Competition: Hindi to English Machine Translation System

Shruti Sharma
20111061
{shruti20}@iitk.ac.in
Indian Institute of Technology Kanpur (IIT Kanpur)

April 23, 2021

Abstract

Write a brief abstract about your models, competition, your rank and scores for various evaluation metrics, etc. Abstract should not be more than 100 words long.

Some Guidelines

1. Please do not show your creativity and do not change the title and follow the format that is there above.
2. Follow the sections given below, you should have the same sections in your report.
3. **The maximum length of the report is maximum 8 pages (without references), not even a single character more!** You can use any number of pages for references. This is a strict limit. If your report exceeds by even one character, you get a zero.
4. The sections given below are compulsory and you should cover the points mentioned below. Also I have written the things point-wise, of course you will write properly in paragraph format. However, you are free to have more things within the sections as well. You can create subsections within a section. Creativity and innovation counts!
5. There should be no grammar or spelling mistakes! Please use Grammarly, you have free subscription!
6. As already pointed out, we are very strict about plagiarism. We will be using softwares to check for plagiarism. So please cite the relevant papers/sources and do not copy paste from somewhere. Explain techniques/methods from your perspective. If your report is found to be plagiarized then the entire group gets a F grade in the course.
7. You can cite papers like this [1, 2]. You can create a footnote like this [1]
8. If you write any mathematical equations please use latex.
9. You can have images in your report but do not forget to cite them, if you have taken from somewhere. Do not paste an image directly from the papers, make sure the images are of high quality and should not get blurred on zooming. So best way to do it use either SVG images or PDF images. In general, it is always better to create images on your own.
10. Use the quotation marks properly, use these “Quote” instead of these ”Quote”
11. In your report, obviously remove this part.

1 Competition Result

Codalab Username: put only your username here
Final leaderboard rank on the test set: put the rank

[1] https://en.wikipedia.org

METEOR Score wrt to the best rank: put the score
BLEU Score wrt to the best rank: put the score
Link to the CoLab/Kaggle notebook: `www.google.com` Note that the code should be very well documented.

2 Problem Description

What are you trying to solve? Brief overview of the problem, the problem setting, etc.

3 Data Analysis

1. Describe the train dataset that has been provided to you.
2. Analysis of the data, e.g., corpus statistics, noise in the corpus, etc.
3. Test data will also be provided to you, so you can do analysis of that as well, e.g., how much does it differ from train data?
4. Some interesting insights about the data? You can use visualizations if you like. Be creative!

4 Model Description

1. Model evolution: Describe the models in detail that you experimented with in each of the phases, what were the key learnings in each phase that lead to making changes to model architecture or switching to a new model. You can have figures for model architectures.
2. If you took inspiration from an existing MT model, please cite the paper(s).
3. Detailed description of the final model that worked best on the test set. You can have figures for model architectures.
4. Model objective (loss) functions.
5. Inference: details of what kind of decoding strategy (e.g., greedy, beam search, etc.), did you use? What was the motivation for it? If you tried different strategies, describe the evolution from one strategy to another.

5 Experiments

1. Data Pre-processing you did both for source and target language. What was the reason for doing this kind of pre-processing.
2. Training procedure: Optimizer, learning rates, epochs, training time, etc for different models you tried
3. Details about different hyper-parameters for different models. You can use tabular format if you like. How did you arrive at these hyper-parameters?

6 Results

1. Results of different models on dev data in three phases in tabular format. If you didn't take part in any phase leave it blank.
2. In the results table show all metrics, as provided in the evaluation scripts. You can also have a column for rank on the leaderboard if you like.
3. Results of different models on the test data in tabular format.
4. Briefly explain the results, e.g., what was the best performing model (both in terms of METEOR and BLEU) on dev set, test set. Why do you think this model worked better than all other models?

7 Error Analysis

1. Error analysis of your different models both on the dev set and the test set. What models worked in general and in what kind of setting? What was the reason for it?
2. As you might have noticed that both the evaluation metrics (BLEU and METEOR) behave differently, maybe some sort of analysis about relationship between models and these scores.
3. Error analysis includes analysis of the model, you can use visualizations for this. For e.g., if you used attention mechanism you can use, attention heat maps, etc.
4. Analyze why models are not perfect? E.g., what kind of mistakes are made by the best model, how could these be overcome?
5. Since you will be provided with test data with true translations, do error analysis on that as well.
6. Any interesting insights!

8 Conclusion

Finally conclude the report with your key findings, what would you recommend? What would be possible future directions, etc.

References

[1] S. Janghorbani, A. Modi, J. Buhmann, and M. Kapadia, "Domain authoring assistant for intelligent virtual agents," in *18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019)*, (Montreal, Canada), 2019.

[2] P. Colombo, W. Witon, A. Modi, J. Kennedy, and M. Kapadia, "Affect-driven dialog generation," in *NAACL-HLT*, (Minneapolis, USA), 2019.