# FTR Tweet embeddings

export GCP\_PROJECT\_NAME='twttr-recos-ml-prod'

## Running Adhoc jobs

### Base ftrat5

```

rm dist/ftr-tweet-adhoc-job-bundle/ftr-tweet-adhoc-job.jar

./bazel bundle src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet:ftr-tweet-adhoc-job && \

bin/d6w create \

${GCP\_PROJECT\_NAME}/us-central1/ftr-tweets-ann-adhoc-job \

src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet/ftr-tweets-ann-adhoc-job.d6w \

--jar dist/ftr-tweet-adhoc-job-bundle/ftr-tweet-adhoc-job.jar \

--bind=profile.project=

${GCP\_PROJECT\_NAME} \

--bind=profile.user\_name=your\_ldap \

--bind=profile.build\_target="src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet:ftr-tweet-index-generation-adhoc-job" \

--bind=profile.date="2022-08-26T12" \

--bind=profile.machine="n2-standard-2" \

--bind=profile.job\_name="ftr-tweets-ann-adhoc-job" --ignore-existing

```

### ClusterToTweet Index with base ftrat5

```

export GCP\_PROJECT\_NAME='twttr-recos-ml-prod'

rm dist/ftr-tweet-index-generation-adhoc-job-bundle/ftr-tweet-index-generation-adhoc-job.jar

./bazel bundle src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet:ftr-tweet-index-generation-adhoc-job && \

bin/d6w create \

${GCP\_PROJECT\_NAME}/us-central1/ftr-tweet-index-generation-adhoc-job \

src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet/ftr-based-simclusters-index-generation-job.d6w \

--jar dist/ftr-tweet-index-generation-adhoc-job-bundle/ftr-tweet-index-generation-adhoc-job.jar \

--bind=profile.project=${GCP\_PROJECT\_NAME} \

--bind=profile.user\_name=your\_ldap \

--bind=profile.date="2022-08-27T12" \

--bind=profile.machine="n2-standard-2" \

--bind=profile.build\_target="src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet:ftr-tweet-index-generation-adhoc-job" \

--bind=profile.job\_name="ftr-tweet-index-generation-adhoc-job" --ignore-existing

```

### OON ftrat5

```

rm dist/oon-ftr-tweet-index-generation-adhoc-job-bundle/oon-ftr-tweet-index-generation-adhoc-job.jar

./bazel bundle src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet:oon-ftr-tweet-index-generation-adhoc-job && \

bin/d6w create \

${GCP\_PROJECT\_NAME}/us-central1/oon-ftr-ann-adhoc-job \

src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet/ftr-based-simclusters-index-generation-job.d6w \

--jar dist/oon-ftr-tweet-index-generation-adhoc-job-bundle/oon-ftr-tweet-index-generation-adhoc-job.jar \

--bind=profile.project=${GCP\_PROJECT\_NAME} \

--bind=profile.user\_name=${USER} \

--bind=profile.build\_target="src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet:oon-ftr-tweet-index-generation-adhoc-job" \

--bind=profile.date="2022-09-21T12" \

--bind=profile.machine="n2-standard-2" \

--bind=profile.job\_name="oon-ftr-ann-adhoc-job" --ignore-existing

```

## Scheduling jobs

### decayed\_sum\_job

```

export SERVICE\_ACCOUNT='cassowary'

export GCP\_PROJECT\_NAME='twttr-recos-ml-prod'

export PROJECT\_DATE='2022-07-24T16'

bin/d6w schedule \

${GCP\_PROJECT\_NAME}/us-central1/iikf2020-decayed-sum-ann-batch-job \

src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet/iikf2020-decayed-sum-ann-batch-job.d6w \

--bind=profile.project=${GCP\_PROJECT\_NAME} \

--bind=profile.user\_name=${SERVICE\_ACCOUNT} \

--bind=profile.machine="n2-highmem-4" \

--bind=profile.job\_name="iikf2020-decayed-sum-ann-batch-job" \

--bind=profile.date=${PROJECT\_DATE} \

--bind=profile.environment=prod

```

### ftrat5 pop1000

```

export SERVICE\_ACCOUNT='cassowary'

export GCP\_PROJECT\_NAME='twttr-recos-ml-prod'

export PROJECT\_DATE='2022-07-24T17'

bin/d6w schedule \

${GCP\_PROJECT\_NAME}/us-central1/iikf2020-ftrat5-pop1000-ann-batch-job \

src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet/iikf2020-ftrat5-pop1000-ann-batch-job.d6w \

--bind=profile.project=${GCP\_PROJECT\_NAME} \

--bind=profile.user\_name=${SERVICE\_ACCOUNT} \

--bind=profile.machine="n2-highmem-4" \

--bind=profile.job\_name="iikf2020-ftrat5-pop1000-ann-batch-job" \

--bind=profile.date=${PROJECT\_DATE} \

--bind=profile.environment=prod

```

### ftrat5 pop10000

```

export SERVICE\_ACCOUNT='cassowary'

export GCP\_PROJECT\_NAME='twttr-recos-ml-prod'

export PROJECT\_DATE='2022-07-24T18'

bin/d6w schedule \

${GCP\_PROJECT\_NAME}/us-central1/iikf2020-ftrat5-pop10000-ann-batch-job \

src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet/iikf2020-ftrat5-pop10000-ann-batch-job.d6w \

--bind=profile.project=${GCP\_PROJECT\_NAME} \

--bind=profile.user\_name=${SERVICE\_ACCOUNT} \

--bind=profile.machine="n2-highmem-4" \

--bind=profile.job\_name="iikf2020-ftrat5-pop10000-ann-batch-job" \

--bind=profile.date=${PROJECT\_DATE} \

--bind=profile.environment=prod

```

### Deschedule

```

export SERVICE\_ACCOUNT='cassowary'

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-iikf2020-decayed-sum-ann-batch-job

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-iikf2020-ftrat5-pop1000-ann-batch-job

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-iikf2020-ftrat5-pop10000-ann-batch-job

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-iikf2020-decayed-sum-ann-batch-job

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-iikf2020-ftrat5-pop1000-ann-batch-job

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-iikf2020-ftrat5-pop10000-ann-batch-job

```

### pop1000-rnkdecay11

```

export SERVICE\_ACCOUNT='cassowary'

export GCP\_PROJECT\_NAME='twttr-recos-ml-prod'

export PROJECT\_DATE='2022-08-27T16'

bin/d6w schedule \

${GCP\_PROJECT\_NAME}/us-central1/ftr-pop1000-rnkdecay11-tweet-index-generation-batch-job \

src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet/ftr-based-simclusters-index-generation-job.d6w \

--bind=profile.project=${GCP\_PROJECT\_NAME} \

--bind=profile.user\_name=${SERVICE\_ACCOUNT} \

--bind=profile.machine="n2-standard-2" \

--bind=profile.job\_name="ftr-pop1000-rnkdecay11-tweet-index-generation-batch-job" \

--bind=profile.build\_target="src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet:ftr-tweet-index-generation-pop1000-rnkdecay11-job" \

--bind=profile.date=${PROJECT\_DATE} \

--bind=profile.environment=prod

```

### pop10000-rnkdecay11

```

export SERVICE\_ACCOUNT='cassowary'

export GCP\_PROJECT\_NAME='twttr-recos-ml-prod'

export PROJECT\_DATE='2022-08-27T16'

bin/d6w schedule \

${GCP\_PROJECT\_NAME}/us-central1/ftr-pop10000-rnkdecay11-tweet-index-generation-batch-job \

src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet/ftr-based-simclusters-index-generation-job.d6w \

--bind=profile.project=${GCP\_PROJECT\_NAME} \

--bind=profile.user\_name=${SERVICE\_ACCOUNT} \

--bind=profile.machine="n2-standard-2" \

--bind=profile.job\_name="ftr-pop10000-rnkdecay11-tweet-index-generation-batch-job" \

--bind=profile.build\_target="src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet:ftr-tweet-index-generation-pop10000-rnkdecay11-job" \

--bind=profile.date=${PROJECT\_DATE} \

--bind=profile.environment=prod

```

### decayed\_sum

```

export SERVICE\_ACCOUNT='cassowary'

export GCP\_PROJECT\_NAME='twttr-recos-ml-prod'

export PROJECT\_DATE='2022-09-05T16'

bin/d6w schedule \

${GCP\_PROJECT\_NAME}/us-central1/decayed-sum-tweet-index-generation-batch-job \

src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet/ftr-based-simclusters-index-generation-job.d6w \

--bind=profile.project=${GCP\_PROJECT\_NAME} \

--bind=profile.user\_name=${SERVICE\_ACCOUNT} \

--bind=profile.machine="n2-standard-2" \

--bind=profile.job\_name="decayed-sum-tweet-index-generation-batch-job" \

--bind=profile.build\_target="src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet:ftr-tweet-index-generation-decayed-sum-job" \

--bind=profile.date=${PROJECT\_DATE} \

--bind=profile.environment=prod

```

### OON ftrat5

```

export SERVICE\_ACCOUNT='cassowary'

export GCP\_PROJECT\_NAME='twttr-recos-ml-prod'

export PROJECT\_DATE='2022-09-21T16'

bin/d6w schedule \

${GCP\_PROJECT\_NAME}/us-central1/oon-ftr-pop1000-rnkdecay-tweet-index-generation-batch-job \

src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet/ftr-based-simclusters-index-generation-job.d6w \

--bind=profile.project=${GCP\_PROJECT\_NAME} \

--bind=profile.user\_name=${SERVICE\_ACCOUNT} \

--bind=profile.machine="n2-standard-2" \

--bind=profile.job\_name="oon-ftr-pop1000-rnkdecay-tweet-index-generation-batch-job" \

--bind=profile.build\_target="src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/ftr\_tweet:oon-ftr-tweet-index-generation-pop1000-rnkdecay-job" \

--bind=profile.date=${PROJECT\_DATE} \

--bind=profile.environment=prod

```

### Deschedule

```

export SERVICE\_ACCOUNT='cassowary'

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-ftr-pop1000-rnkdecay11-tweet-index-generation-batch-job

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-ftr-pop1000-rnkdecay11-tweet-index-generation-batch-job

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-ftr-pop10000-rnkdecay11-tweet-index-generation-batch-job

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-ftr-pop10000-rnkdecay11-tweet-index-generation-batch-job

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-decayed-sum-tweet-index-generation-batch-job

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-decayed-sum-tweet-index-generation-batch-job

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-oon-ftr-pop1000-rnkdecay-tweet-index-generation-batch-job

aurora cron deschedule atla/${SERVICE\_ACCOUNT}/prod/twttr-recos-ml-prod-us-central1-oon-ftr-pop1000-rnkdecay-tweet-index-generation-batch-job

```