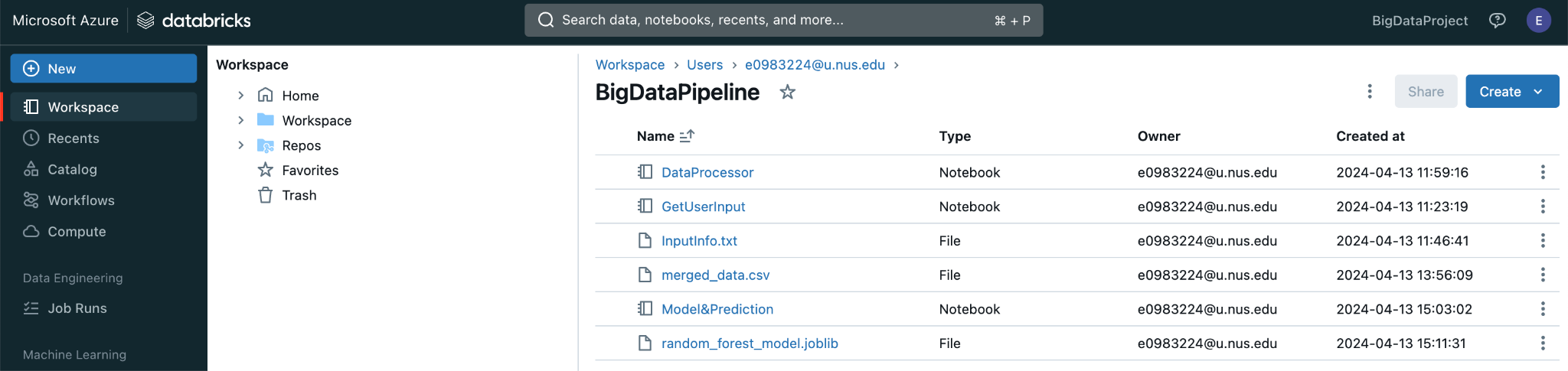
Our Workflow is deployed in an Azure Databricks environment. Please ensure that these files are placed in the same directory in the workspace, as shown in the figure below.

* **BigDataPipeline.dbc** contains three tasks in the workflow.
* **InputInfo.txt** used to get user input.If you want to change the start and end stations for the query, modify them in this file.
* **random\_forest\_model.joblib** is a locally trained prediction model. It uses "distance," "total\_volume," and "rainfall\_value" as independent variables, not including the additional "interval\_current" variable mentioned in the Additional Effort.
* **merged\_data.csv** is used to extract the predictors needed for the model prediction (see Report-Workflow Building-Data Processor).