", and its task type is "PREDICTION". The path for this service is `"/multiple_linear_regression"` with a POST request. According to the provided OpenAPI documentation, this service accepts a dataframe as input along with the names of independent variables (temperature features) and the dependent variable (displacements feature). It returns the model summary and coefficients as output.

Since you only want the real and predicted displacements, you can extract these columns from the output provided by the ``srv12_multiple_linear_regression`` service. This will be the input for the next step.

STEP 4: A Monitor service for comparing the real and predicted displacements on the dataset obtained from STEP 2.

For STEP 4, you can use a service that compares the real and predicted displacements, calculates some statistical metrics, and provides a summary of the results. There isn't an exact match for such a service in the provided context, but you can create a custom Monitor service using a programming language like Python or R to implement the comparison logic and calculate the required statistics.

Service name: ``srv_compare_displacements`` (Monitor)

This custom service would take the real and predicted displacements as input from STEP 3, perform some statistical comparisons, and produce

output containing relevant metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Percentage Error (MPE), etc.

You will need to write custom code in a Transform or Monitor service to implement the comparison logic and calculate these statistics based on your requirements. You can use libraries like NumPy, pandas, or scipy for statistical analysis in Python or similar packages in R.

Look better in the provided context.

I apologize for the misunderstanding.

By looking better into the context I found the following service.

Service name: `srv9_compare_model_output_with_real_data`

Use this service to perform some statistical comparisons, and produce output containing relevant metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Percentage Error (MPE), etc.