Steps to run the Fril – SVM experiment:

1. Install Jetty Server
2. Download the War file for Fril Connectivity and place it in webapps folder in Jetty.
   https://dl.dropboxusercontent.com/u/103236727/Linkage-0.2.war

3. Download the War file for SVM Connectivity and place it in webapps folder in Jetty.
   https://dl.dropboxusercontent.com/u/103236727/orders-server-example.war

4. Download the Jar file for running standalone program to calculate Precision and Recall.

   https://dl.dropboxusercontent.com/u/103236727/PrecisonRecallCalculator.jar

5. Download the configuration file required for step4.
   https://dl.dropboxusercontent.com/u/103236727/RecordLinkage.Properties

6. The Files obtained in step 4 and 5 can be placed in one folder, since the jar at step 4 requires property file at step 5.

7. To start the experiment, enable Google Advance Rest Client in Chrome operating system.

8. **Phase1 Experiment:**

   Use the following link in the URL section of Google Advance Rest Client:
   http://localhost:8080/Linkage-0.2/v1/link

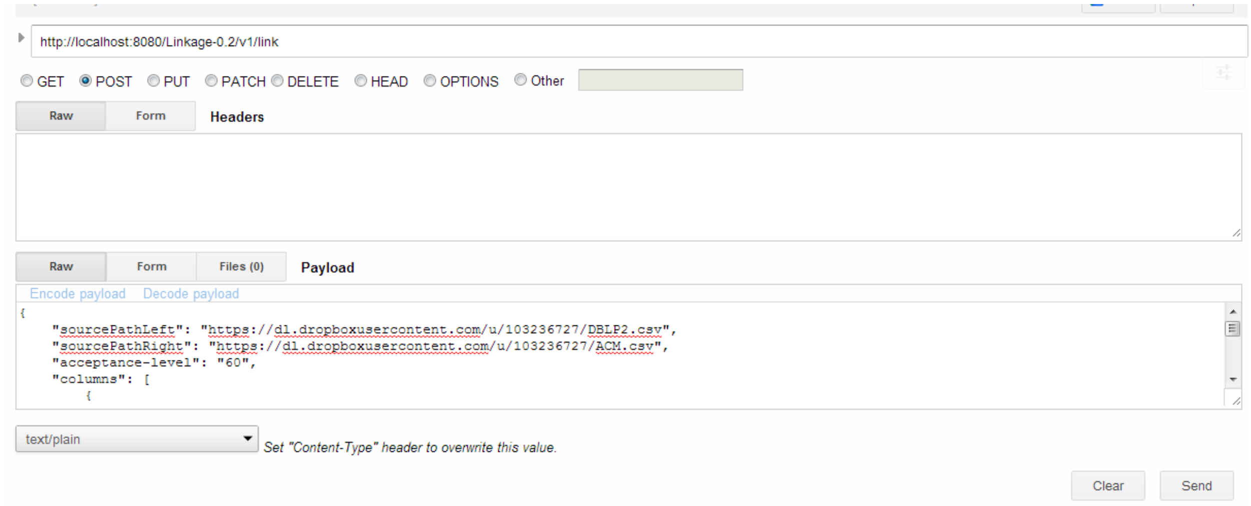
   Use the POST method

   Paste the following JSON in the payload section

```
{
  "sourcePathLeft": "https://dl.dropboxusercontent.com/u/103236727/DBLP2.csv",
  "sourcePathRight": "https://dl.dropboxusercontent.com/u/103236727/ACM.csv",
  "acceptance-level": "60",
  "columns": [
    {
      "columnName": "title",
      "algorithm": "EditDistance",
      "weight": "",
      "params": {
        "match-level-start": "",
        "math-level-end": ""
      }
    },
    {
      "columnName": "authors",
      "algorithm": "EditDistance",
      "weight": "",
      "params": {
        "match-level-start": "",
        "math-level-end": ""
      }
    },
    {
      "columnName": "venue",
      "algorithm": "EditDistance",
      "weight": "",
      "params": {
        "match-level-start": "",
        "math-level-end": ""
      }
    },
    {
      "columnName": "year",
      "algorithm": "NumericDistance",
      "weight": "",
      "params": {
        "use-lineral-approximation": "true",
        "percent-difference": "",
        "numeric-difference": ""
      }
    }
  ]
}
```

The **result.csv** file will be placed in the Jetty Distribution Folder.

Following is the screenshot to do the above mentioned steps:



http://localhost:8080/Linkage-0.2/v1/link

○ GET ● POST ○ PUT ○ PATCH ○ DELETE ○ HEAD ○ OPTIONS ○ Other

```
{
    "sourcePathLeft": "https://dl.dropboxusercontent.com/u/103236727/DBLP2.csv",
    "sourcePathRight": "https://dl.dropboxusercontent.com/u/103236727/ACM.csv",
    "acceptance-level": "60",
    "columns": [
        {
```

text/plain

Set "Content-Type" header to overwrite this value.

Phase 2:

This step will calculate the precision and recall of results obtained. Please note that this program is data specific. This would only work for this data type or data type with similar attributes.

Place the configuration file and Jar file from step 4 and 5.

Following are the contents of configuration file:

PATH_JSON_PHASE1=C:\\Users\\asant\\Desktop\\firstphase.json
For the first iteration, this is the default JSON mentioned above. The standalone program will update this json after the phase 2 with new weights.

PATH_SRC1_LOCAL=C:\\Users\\asant\\Desktop\\Directed Research\\DBLP-ACM\\DBLP2.csv
This is the data source 1 file can be obtained from https://dl.dropboxusercontent.com/u/103236727/DBLP2.csv. Place this file on the path mentioned.

PATH_SRC2_LOCAL=C:\\Users\\asant\\Desktop\\Directed Research\\DBLP-ACM\\ACM.csv
This is data source 1 file can be obtained from https://dl.dropboxusercontent.com/u/103236727/ACM.csv. Place this file on the path mentioned.

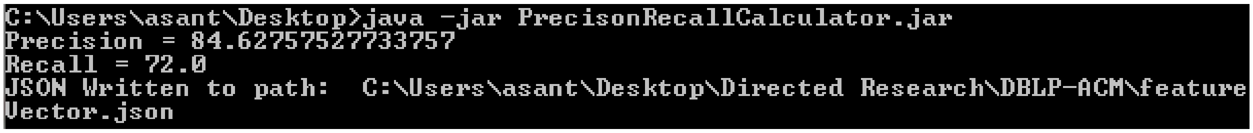
PATH_PERFECT_MAPPING_LOCAL=C:\\Users\\asant\\Desktop\\Directed Research\\DBLP-ACM\\DBLP-ACM_perfectMapping.csv
This is data source 1 file can be obtained from https://dl.dropboxusercontent.com/u/103236727/DBLP-ACM_perfectMapping.csv
Place this file on the path mentioned. This is the perfect mapping.

PATH_RESULT_PHASE1=C:\\Users\\asant\\Desktop\\Directed Research\\DBLP-ACM\\result.csv
This is the result obtained from the Phase1 by FRIL.

PATH_JSON_PHASE2=C:\\Users\\asant\\Desktop\\Directed Research\\DBLP-ACM\\featureVector.json
This is the path where the new json for PHASE2 having the feedback and the rows are present. The standalone program writes in this file.

TOTAL_CORRECT_ANS_COUNT=2225
This is the total answer count.

Command to run this standalone program:
Java –jar PrecisionRecallCalculator.jar

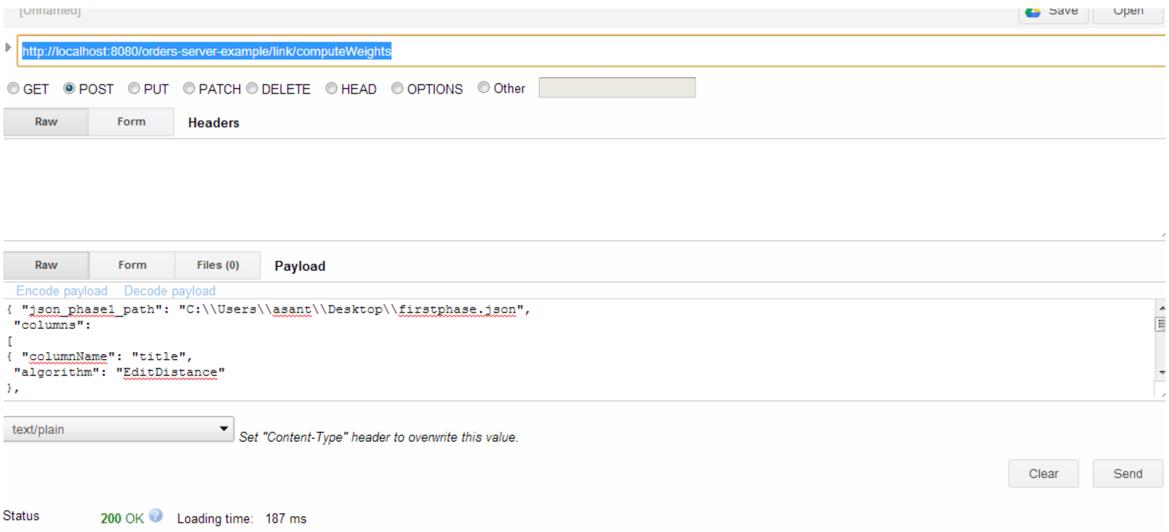


```
C:\Users\asant\Desktop>java -jar PrecisonRecallCalculator.jar
Precision  = 84.62757527733757
Recall = 72.0
JSON Written to path:  C:\Users\asant\Desktop\Directed Research\DBLP-ACM\feature
Vector.json
```

Phase 3: Take the JSON produced above.
Path: [PATH_JSON_PHASE2=C:\\Users\\asant\\Desktop\\Directed Research\\DBLP-ACM\\featureVector.json]

And run the phase 3 using Google Advance Rest Client.

Link: http://localhost:8080/orders-server-example/link/computeWeights
Method: POST
Payload: The contents of featureVector.json



http://localhost:8080/orders-server-example/link/computeWeights

○ GET ● POST ○ PUT ○ PATCH ○ DELETE ○ HEAD ○ OPTIONS ○ Other

```
{ "json_phase1_path": "C:\\Users\\asant\\Desktop\\firstphase.json",
 "columns":
[
{ "columnName": "title",
 "algorithm": "EditDistance"
},
```

text/plain *Set "Content-Type" header to overwrite this value.*

Status 200 OK Loading time: 187 ms

It produces an output like the following; this can be again fed as a configuration to phase 1 and the steps can be repeated.

```
{ "sourcePathLeft":"https://dl.dropboxusercontent.com/u/103236727/DBLP2.csv",
"sourcePathRight":"https://dl.dropboxusercontent.com/u/103236727/ACM.csv",
"acceptance-level":"60",
"columns": [{
"columnName":"title",
"algorithm":"EditDistance",
"weight":"30",
"params": {
 "match-level-start": "0.1",
"math-level-end": "0.9"
}
},{
"columnName":"authors",
"algorithm":"EditDistance",
"weight":"25",
"params": {
 "match-level-start": "0.1",
"math-level-end": "0.9"
}
},{
"columnName":"venue",
"algorithm":"EditDistance",
"weight":"24",
"params": {
 "match-level-start": "0.1",
"math-level-end": "0.9"
}
},{
"columnName":"year",
"algorithm":"NumericDistance",
"weight":"18",
"params": {
 "use-lineral-approximation": "true",
"percent-difference": "5.0,5.0",
"numeric-difference": ""
}
}]}
```