DECLARE date\_end, date\_latest\_follows DATE;

SET date\_end = (

SELECT PARSE\_DATE('%Y%m%d', MAX(partition\_id)) AS partition\_id

FROM `twttr-bq-cassowary-prod.user.INFORMATION\_SCHEMA.PARTITIONS`

WHERE partition\_id IS NOT NULL AND partition\_id != '\_\_NULL\_\_' AND table\_name="interaction\_graph\_labels\_daily"

);

SET date\_latest\_follows = (

SELECT PARSE\_DATE('%Y%m%d', MAX(partition\_id)) AS partition\_id

FROM `twttr-recos-ml-prod.user\_events.INFORMATION\_SCHEMA.PARTITIONS`

WHERE partition\_id IS NOT NULL AND partition\_id != '\_\_NULL\_\_' AND table\_name="valid\_user\_follows");

DELETE

FROM `twttr-recos-ml-prod.realgraph.scores`

WHERE ds = date\_end;

-- score candidates (59m)

INSERT INTO `twttr-recos-ml-prod.realgraph.scores`

WITH predicted\_scores AS (

SELECT

source\_id,

destination\_id,

p1.prob AS prob,

p2.prob AS prob\_explicit

FROM ML.PREDICT(MODEL `twttr-recos-ml-prod.realgraph.prod`,

(

SELECT

\*

FROM

`twttr-recos-ml-prod.realgraph.candidates` ) ) S1

CROSS JOIN UNNEST(S1.predicted\_label\_probs) AS p1

JOIN ML.PREDICT(MODEL `twttr-recos-ml-prod.realgraph.prod\_explicit`,

(

SELECT

\*

FROM

`twttr-recos-ml-prod.realgraph.candidates` ) ) S2

USING (source\_id, destination\_id)

CROSS JOIN UNNEST(S2.predicted\_label\_probs) AS p2

WHERE p1.label=1 AND p2.label=1

)

SELECT

COALESCE(predicted\_scores.source\_id, tweeting\_follows.source\_id) AS source\_id,

COALESCE(predicted\_scores.destination\_id, tweeting\_follows.destination\_id) AS destination\_id,

COALESCE(prob, 0.0) AS prob,

COALESCE(prob\_explicit, 0.0) AS prob\_explicit,

(tweeting\_follows.source\_id IS NOT NULL) AND (tweeting\_follows.destination\_id IS NOT NULL) AS followed,

date\_end AS ds

FROM

predicted\_scores

FULL JOIN

`twttr-recos-ml-prod.realgraph.tweeting\_follows` tweeting\_follows

USING (source\_id, destination\_id)