

Yunting LU



| Version | Author | Date |
|---------|------------|---------------------|
| 1.0 | yunting.lu | 2023-04-27 09:05:06 |

Table of contents

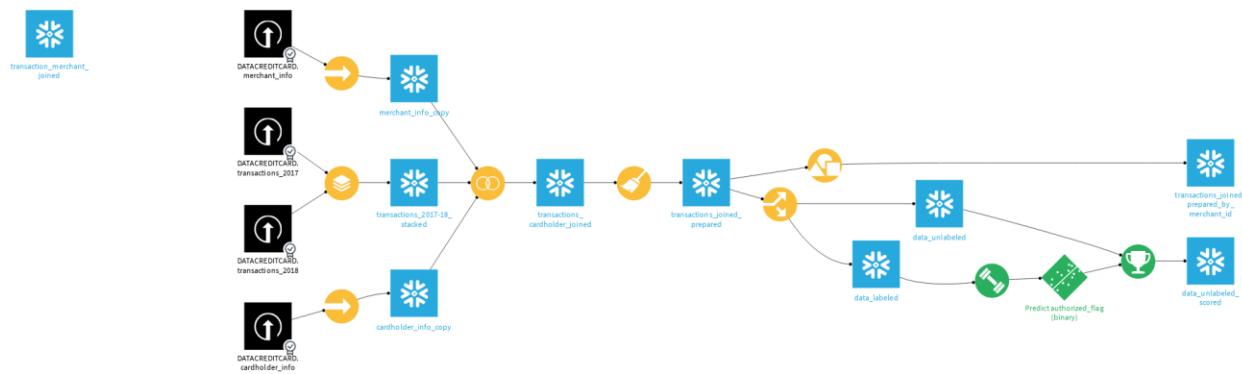
| | | |
|--------------|---|------------------------------|
| I. | Project general description | 4 |
| II. | Graphical Representation of the Flow | 5 |
| III. | Datasets | 6 |
| 1. | {{\\$dataset.name}} | Error! Bookmark not defined. |
| a) | Status of the dataset {{\\$dataset.name}} | Error! Bookmark not defined. |
| b) | Related Recipes | 24 |
| c) | Schema | 24 |
| IV. | Managed folders | Error! Bookmark not defined. |
| 1. | {{\\$folder.name}} | Error! Bookmark not defined. |
| a) | Status | Error! Bookmark not defined. |
| b) | Derived datasets | Error! Bookmark not defined. |
| c) | Related Recipes | Error! Bookmark not defined. |
| V. | Recipes | 25 |
| 1. | {{ \\$recipe.name }} | Error! Bookmark not defined. |
| a) | Status | 30 |
| b) | Inputs / Outputs | 30 |
| c) | Prepare Steps | 28 |
| d) | Join | 27 |
| e) | Code | Error! Bookmark not defined. |
| f) | Windows | Error! Bookmark not defined. |
| g) | Sorted Columns | Error! Bookmark not defined. |
| h) | Top rows and sorting direction | Error! Bookmark not defined. |
| i) | Parameters | Error! Bookmark not defined. |
| VI. | Saved Models | 31 |
| 1. | {{ \\$model.name }} | Error! Bookmark not defined. |
| a) | Status | 31 |
| b) | Related Recipes | 31 |
| VII. | Model Evaluation Stores | Error! Bookmark not defined. |
| 1. | {{ \\$mes.name }} | Error! Bookmark not defined. |
| a) | Related Recipes | Error! Bookmark not defined. |
| VIII. | Labeling Tasks | Error! Bookmark not defined. |
| 1. | {{ \\$labeling_task.name }} | Error! Bookmark not defined. |
| a) | Status | Error! Bookmark not defined. |
| b) | Inputs / Outputs | Error! Bookmark not defined. |
| IX. | Deployment and Monitoring | 32 |
| A. | Implementation Details | 32 |
| B. | Version Control | 32 |

I. Project general description

Project references for the created flow:

- Project name: **Yunting LU**
- Project key: **YUNTINGLU**
- Project short description:
*The project *Yunting LU* was created by yunting.lu on Apr 19th 2023*
- Status: **Sandbox**
- Current git branch: **master**
- Created by **yunting.lu** on the 2023-04-19 12:01:49
- Last Modification by **yunting.lu** on the 2023-04-19 15:00:05

II. Graphical Representation of the Flow



III. Datasets

List all the datasets in the flow.

1. transactions_2017

Summarized information about your dataset transactions_2017:

- This dataset has not been created in this project. Its origin project is **data credit card**

a) Status of the dataset transactions_2017

Characteristics of the dataset transactions_2017:

- Source project name: **data credit card**
- Type: **Uploaded Files**
- Connection: **dataiku-managed-storage**
- Total Size: **Not computed**
- Record Count: **Not computed**
- Created by **maxime.capron** on the 2023-04-19 06:42:26
- Last Modification by **maxime.capron** on the 2023-04-19 06:42:49

b) Related Recipes

The dataset transactions_2017 relates to the following recipes.

Successor Recipes

List of recipes with transactions_2017 as an input:

- compute_transactions_2017-18_stacked (*Stack*)

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|-----------------------|-------------|-------------|
| <u>transaction_id</u> | string | |

| | | |
|-----------------------------|--------|--|
| authorized_flag | string | |
| purchase_date | string | |
| card_id | string | |
| merchant_id | string | |
| merchant_category_id | string | |
| item_category | string | |
| purchase_amount | string | |
| signature_provided | string | |

2. transactions_2018

Summarized information about your dataset transactions_2018:

- This dataset has not been created in this project. Its origin project is **data credit card**

a) Status of the dataset transactions_2018

Characteristics of the dataset transactions_2018:

- Source project name: **data credit card**
- Type: **Uploaded Files**
- Connection: **dataiku-managed-storage**
- Total Size: **Not computed**
- Record Count: **Not computed**
- Created by **maxime.capron** on the 2023-04-19 06:41:55
- Last Modification by **maxime.capron** on the 2023-04-19 06:43:02

b) Related Recipes

The dataset transactions_2018 relates to the following recipes.

Successor Recipes

List of recipes with transactions_2018 as an input:

- compute_transactions_2017-18_stacked (*Stack*)

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|-----------------------------------|-------------|-------------|
| <code>transaction_id</code> | string | |
| <code>authorized_flag</code> | string | |
| <code>purchase_date</code> | string | |
| <code>card_id</code> | string | |
| <code>merchant_id</code> | string | |
| <code>merchant_category_id</code> | string | |
| <code>item_category</code> | string | |
| <code>purchase_amount</code> | string | |
| <code>signature_provided</code> | string | |

3. transactions_2017-18_stacked

Summarized information about your dataset transactions_2017-18_stacked:

a) Status of the dataset transactions_2017-18_stacked

Characteristics of the dataset transactions_2017-18_stacked:

- Type: **Snowflake**
- Connection: **managed-snowflake**
- Total Size: **Not computed**
- Record Count: **Not computed**
- Created by **yunting.lu** on the 2023-04-19 12:08:27
- Last Modification by **yunting.lu** on the 2023-04-19 12:08:27

b) Related Recipes

The dataset transactions_2017-18_stacked relates to the following recipes.

Parent Recipe

List of recipes with transactions_2017-18_stacked as an output:

- compute_transactions_2017-18_stacked ([Stack](#))

Successor Recipes

List of recipes with transactions_2017-18_stacked as an input:

- compute_transactions_cardholder_joined ([Join](#))

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|----------------------|-------------|-------------|
| transaction_id | string | |
| authorized_flag | string | |
| purchase_date | string | |
| card_id | string | |
| merchant_id | string | |
| merchant_category_id | string | |
| item_category | string | |
| purchase_amount | string | |
| signature_provided | string | |

4. cardholder_info

Summarized information about your dataset cardholder_info:

- This dataset has not been created in this project. Its origin project is **data credit card**

a) Status of the dataset cardholder_info

Characteristics of the dataset cardholder_info:

- Source project name: **data credit card**
- Type: **Uploaded Files**
- Connection: **dataiku-managed-storage**

- Total Size: **Not computed**
- Record Count: **Not computed**
- Created by **maxime.capron** on the 2023-04-19 06:42:42
- Last Modification by **maxime.capron** on the 2023-04-19 06:42:58

b) Related Recipes

The dataset cardholder_info relates to the following recipes.

Successor Recipes

List of recipes with cardholder_info as an input:

- compute_cardholder_info_copy (*Sync*)

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|------------------------------|-------------|-------------|
| internal_card_mapping | string | |
| first_active_month | string | |
| reward_program | string | |
| latitude | string | |
| longitude | string | |
| fico_score | string | |
| age | string | |

5. cardholder_info_copy

Summarized information about your dataset cardholder_info_copy:

a) Status of the dataset cardholder_info_copy

Characteristics of the dataset cardholder_info_copy:

- Type: **Snowflake**
- Connection: **managed-snowflake**

- Total Size: **Not computed**
- Record Count: **Not computed**
- Created by **yunting.lu** on the 2023-04-19 12:14:44
- Last Modification by **yunting.lu** on the 2023-04-19 12:14:44

b) Related Recipes

The dataset cardholder_info_copy relates to the following recipes.

Parent Recipe

List of recipes with cardholder_info_copy as an output:

- compute_cardholder_info_copy (*Sync*)

Successor Recipes

List of recipes with cardholder_info_copy as an input:

- compute_transactions_cardholder_joined (*Join*)

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|-----------------------|-------------|-------------|
| internal_card_mapping | string | |
| first_active_month | string | |
| reward_program | string | |
| latitude | string | |
| longitude | string | |
| fico_score | string | |
| age | string | |

6. merchant_info

Summarized information about your dataset merchant_info:

- This dataset has not been created in this project. Its origin project is **data credit card**

a) Status of the dataset merchant_info

Characteristics of the dataset merchant_info:

- Source project name: **data credit card**
- Type: **Uploaded Files**
- Connection: **dataiku-managed-storage**
- Total Size: **Not computed**
- Record Count: **Not computed**
- Created by **maxime.capron** on the 2023-04-19 06:42:37
- Last Modification by **maxime.capron** on the 2023-04-19 06:42:54

b) Related Recipes

The dataset merchant_info relates to the following recipes.

Successor Recipes

List of recipes with merchant_info as an input:

- compute_merchant_info_copy (*Sync*)

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|-----------------------|-------------|-------------|
| merchant_id | string | |
| merchant_category_id | string | |
| subsector_description | string | |
| latitude | string | |
| longitude | string | |

7. merchant_info_copy

Summarized information about your dataset merchant_info_copy:

a) Status of the dataset merchant_info_copy

Characteristics of the dataset merchant_info_copy:

- Type: **Snowflake**
- Connection: **managed-snowflake**
- Total Size: **Not computed**
- Record Count: **Not computed**
- Created by **yunting.lu** on the 2023-04-19 12:14:04
- Last Modification by **yunting.lu** on the 2023-04-19 12:14:04

b) Related Recipes

The dataset merchant_info_copy relates to the following recipes.

Parent Recipe

List of recipes with merchant_info_copy as an output:

- compute_merchant_info_copy (*Sync*)

Successor Recipes

List of recipes with merchant_info_copy as an input:

- compute_transactions_cardholder_joined (*Join*)

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|-----------------------|-------------|-------------|
| merchant_id | string | |
| merchant_category_id | string | |
| subsector_description | string | |
| latitude | string | |
| longitude | string | |

8. transactions_cardholder_joined

Summarized information about your dataset transactions_cardholder_joined:

a) Status of the dataset transactions_cardholder_joined

Characteristics of the dataset transactions_cardholder_joined:

- Type: **Snowflake**
- Connection: **managed-snowflake**
- Total Size: **Not computed**
- Record Count: **347,928**
- Created by **yunting.lu** on the 2023-04-19 12:17:24
- Last Modification by **yunting.lu** on the 2023-04-19 13:00:44

b) Related Recipes

The dataset transactions_cardholder_joined relates to the following recipes.

Parent Recipe

List of recipes with transactions_cardholder_joined as an output:

- compute_transactions_cardholder_joined (*Join*)

Successor Recipes

List of recipes with transactions_cardholder_joined as an input:

- compute_transactions_joined_prepared (*Prepare*)

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|-----------------|-------------|-------------|
| transaction_id | string | |
| authorized_flag | string | |
| purchase_date | string | |
| card_id | string | |

| | |
|--------------------------------|--------|
| merchant_id | string |
| merchant_category_id | string |
| item_category | string |
| purchase_amount | string |
| signature_provided | string |
| card_first_active_month | string |
| card_reward_program | string |
| card_latitude | string |
| card_longitude | string |
| card_fico_score | string |
| card_age | string |
| merchant_merchant_category_id | string |
| merchant_subsector_description | string |
| merchant_latitude | string |
| merchant_longitude | string |

9. transactions_joined_prepared

Summarized information about your dataset transactions_joined_prepared:

a) Status of the dataset transactions_joined_prepared

Characteristics of the dataset transactions_joined_prepared:

- Type: **Snowflake**
- Connection: **managed-snowflake**
- Total Size: **Not computed**
- Record Count: **347,928**
- Created by **yunting.lu** on the 2023-04-19 12:33:07
- Last Modification by **yunting.lu** on the 2023-04-19 13:06:15

b) Related Recipes

The dataset transactions_joined_prepared relates to the following recipes.

Parent Recipe

List of recipes with transactions_joined_prepared as an output:

- compute_transactions_joined_prepared (*Prepare*)

Successor Recipes

List of recipes with transactions_joined_prepared as an input:

- compute_transactions_joined_prepared_by_merchant_id (*Group*)
- split_transactions_joined_prepared (*Split*)

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|---------------------------|-------------|-------------|
| transaction_id | bigint | |
| authorized_flag | bigint | |
| purchase_date | string | |
| purchase_DayOfWeek | bigint | |
| weekend | boolean | |
| card_id | string | |
| merchant_id | string | |
| merchant_category_id | bigint | |
| item_category | string | |
| purchase_amount | double | |
| signature_provided | bigint | |
| card_first_active_month | string | |
| since_purchase_date_years | bigint | |
| card_reward_program | string | |

| | | |
|--------------------------------|--------|--|
| card_latitude | double | |
| card_longitude | double | |
| card_fico_score | bigint | |
| card_age | bigint | |
| merchant_merchant_category_id | bigint | |
| merchant_subsector_description | string | |
| merchant_latitude | double | |
| merchant_longitude | double | |
| card_geopoint | string | |
| card_state | string | |
| merchant_geopoint | string | |
| distance_betw_geopoint | double | |
| merchant_state | string | |

10. transactions_joined_prepared_by_merchant_id

Summarized information about your dataset transactions_joined_prepared_by_merchant_id:

a) Status of the dataset transactions_joined_prepared_by_merchant_id

Characteristics of the dataset transactions_joined_prepared_by_merchant_id:

- Type: **Snowflake**
- Connection: **managed-snowflake**
- Total Size: **Not computed**
- Record Count: **50,339**
- Created by **yunting.lu** on the 2023-04-19 13:36:00
- Last Modification by **yunting.lu** on the 2023-04-19 14:49:02

b) Related Recipes

The dataset transactions_joined_prepared_by_merchant_id relates to the following recipes.

Parent Recipe

List of recipes with transactions_joined_prepared_by_merchant_id as an output:

- compute_transactions_joined_prepared_by_merchant_id (*Group*)

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|----------------------------|-------------|-------------|
| merchant_id | string | |
| purchase_amount_avg | double | |
| merchant_geopoint_distinct | bigint | |
| fraud_sum | double | |
| count | bigint | |

11. data_labeled

Summarized information about your dataset data_labeled:

a) Status of the dataset data_labeled

Characteristics of the dataset data_labeled:

- Type: **Snowflake**
- Connection: **managed-snowflake**
- Total Size: **Not computed**
- Record Count: **201,941**
- Created by **yunting.lu** on the 2023-04-19 13:49:18
- Last Modification by **yunting.lu** on the 2023-04-19 13:49:43

b) Related Recipes

The dataset data_labeled relates to the following recipes.

Parent Recipe

List of recipes with data_labeled as an output:

- split_transactions_joined_prepared (*Split*)

Successor Recipes

List of recipes with data_labeled as an input:

- train_PredictAuthorizedFlagBinary_ (*Prediction Training*)

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|---------------------------|-------------|-------------|
| transaction_id | bigint | |
| authorized_flag | bigint | |
| purchase_date | string | |
| purchase_DayOfWeek | bigint | |
| weekend | boolean | |
| card_id | string | |
| merchant_id | string | |
| merchant_category_id | bigint | |
| item_category | string | |
| purchase_amount | double | |
| signature_provided | bigint | |
| card_first_active_month | string | |
| since_purchase_date_years | bigint | |
| card_reward_program | string | |
| card_latitude | double | |
| card_longitude | double | |
| card_fico_score | bigint | |

| | | |
|---------------------------------------|--------|--|
| card_age | bigint | |
| merchant_merchant_category_id | bigint | |
| merchant_subsector_description | string | |
| merchant_latitude | double | |
| merchant_longitude | double | |
| card_geopoint | string | |
| card_state | string | |
| merchant_geopoint | string | |
| distance_betw_geopoint | double | |
| merchant_state | string | |

12. data_unlabeled

Summarized information about your dataset data_unlabeled:

a) Status of the dataset data_unlabeled

Characteristics of the dataset data_unlabeled:

- Type: **Snowflake**
- Connection: **managed-snowflake**
- Total Size: **Not computed**
- Record Count: **Not computed**
- Created by **yunting.lu** on the 2023-04-19 13:49:27
- Last Modification by **yunting.lu** on the 2023-04-19 13:49:43

b) Related Recipes

The dataset data_unlabeled relates to the following recipes.

Parent Recipe

List of recipes with data_unlabeled as an output:

- **split_transactions_joined_prepared** (*Split*)

Successor Recipes

List of recipes with data_unlabeled as an input:

- score_data_unlabeled (*Prediction Scoring*)

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|--------------------------------|-------------|-------------|
| transaction_id | bigint | |
| authorized_flag | bigint | |
| purchase_date | string | |
| purchase_DayOfWeek | bigint | |
| weekend | boolean | |
| card_id | string | |
| merchant_id | string | |
| merchant_category_id | bigint | |
| item_category | string | |
| purchase_amount | double | |
| signature_provided | bigint | |
| card_first_active_month | string | |
| since_purchase_date_years | bigint | |
| card_reward_program | string | |
| card_latitude | double | |
| card_longitude | double | |
| card_fico_score | bigint | |
| card_age | bigint | |
| merchant_merchant_category_id | bigint | |
| merchant_subsector_description | string | |
| merchant_latitude | double | |

| | | |
|-------------------------------|--------|--|
| merchant_longitude | double | |
| card_geopoint | string | |
| card_state | string | |
| merchant_geopoint | string | |
| distance_betw_geopoint | double | |
| merchant_state | string | |

13. data_unlabeled_scored

Summarized information about your dataset `data_unlabeled_scored`:

a) Status of the dataset `data_unlabeled_scored`

Characteristics of the dataset `data_unlabeled_scored`:

- Type: **Snowflake**
- Connection: **managed-snowflake**
- Total Size: **Not computed**
- Record Count: **145,987**
- Created by **yunting.lu** on the 2023-04-19 14:39:18
- Last Modification by **yunting.lu** on the 2023-04-19 14:39:18

b) Related Recipes

The dataset `data_unlabeled_scored` relates to the following recipes.

Parent Recipe

List of recipes with `data_unlabeled_scored` as an output:

- `score_data_unlabeled` (*Prediction Scoring*)

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|-----------------------|-------------|-------------|
| transaction_id | bigint | |

| | |
|--------------------------------|---------|
| authorized_flag | bigint |
| purchase_date | date |
| purchase_DayOfWeek | bigint |
| weekend | boolean |
| card_id | string |
| merchant_id | string |
| merchant_category_id | bigint |
| item_category | string |
| purchase_amount | double |
| signature_provided | bigint |
| card_first_active_month | string |
| since_purchase_date_years | bigint |
| card_reward_program | string |
| card_latitude | double |
| card_longitude | double |
| card_fico_score | bigint |
| card_age | bigint |
| merchant_merchant_category_id | bigint |
| merchant_subsector_description | string |
| merchant_latitude | double |
| merchant_longitude | double |
| card_geopoint | string |
| card_state | string |
| merchant_geopoint | string |
| distance_betw_geopoint | double |
| merchant_state | string |
| proba_0 | double |

| | | |
|-------------------|--------|--|
| proba_1 | double | |
| prediction | string | |

14. transaction_merchant_joined

Summarized information about your dataset transaction_merchant_joined:

a) Status of the dataset transaction_merchant_joined

Characteristics of the dataset transaction_merchant_joined:

- Type: **Snowflake**
- Connection: **managed-snowflake**
- Total Size: **Not computed**
- Record Count: **Not computed**
- Created by **yunting.lu** on the 2023-04-19 12:21:24
- Last Modification by **yunting.lu** on the 2023-04-19 12:26:04

b) Related Recipes

The dataset transaction_merchant_joined relates to the following recipes.

c) Schema

Schema details of the dataset.

| Column Name | Column Type | Description |
|------------------------------|-------------|-------------|
| merchant_id | string | |
| merchant_category_id | string | |
| subsector_description | string | |
| latitude | string | |
| longitude | string | |

IV. Recipes

List all the recipes in the flow.

1. compute_transactions_2017-18_stacked

Summarized information about the recipe compute_transactions_2017-18_stacked.

a) Status

Characteristics of the recipe compute_transactions_2017-18_stacked:

- Type: **Stack**
- Created by **yunting.lu** on the 2023-04-19 12:08:27
- Last Modification by **yunting.lu** on the 2023-04-19 12:12:44

b) Inputs / Outputs

List of the inputs of the recipe:

- Dataset transactions_2017 (*Uploaded Files*)
- Dataset transactions_2018 (*Uploaded Files*)

List of the outputs of the recipe:

- Dataset transactions_2017-18_stacked (*Snowflake*)

Stacking mode: Union of input schemas

2. compute_cardholder_info_copy

Summarized information about the recipe compute_cardholder_info_copy.

a) Status

Characteristics of the recipe compute_cardholder_info_copy:

- Type: **Sync**

- Created by **yunting.lu** on the 2023-04-19 12:14:44
- Last Modification by **yunting.lu** on the 2023-04-19 12:14:46

b) Inputs / Outputs

List of the inputs of the recipe:

- Dataset cardholder_info (*Uploaded Files*)

List of the outputs of the recipe:

- Dataset cardholder_info_copy (*Snowflake*)

3. compute_merchant_info_copy

Summarized information about the recipe compute_merchant_info_copy.

a) Status

Characteristics of the recipe compute_merchant_info_copy:

- Type: **Sync**
- Created by **yunting.lu** on the 2023-04-19 12:14:04
- Last Modification by **yunting.lu** on the 2023-04-19 12:14:18

b) Inputs / Outputs

List of the inputs of the recipe:

- Dataset merchant_info (*Uploaded Files*)

List of the outputs of the recipe:

- Dataset merchant_info_copy (*Snowflake*)

4. compute_transactions_cardholder_joined

Summarized information about the recipe compute_transactions_cardholder_joined.

a) Status

Characteristics of the recipe compute_transactions_cardholder_joined:

- Type: **Join**
- Created by **yunting.lu** on the 2023-04-19 12:17:24
- Last Modification by **yunting.lu** on the 2023-04-19 13:00:51

b) Inputs / Outputs

List of the inputs of the recipe:

- Dataset transactions_2017-18_stacked ([Snowflake](#))
- Dataset cardholder_info_copy ([Snowflake](#))
- Dataset merchant_info_copy ([Snowflake](#))

List of the outputs of the recipe:

- Dataset transactions_cardholder_joined ([Snowflake](#))

c) Join

Left transactions_2017-18_stacked - cardholder_info_copy And

| Left input | Condition | Right input |
|------------|-----------|-----------------------|
| card_id | = | internal_card_mapping |

Left cardholder_info_copy - merchant_info_copy And

| Left input | Condition | Right input |
|------------|-----------|-------------|
| latitude | = | latitude |
| longitude | = | longitude |

5. compute_transactions_joined_prepared

Summarized information about the recipe compute_transactions_joined_prepared.

a) Status

Characteristics of the recipe compute_transactions_joined_prepared:

- Type: **Prepare**
- Created by **yunting.lu** on the 2023-04-19 12:33:07
- Last Modification by **yunting.lu** on the 2023-04-19 13:06:17

b) Inputs / Outputs

List of the inputs of the recipe:

- Dataset transactions_cardholder_joined (*Snowflake*)

List of the outputs of the recipe:

- Dataset transactions_joined_prepared (*Snowflake*)

c) Prepare Steps

| Type | Comment |
|--------------------------|---------|
| DateParser | |
| DateComponentsExtractor | |
| VisualIfRule | |
| DateDifference | |
| GeoPointCreator | |
| GeoPointCreator | |
| CityLevelReverseGeocoder | |
| CityLevelReverseGeocoder | |
| ColumnsSelector | |
| GeoDistanceProcessor | |

6. compute_transactions_joined_prepared_by_merchant_id

Summarized information about the recipe
compute_transactions_joined_prepared_by_merchant_id.

a) Status

Characteristics of the recipe **compute_transactions_joined_prepared_by_merchant_id**:

- Type: **Group**
- Created by **yunting.lu** on the 2023-04-19 13:36:00
- Last Modification by **yunting.lu** on the 2023-04-19 14:49:04

b) Inputs / Outputs

List of the inputs of the recipe:

- Dataset transactions_joined_prepared (*Snowflake*)

List of the outputs of the recipe:

- Dataset transactions_joined_prepared_by_merchant_id (*Snowflake*)

Group keys: merchant_id

7. split_transactions_joined_prepared

Summarized information about the recipe split_transactions_joined_prepared.

a) Status

Characteristics of the recipe split_transactions_joined_prepared:

- Type: **Split**
- Created by **yunting.lu** on the 2023-04-19 13:49:30
- Last Modification by **yunting.lu** on the 2023-04-19 13:50:58

b) Inputs / Outputs

List of the inputs of the recipe:

- Dataset transactions_joined_prepared (*Snowflake*)

List of the outputs of the recipe:

- Dataset data_labeled (*Snowflake*)
- Dataset data_unlabeled (*Snowflake*)

Split mode: Specific filter for each output dataset

8. train_Predict_authorized_flag_binary_

Summarized information about the recipe train_Predict_authorized_flag_binary_.

a) Status

Characteristics of the recipe train_PredictAuthorizedFlagBinary:

- Type: **Prediction Training**
- Created by **yunting.lu** on the 2023-04-19 14:38:48
- Last Modification by **yunting.lu** on the 2023-04-19 14:38:48

b) Inputs / Outputs

List of the inputs of the recipe:

- Dataset data_labeled ([Snowflake](#))

List of the outputs of the recipe:

- Saved Model Predict authorized_flag (binary) ([Random forest \(Random Forest\) - v1](#))

9. score_data_unlabeled

Summarized information about the recipe score_data_unlabeled.

a) Status

Characteristics of the recipe score_data_unlabeled:

- Type: **Prediction Scoring**
- Created by **yunting.lu** on the 2023-04-19 14:39:18
- Last Modification by **yunting.lu** on the 2023-04-19 14:39:30

b) Inputs / Outputs

List of the inputs of the recipe:

- Dataset data_unlabeled ([Snowflake](#))
- Saved Model Predict authorized_flag (binary) ([Random forest \(Random Forest\) - v1](#))

List of the outputs of the recipe:

- Dataset data_unlabeled_scored ([Snowflake](#))

V. Saved Models

List all the saved models in the flow.

1. Predict authorized_flag (binary)

Summarized information about the saved model Predict authorized_flag (binary):.

a) Status

Model id: **kLxujjCN**

Active version: **Random forest (Random Forest) - v1**

b) Related Recipes

List of the recipes related to the saved model.

Parent Recipe

Recipe with Predict authorized_flag (binary) as an output:

- train_Predict_authorized_flag__binary_ (*Prediction Training*)

Successor Recipes

List of recipes with Predict authorized_flag (binary) as an input:

- score_data_unlabeled (*Prediction Scoring*)

VI. Deployment and Monitoring

A. Implementation Details

- The flow can be found here: <https://dss-44270426-726225ce-dku.eu-west-3.app.dataiku.io/projects/YUNTINGLU/>
- The name of the generated file is: Dataiku Flow Documentation - YUNTINGLU.docx

B. Version Control

- The flow was executed with the following version of DSS: 11.4.1