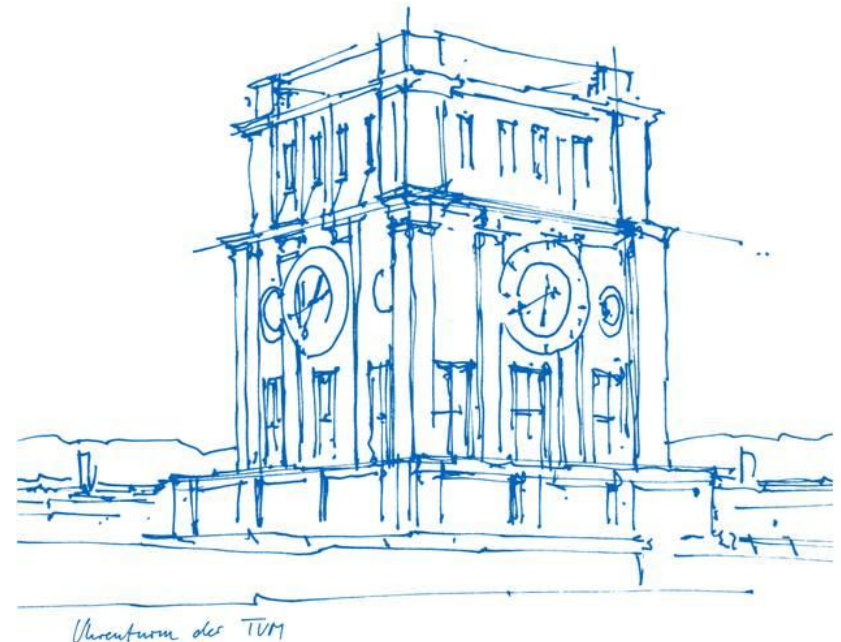# NLP - Opinion Mining Lab (Practicum)

1.4 End-to-End Modeling and Characterization of Crowdsourcing Annotators

Team : Soner Durmaz, Alex Marin Felices and

Saumya Goyal

Technische Universität München

Fakultät für Informatik

Munich, 26. July 2021

Uhrenturm der TUM

Topics

1. Problem Statement

2. Dataset

3. Approach

4. Results

5. Conclusion

1. Problem Statement

- **Primary**: Finding Bias matrix of all the annotators

- **Secondary**:
  - Using the above extracted feature in detecting Spammers

  - Clustering the like-minded annotators

Spammer Detection

Cluster annotators

Bias Matrices
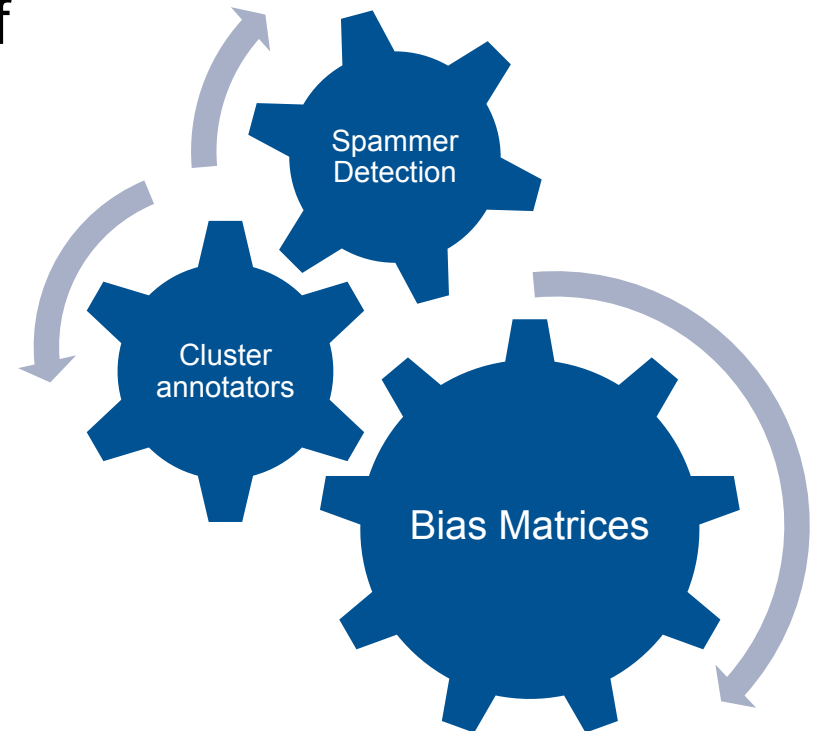
2. Dataset

- 1,373 comments

- 10,439 sentences - 4,616 relevant

- 10 different annotators

- Singly labelled

- Preprocessing

  - Stop words

  - Lower case

  - Lemmatization

  - GloVe50
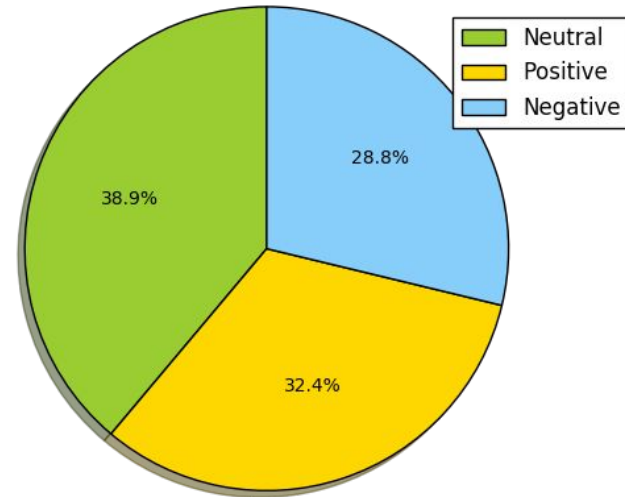
3. Our Approach

1) End-to-End Annotator <u>Bias</u> Approximation (**Primary** Goal)
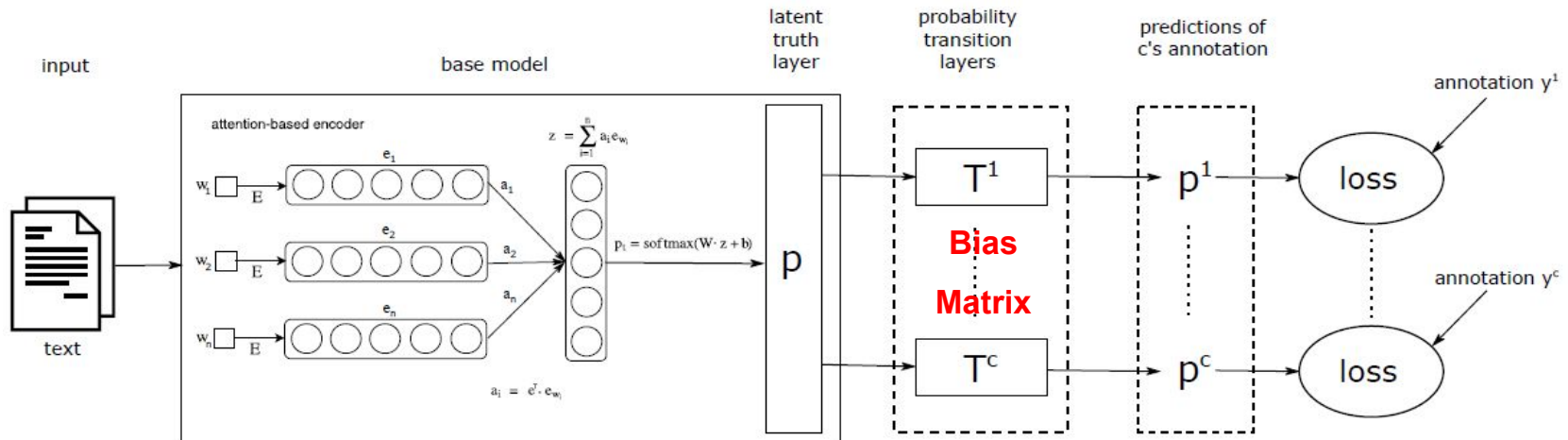
2) Finding Ground Truth

3) Our Hybrid Architecture

Predicted labels

	P	N	NT
P	✓		
N		✓	
NT			✓

Observed labels

BIAS MATRIX

1) **End-to-End** Annotator Bias Approximation

While training :

- **Input Text to prediction -** different from other approaches on crowdsourced data

- **Learned parameters** : Attention Vectors, Linear Layer - W and b, **Bias Matrix**

  On test data :

- **Input -** One Sentence

- **Output -** What will the annotator A1, A2, … A10 will annotate it as - P, N, NT

2) Finding **Ground Truth**

Algorithm 1 The Fast Dawid-Skene Algorithm

**Input** Crowdsourced choices of Q questions by A participants (annotators) from C choices

Output Proposed true choices - T_{qc}

1: Estimate Ts using majority voting.
2: **repeat**
3:      *M-step:* Obtain the parameters, $P(c_a|Y_q = c)$ and $P(Y_q = c)$ using Equations 3 and 4
4:      *E-step:* Estimate Ts using the parameters, $P(c_a|Y_q = c)$ and $P(Y_q = c)$, and with the help of Equations 2 and 1.
5: **until** convergence

- **Why? :** Intuitive!
- **Why Fast David Skene**: It is famously used for ground truth estimation in crowdsourced data
- **Problem** : Needed multiple-annotations for each sentence, and we had single annotation
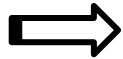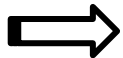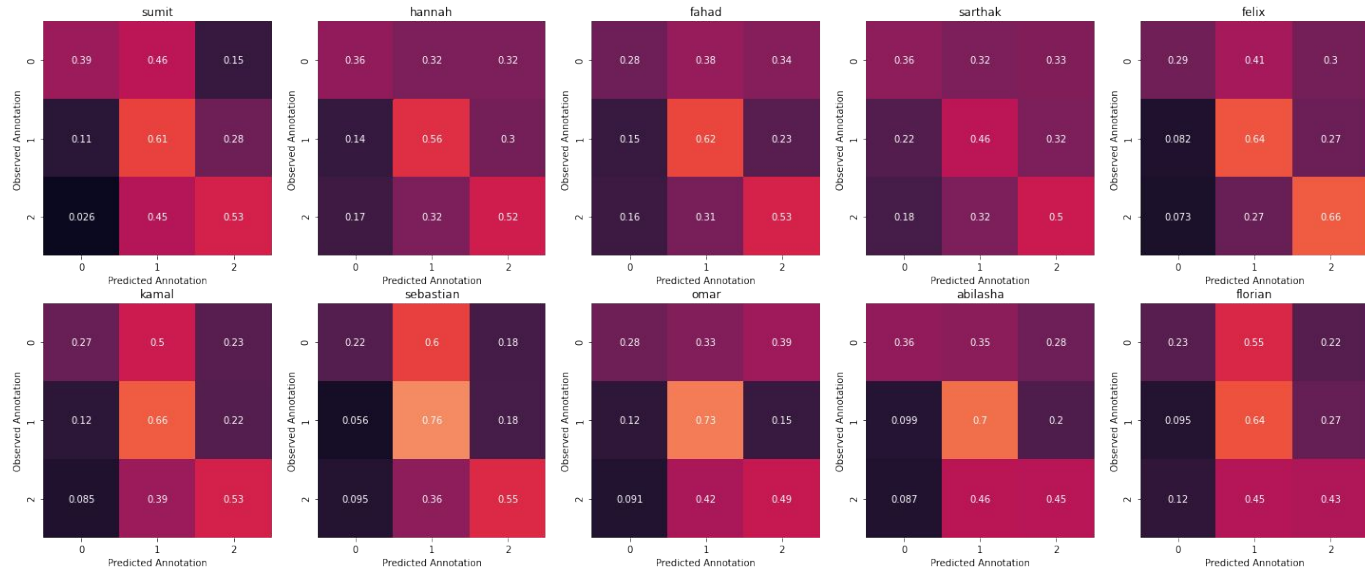
Did we succeed?

3) Our **Hybrid** Architecture
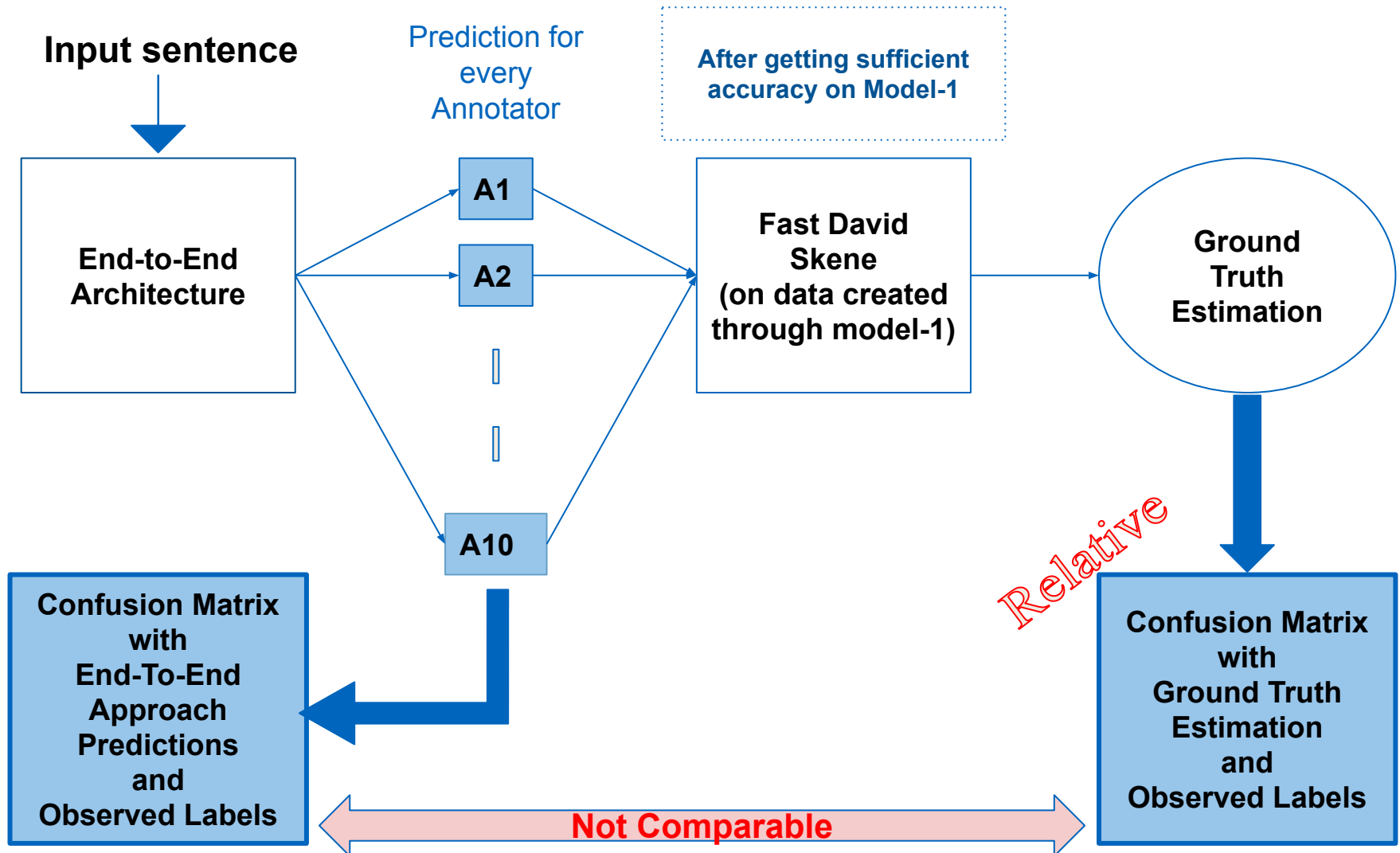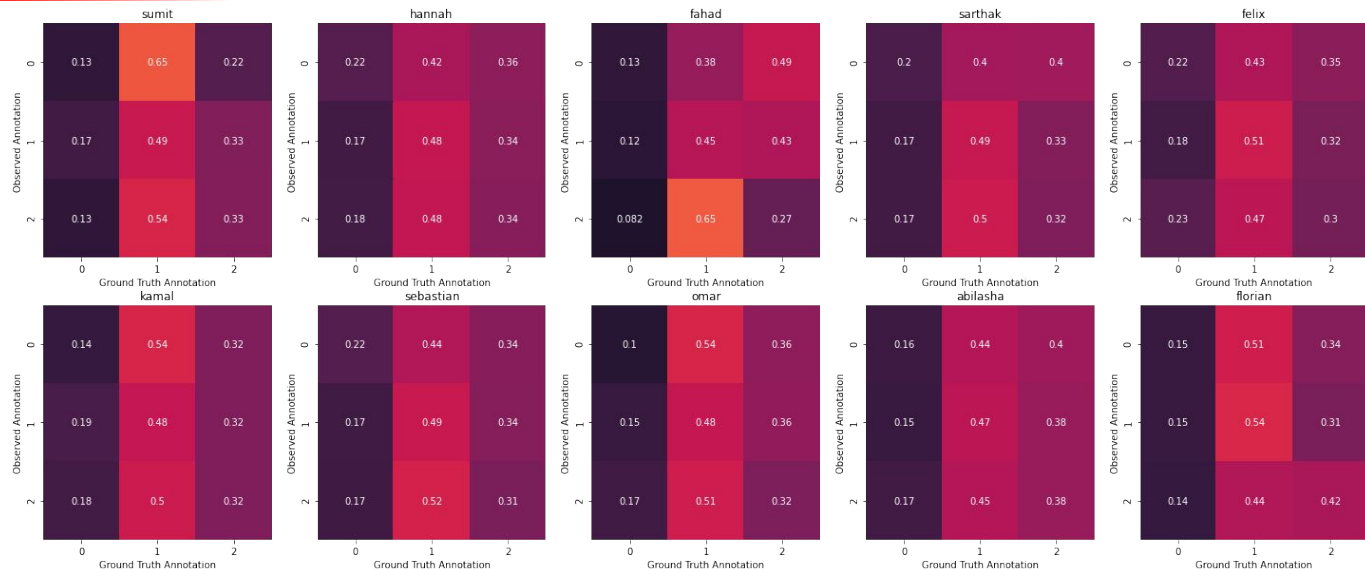
Confusion matrices of both methods

**Confusion Matrix - 1
End-to-End Approach**
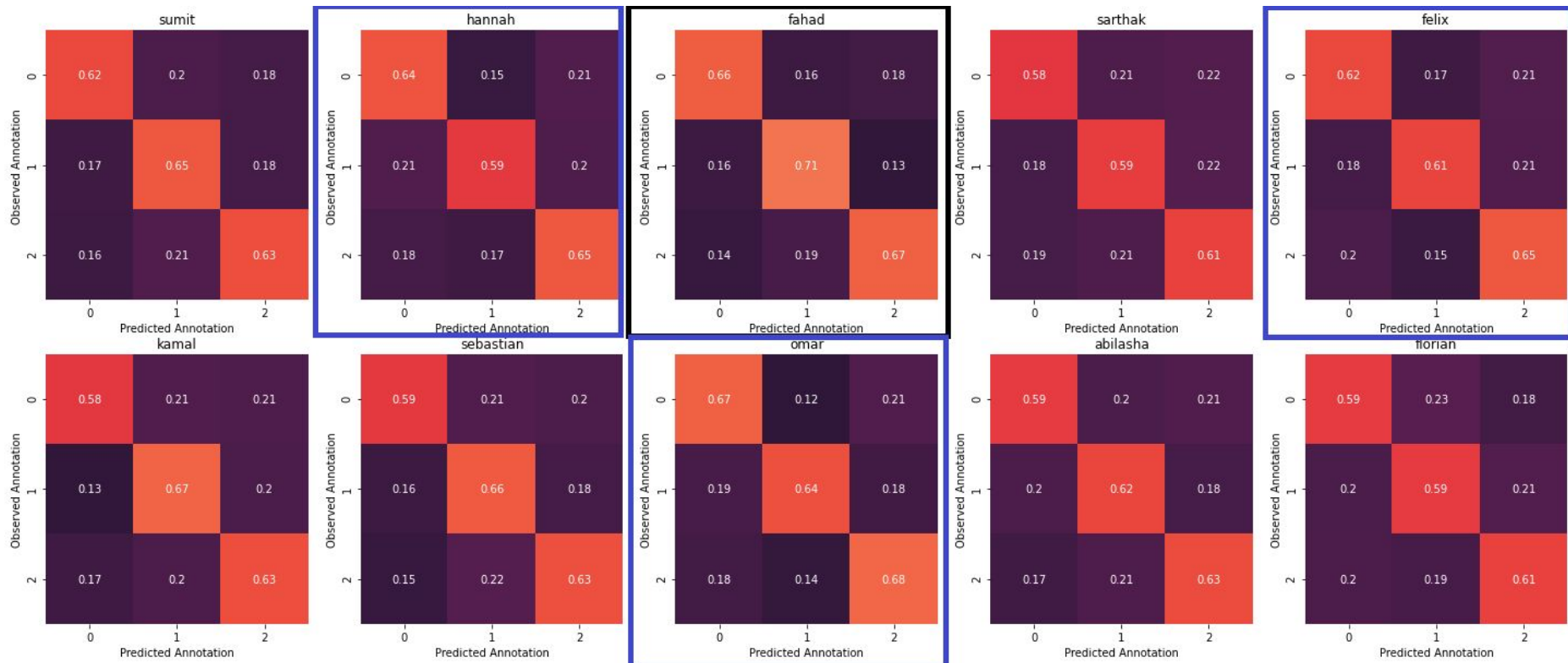
Accuracy :
46%
F1-score:
38%

**Confusion Matrix - 2
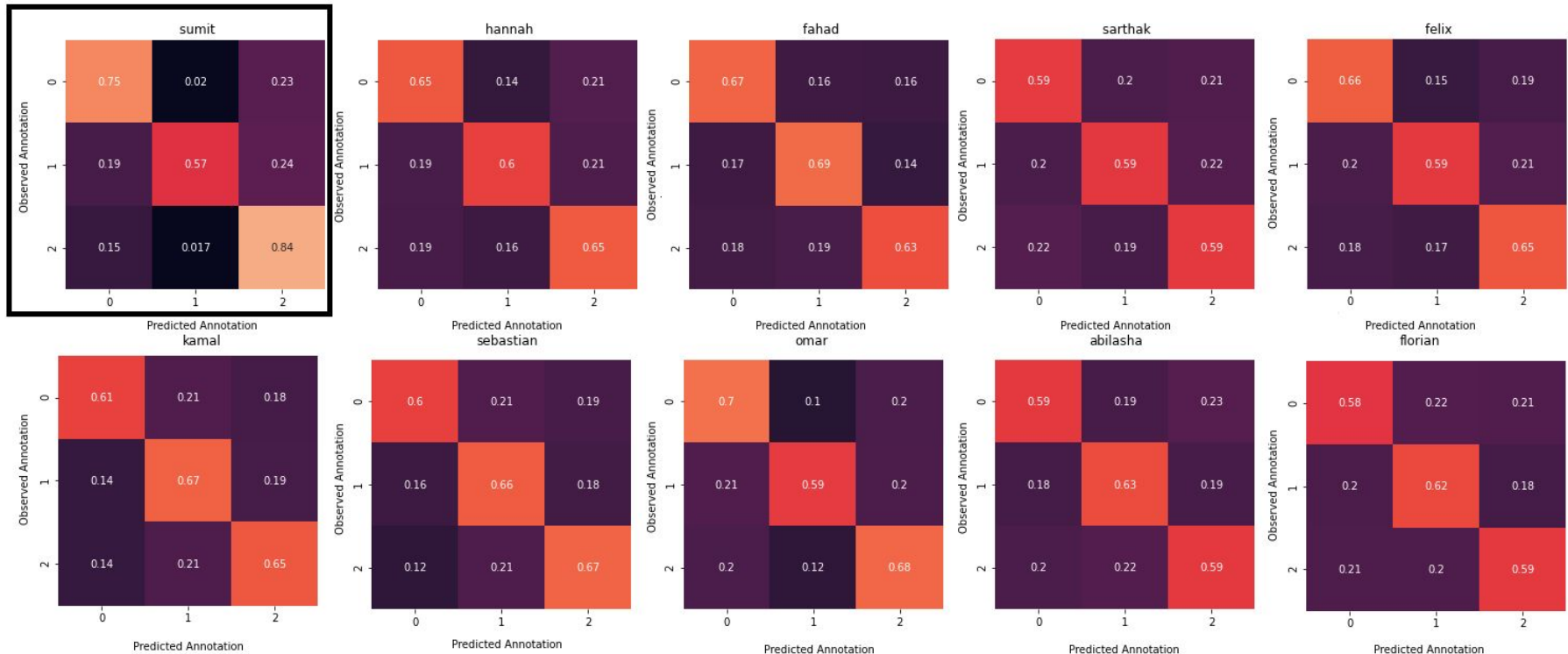After finding Ground
Truth**

4. **Results**:
Annotator Clustering Based On Bias Matrices

USING BIAS MATRIX

4. **Results**:
Spammer Detection with Manipulated Data

USING BIAS MATRIX

5. Conclusion

Our Limitations:

1) The dataset is small for a singly labeled dataset, on an average there were 450(approx) samples.

2) Our Computer resources were less as we had to run many epochs for our experiments - Google Colab also took 2 hours/run

Conclusions:

1) Both the approaches are not comparable -

   a) **End-to-End** is more like percentage - individual bias finding
   b) **FDS** is more like percentile - relative to all the annotators
2) To get more accurate results we need more data.

Thank You!