

Appendix

1 Evaluation Tool

There have been many efforts in Question Answering in multiple languages. To evaluate each model with the actual and predicted answers we have created an Evaluation Tool¹. The tool provides many features to the researcher to analyze the results with user-friendly UI/UX. We have used Flask² as a backend by synchronizing with the evaluation script. Our tool can list the best and worst-performing samples for further analysis and can display collective EM and F1 score for the input samples. Preview images of the tool are given in the Figures 1, 2, 3.

Our code and evaluation scripts are available at : <https://github.com/Question-Answering-Analytic-Tool>

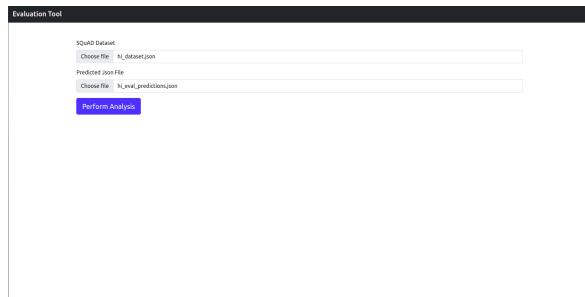


Figure 1: Upload training dataset and evaluated prediction json file.

¹Concerning evaluation script available for SQuAD <https://rajpurkar.github.io/SQuAD-explorer/>

²Jaccard score <https://flask.palletsprojects.com/en/2.0.x/>

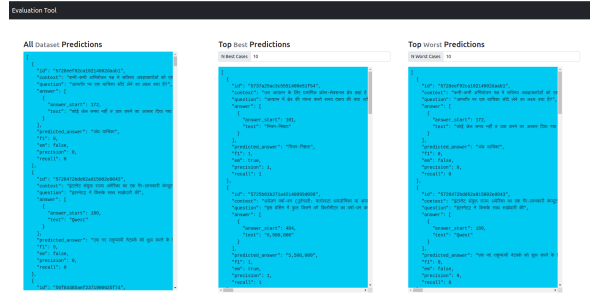


Figure 2: Generating preview of- All Prediction, n-Best predictions and n-Worst predictions after evaluation. predictions

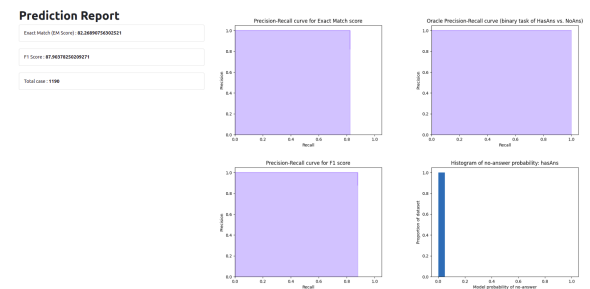


Figure 3: Prediction Report with F1 score and Exact Match

2 Mathematical representation of Jaccard Score and WER:

Formula of Jaccard Score ¹ and Word error rate (WER) ² are mentioned below:

$$J(A^A, A^P) = \frac{|A^A \cap A^P|}{|A^A \cup A^P|} \quad (1)$$

Jaccard Score: A^A and A^P represents the actual and predicted answers respectively.

$$WER = \frac{S + D + I}{N} \quad (2)$$

WER: S is the number of substitutions, D is the number of deletions, I is the number of insertions, N is the number of words in the reference

3 Dataset Format

Figure 4 shows the format of dataset which we have compiled by combining MLQA³ and XQuAD⁴.

```
{
  "answers": {
    "answer_start": [
      378
    ],
    "text": [
      "उत्तर पूर्व"
    ]
  },
  "context": "उत्ती \ "एरिया XX \ " नामकरण प्रणाली का प्रयोग नेवादा परीक्षण स्थल के अन्य भागों के लिए किया गया है।मूल रूप में 6 बटे 10 मील का यह आयताकार अड्डा अब तथाकथित 'ग्रूम बॉक्स \ " का एक भाग है, जो कि 23 बटे 25.3 मील का एक प्रतिबंधित हवाई क्षेत्र है। यह क्षेत्र NTS के आंतरिक सड़क प्रबंधन से जुड़ा है, जिसकी पक्की सड़कें दक्षिण में मरकरी की ओर और पश्चिम में युक्का फ्लैट की ओर जाती हैं। झील से उत्तर पूर्व की ओर बढ़ते हुए व्यापक और और सुव्यवस्थित ग्रूम झील की सड़कें एक दर्रे के जरिये पेचीदा पहाड़ियों से होकर गुजरती हैं। पहले सड़कें ग्रूम घाटी",
  "id": "eeb8dbd25efe5221dc6723dde95daa07d2c8478",
  "question": "झील के सापेक्ष ग्रूम लेक रोड कहाँ जाती थी?",
  "title": "एरिया 51"
}
```

Figure 4: Json structure of dataset which we have used to fine-tune and evaluate the model.

³MLQA <https://github.com/facebookresearch/MLQA>

⁴XQuAD <https://github.com/deepmind/xquad>