# Hybrid Forecast Generation: Step Description

Forecast Consolidation step is aimed to merge several forecast values for a triple product/location/day and provide one (hybrid forecast) value for it. Input forecasts are to be in one granularity level due to Reconciliation step.

## Code Realization Requirements

The code should be created in a form TODO

# Input Data

The initial data for the step is listed below. These tables should be present in the system before this step for example in STG or in DDS area.

## RECONCILED\_FORECAST

Forecast values from ts\_forecast step that is reconciled to sku/location/day level of granularity.

|  |  |
| --- | --- |
| **RECONCILED\_FORECAST** | |
| Column Name | Description |
| **product\_lvl\_id<M>** | ID of product group for which specific ML model will be trained, M = max(**ml\_product\_lvl, ts\_product\_lvl)** |
| **location\_lvl\_id<N>** | ID of location group for which specific ML model will be trained, N = **ml\_location\_lvl** |
| **customer\_lvl\_id<L>** | ID of customer group for which specific ML model will be trained, L = **ml\_customer\_lvl** |
| **distr\_channel\_lvl\_id<K>** | ID of distr\_channel group for which specific ML model will be trained, K = **ml\_distr\_channel\_lvl** |
| **PERIOD\_DT** | Date of sales (calendar day) |
| **PERIOD\_END\_DT** | End date of the period |
| **TS\_FORECAST\_VALUE\_REC** | TS forecast value |
| **SEGMENT\_NAME** | Name of segment that was linked to a pair product/location within TS Project (can be missing) |
| **DEMAND\_TYPE** | ‘promo’ or ‘regular’ – type ML model is used to forecast |
| **ASSORTMENT\_TYPE** | new, or old |
| **ML\_FORECAST\_VALUE** | ML forecasted value |

## CONFIG\_PARAMETERS

The following config parameters are used within the step.

|  |  |
| --- | --- |
| **CONFIG.CONFIG\_PARAMETERS** | |
| Column Name | Description |
| **IB\_ZERO\_DEMAND\_THRESHOLD** | If TS Forecast value is less than that value it means that future demand will be zero. Default value = 0.01 |

## INITIAL\_GLOBAL parameters

All parameters are listed in initial\_global file.

|  |  |
| --- | --- |
| **INITIAL\_GLOBAL parameters init** | |
| Column Name | Description |

## Other Dependencies

All string comparisons are case-insensitive.

# Algorithm Definition

**4.0 Create mid-term hybrid forecast**

**Inputs:** MID\_RECONCILED\_FORECAST

SELECT PRODUCT\_LVL\_ID,

LOCATION\_LVL\_ID,

CUSTOMER\_LVL\_ID

DISTR\_CHANNEL\_LVL\_ID,

PERIOD\_DT,

PERIOD\_END\_DT,

TS\_FORECAST\_VALUE\_REC as TS\_FORECAST\_VALUE,

SEGMENT\_NAME,

DEMAND\_TYPE,

ASSORTMENT\_TYPE,

ML\_FORECAST\_VALUE,

TS\_FORECAST\_VALUE\_REC as HYBRID\_FORECAST\_VALUE,

ENSEMBLE\_FORECAST\_VALUE

'ts’ as FORECAST\_SOURCE

from MID\_RECONSILED\_FORECAST

## Join TS and ML forecasts

**Inputs:** RECONCILED\_FORECAST

**Transformation algorithm:**

1. Filling missing:
   1. TS\_FORECAST\_VALUE\_F = COALESCE(TS\_FORECAST\_VALUE\_REC, ML\_FORECAST\_VALUE).
   2. ML\_FORECAST\_VALUE\_F = COALESCE(ML\_FORECAST\_VALUE, TS\_FORECAST\_VALUE\_REC).
2. Add new column HYBRID\_FORECAST\_VALUE based on rules:

=

1. CASE WHEN

(**DEMAND\_TYPE** = ‘promo’ and **SEGMENT\_NAME** <> ‘Retired’**)** OR

**SEGMENT\_NAME** = ‘Short’ OR

**ASSORTMENT\_TYPE** = ‘new’OR

THEN **ML\_FORECAST\_VALUE\_F**

1. ELSE CASE WHEN

**SEGMENT\_NAME** = ‘Retired’ OR **SEGMENT\_NAME** = ‘Low Volume’

**TS\_FORECAST\_VALUE\_F <= IB\_ZERO\_DEMAND\_THRESHOLD**

THEN **TS\_FORECAST\_VALUE\_F**

1. ELSE

THEN AVERAGE[[1]](#footnote-2)(**TS\_FORECAST\_VALUE\_F, ML\_FORECAST\_VALUE\_F**)

1. Add new column FORECAST\_SOURCE based on rules:

FORECAST\_SOURCE=

1. CASE WHEN

(**DEMAND\_TYPE** = ‘promo’ and **SEGMENT\_NAME** <> ‘Retired’**)** OR

**SEGMENT\_NAME** = ‘Short’ OR

**ASSORTMENT\_TYPE** = ‘new’OR

THEN **‘ml’**

1. ELSE CASE WHEN

**SEGMENT\_NAME** = ‘Retired’ OR **SEGMENT\_NAME** = ‘Low Volume’

**TS\_FORECAST\_VALUE\_F <= IB\_ZERO\_DEMAND\_THRESHOLD**

THEN **‘ts’**

1. ELSE **‘ensemble’**
2. Add new column ENSEMBLE\_FORECAST\_VALUE based on rules:

ENSEMBLE\_FORECAST\_VALUE=

1. CASE WHEN

(**DEMAND\_TYPE** = ‘promo’ and **SEGMENT\_NAME** <> ‘Retired’**)** OR

**SEGMENT\_NAME** = ‘Short’ OR

**ASSORTMENT\_TYPE** = ‘new’OR

THEN **missing**

1. ELSE CASE WHEN

**SEGMENT\_NAME** = ‘Retired’ OR **SEGMENT\_NAME** = ‘Low Volume’

**TS\_FORECAST\_VALUE\_F <= IB\_ZERO\_DEMAND\_THRESHOLD**

THEN **missing**

1. ELSE

THEN AVERAGE[[2]](#footnote-3)(**TS\_FORECAST\_VALUE\_F, ML\_FORECAST\_VALUE\_F**)

1. Rename TS\_FORECAST\_VALUE\_REC to TS\_FORECAST\_VALUE

Output: As a result of this step, HYBRID\_FORECAST table is constructed, (detailed description see in section 4.2)*.*

## Output from the Algorithm

1. Table with all types of forecast

|  |  |
| --- | --- |
| **HYBRID\_FORECAST** | |
| Column Name | Description |
| **product\_lvl\_id<M>** | ID of product group for which specific ML model will be trained, M = **ml\_product\_lvl** |
| **location\_lvl\_id<N>** | ID of location group for which specific ML model will be trained, N = **ml\_location\_lvl** |
| **customer\_lvl\_id<L>** | ID of customer group for which specific ML model will be trained, L = **ml\_customer\_lvl** |
| **distr\_channel\_lvl\_id<K>** | ID of distr\_channel group for which specific ML model will be trained, K = **ml\_distr\_channel\_lvl** |
| **PERIOD\_DT** | Date of sales (calendar day) |
| **PERIOD\_END\_DT** | End date of the period |
| **TS\_FORECAST\_VALUE** | TS forecast value |
| **SEGMENT\_NAME** | Name of segment that was linked to a pair product/location within TS Project (can be missing) |
| **DEMAND\_TYPE** | ‘promo’ or ‘regular’ – type ML model is used to forecast |
| **ASSORTMENT\_TYPE** | new, or old |
| **ML\_FORECAST\_VALUE** | ML forecasted value |
| **HYBRID\_FORECAST\_VALUE** | Hybrid Forecast Value |
| **ENSEMBLE\_FORECAST\_VALUE** | Ensemble Forecast Value |
| **FORECAST\_SOURCE** | Forecast Source |

# Open questions and next steps

|  |  |
| --- | --- |
| № | OQ or NS |
| 1 | Evolve ensemble method (e.g. Adaptive Selection instead simple average of ML and TS forecast) |
| 2 | Provide a config file to handle consolidation rules (for example to specify consolidation rules for particular category). |
| 3 | ML Forecast can be reconciled with TS forecast on product/location/customer/dist\_channel level |

1. Taking into account missing values: AVERAGE(missing, 1) = AVERAGE(1, missing) = 1; AVERAGE(missing, missing) = missing [↑](#footnote-ref-2)
2. Taking into account missing values: AVERAGE(missing, 1) = AVERAGE(1, missing) = 1; AVERAGE(missing, missing) = missing [↑](#footnote-ref-3)