



**Community Meetup - 7 June, 2022**

# Action Items from last meetup

**Development environment setup guide**

**<https://docs.zingg.ai/zingg/contributing/settingupzingg>**

**Code walk through - <https://github.com/zinggAI/zingg/wiki/Code-Walkthrough>**

**Planned for next release**

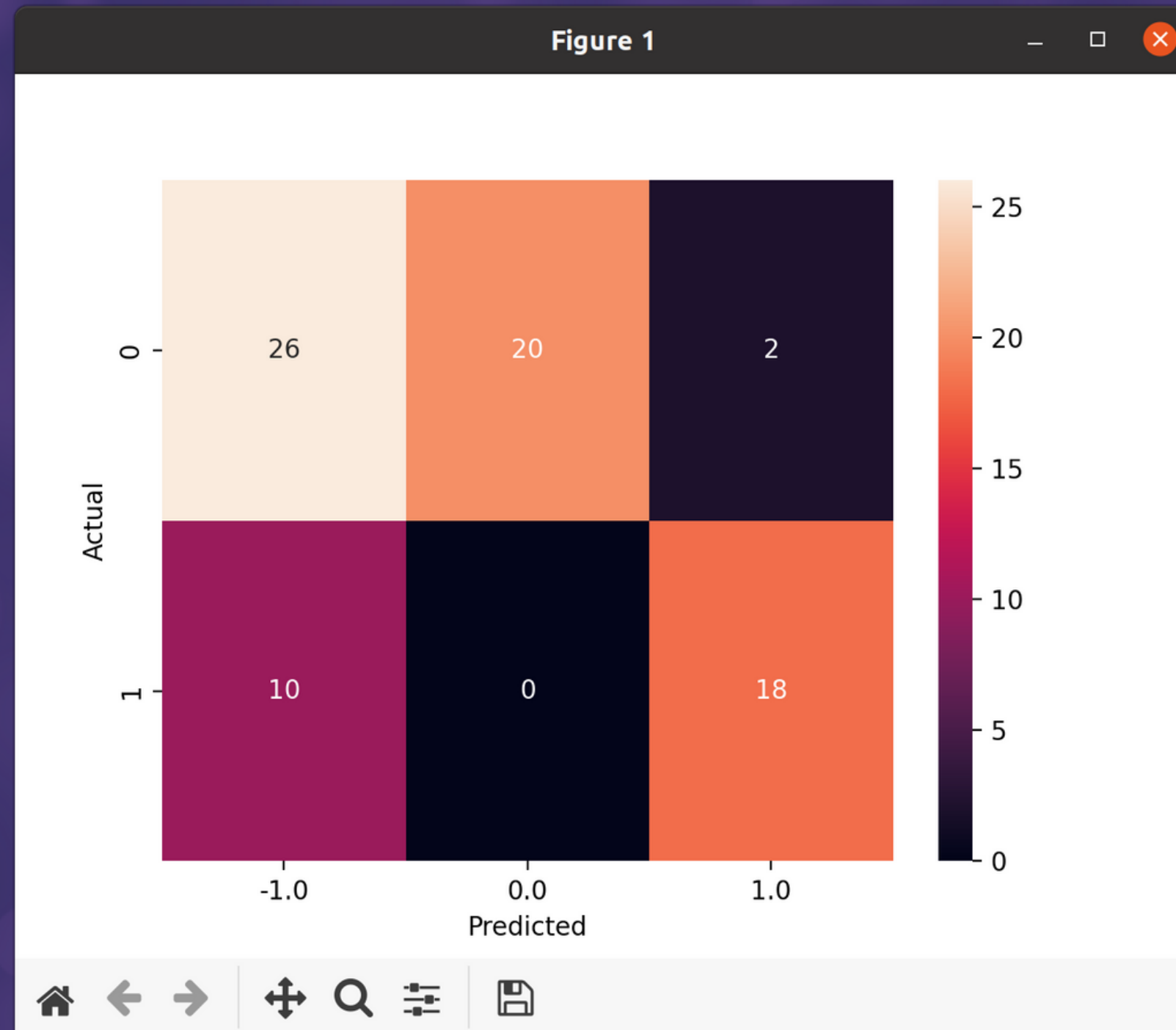


# Python Interface

```
1 from zingg import *
2 from pyspark.sql import DataFrame
3 import pandas as pd
4 import seaborn as sn
5 import matplotlib.pyplot as plt
6 from IPython.display import display
7
8 args = Arguments()
9 fname = FieldDefinition("fname", "string", MatchType.type("FUZZY"), MatchType.type("EXACT"), MatchType.type("PINCODE"))
10 lname = FieldDefinition("lname", "string", MatchType.type("FUZZY"), MatchType.type("EXACT"))
11 fieldDef = [fname, lname]
12 options = ClientOptions()
13 inputPipe = Pipe("test", "csv")
14 inputPipe.addProperty("location", "examples/febr1/test.csv")
15 args.setData(inputPipe)
16 args.setModelId("100")
17 args.setZinggDir("models")
18 args.setNumPartitions(4)
19 args.setLabelDataSampleSize(0.5)
20 args.setFieldDefinition(fieldDef)
21 Arguments.writeArgumentsToJSON("python/ArgumentsToFile.json", args.getArgs())
22 #Zingg execution for the given phase
23 client = Zingg(args, options)
24 client.init()
25 client.execute()
26
27 jMarkedDF = client.getMarkedRecords()
28 markedDF = client.getDfFromDs(jMarkedDF)
29 pMarkedDF = client.getPandasDfFromDs(jMarkedDF)
30 display(pMarkedDF)
31
32 jUnMarkedDF = client.getUnMarkedRecords()
33 unMarkedDF = client.getDfFromDs(jUnMarkedDF)
34 pUnMarkedDF = client.getPandasDfFromDs(jUnMarkedDF)
35
```

**pip install zingg**

# How well is Zingg learning?



# Stopwords

- **Recommendations for common words like LLP, LLC, Pvt, Ltd, Mr etc which can be ignored during matching**
- **configured per column**

# **More AND Multiple Match types**

- Partially populated fields**
- Email, Zip code**
- Alphanumeric**

# **Big Query/MySQL/PostGres...**

- Improvement to generic read/write pipes**



# Other useful things

- **smaller jar**
- **better error messages(hopefully!)**
- **more blocking functions for numeric types**
- **junits(for our sanity :-))**

**Thanks - we just  
crossed 525  
stargazers!**

