

ChatRTDIP

Leveraging the power of large language models (LLM) and generative AI

I would like to use RTDIP components to read from SparkDeltaSource, transform using PandasToPySparkTransformer, then write to SparkEventhubDestination, return python code

Here is the python code to use RTDIP components to read from SparkDeltaSource, transform using PandasToPySparkTransformer, and then write to SparkEventhubDestination:

```
from rtdip_sdk.pipelines.sources import SparkDeltaSource
from rtdip_sdk.pipelines.destinations import SparkEventhubDestination
from rtdip_sdk.pipelines.transformers import PandasToPySparkTransformer
from rtdip_sdk.pipelines.utilities import SparkSessionUtility

# Not required if using Databricks
spark = SparkSessionUtility(config={}).execute()

# Read from SparkDeltaSource
delta_source = SparkDeltaSource(
    spark=spark,
    path="DELTA-TABLE-PATH",
    version=None
)

df = delta_source.read_batch()

# Transform using PandasToPySparkTransformer
transformer = PandasToPySparkTransformer(df)

df_transformed = transformer.transform()

# Write to SparkEventhubDestination
eventhub_destination = SparkEventhubDestination(
    spark=spark,
    data=df_transformed,
    options={
        "delta.enable.checkpoint.write": "true",
        "delta.enable.merge": "true",
        "delta.enable.write": "true"
    }
)
```

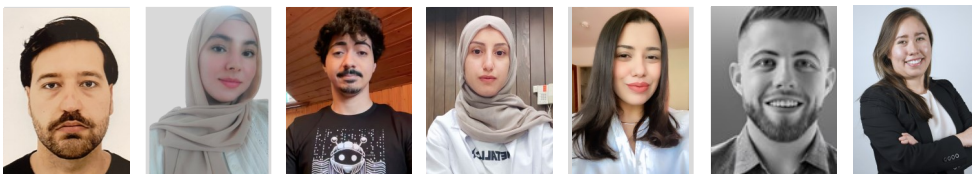
The project focuses on creating a platform where users can easily set up data pipelines using a chat-like interface.

The platform's user-friendly interface, enables users of diverse backgrounds, both technical and non-technical, to easily create complex RTDIP configurations.

Integration of the RTDIP and RTDIP SDK as a knowledge base, combined with a chat engine.



Demo Day on 07.02.2024
at 10:15 via ZOOM.



This project is a collaborative effort between FAU, TU Berlin and an Industry Partner.

AMOS Team 5—Zain Hazzouri, Sara El Brak, Obeidah Smadi, Nahrain Gtari, Linda Najar, Avi Katziuk, Cecilia Betancourt.