

An example for Association Rule Mining using the Eunomia package

11/12/2020

Contents

0.1	Run Apriori	2
0.2	Run Eclat	3
0.3	Run FP-Growth	3
0.4	Run Relim	3

0.0.1 Connect to the database

```
### Define database parameters
cdmdatabaseschema = "main"
resultsdatabaseschema = "main"
arm_inputFile = "testing.txt"
arm_outputFile = "testingResults.txt"

connectionDetails <- Eunomia::getEunomiaConnectionDetails()
connection <- connect(connectionDetails)
#on.exit(DatabaseConnector::disconnect(connection)) #Close db connection on error or exit
```

0.0.1.1 Define cohort There are two options for defining a cohort: - 1) Define the cohort in ATLAS and export the SQL file - 2) define it locally within R

```
cohort <- "SELECT person_id AS subject_id,
  condition_start_date AS cohort_start_date
INTO #diagnoses
FROM @cdm.condition_occurrence
WHERE condition_concept_id IN (
  SELECT descendant_concept_id
  FROM @cdm.concept_ancestor
  WHERE ancestor_concept_id = 4329847 -- Myocardial infarction
)
AND condition_concept_id NOT IN (
  SELECT descendant_concept_id
  FROM @cdm.concept_ancestor
  WHERE ancestor_concept_id = 314666 -- Old myocardial infarction
);
INSERT INTO @cdm.cohort (subject_id, cohort_start_date, cohort_definition_id)
```

```

SELECT subject_id,
       cohort_start_date,
       CAST (1 AS INT) AS cohort_definition_id
FROM #diagnoses
INNER JOIN @cdm.visit_occurrence
  ON subject_id = person_id
   AND cohort_start_date >= visit_start_date
   AND cohort_start_date <= visit_end_date
WHERE visit_concept_id IN (9201, 9203, 262); -- Inpatient or ER;"

renderTranslateExecuteSql(connection, cohort, cdm = cdmdatabaseschema)

sql <- "ALTER TABLE #diagnoses ADD cohort_definition_id INT NOT NULL DEFAULT(1)"

# Execute the script to receive the data
renderTranslateExecuteSql(connection, sql)

querySql(connection, "SELECT count(*) FROM diagnoses;")

```

0.0.2 Get the data and close the connection

```

#### Feature Extraction ####
covariateSettings <- createCovariateSettings(useConditionOccurrenceAnyTimePrior = TRUE,
                                           #includedCovariateIds = c(),
                                           #includedCovariateConceptIds = c()
                                           )

covariateData_eunomia <- getDbCovariateData(connection = connection, cdmDatabaseSchema = cdmdatabaseschema)

disconnect(connection)

getInputFileForAssociationRules(covariateDataObject = covariateData_eunomia, fileToSave = arm_inputFile)

## Getting covariate names from covariate data object and writing text file...

## Preparing input data for Association Rules...

## Input data has been created succesfully and saved in testing.txt

```

0.1 Run Apriori

```

apriori_associationSets <- runAssociationRules(algorithm = "Apriori",
                                              inputFile = arm_inputFile,
                                              outputFile = arm_outputFile,
                                              minsup = 0.5 )

head(apriori_associationSets)

```

0.2 Run Eclat

```
eclat_associationSets <- runAssociationRules(algorithm = "Eclat",
                                             inputFile = arm_inputFile,
                                             outputFile = arm_outputFile,
                                             minsup = 0.5 )
head(eclat_associationSets)
```

0.3 Run FP-Growth

```
fpgrowth_associationSets <- runAssociationRules(algorithm = "FP-Growth",
                                                inputFile = arm_inputFile,
                                                outputFile = arm_outputFile,
                                                minsup = 0.5 )
head(fpgrowth_associationSets)
```

0.4 Run Relim

```
relim_associationSets <- runAssociationRules(algorithm = "Relim",
                                             inputFile = arm_inputFile,
                                             outputFile = arm_outputFile,
                                             minsup = 0.5 )
head(relim_associationSets)
```