Adequacy-Fluency Metrics for Machine Translation Evaluation Project Scope Document

Karthikeyan Arumugam(2018900074)
Nikhil Bishnoi (2019114021)

Monsoon 2020

1   Abstract

In this project we would like to explore all Evaluation approaches and metrics. Our main objective is study and understand how Adequacy- Fluency Metrics We would like to experiment with the approached mentioned in our reference paper and provide a balanced approach towards balanced view on machine translation quality.

2   Approaches to this project work

1. To explore current Evaluation metrics and approaches

2. Study Traditional Evaluation approaches (WER, PER , BLEU , NIST)

3. Explore recent proposals such as METEOR , TER and MEANT.

4. Study Adequacy Evaluation Metrics

5. Fluency Evaluation Metrics

6. Experiment xAM-FM , mAM-FM

3 Proposed Experiments

- Language Pairs

    - Hindi - English
    - Tamil - English
    - Tamil - Hindi

- Machine translation System - To be Provided by Mentors/TAs

- Alternative - Google Translation APIs

4 References

1. Presentation by Haizhou Li - "https://lotus.kuee.kyoto-u.ac.jp/WAT/WAT2015/talk/IT_2.pd

2. Callison-Burch C., Fordyce C., Koehn P., Monz C. and Schroeder J. (2007), (Meta-) evaluation of machine translation, in Proceedings of Statistical Machine Translation Workshop, pp. 136-158

3. C. Tillmann et al., "Accelerated DP Based Search for Statistical Translation", in Proc. of the 5th European Conf. on Speech Commun. and Tech., Rhodos, Greece, Sept 1997, pp. 2667–2670.

4. K. Papineni et al., "BLEU: a method for automatic evaluation of machine translation", in Proc. of the 40th Annu. Meeting of the Assoc. for Computational Linguistics, Philadelphia, PA, USA, Jul 2002, pp. 311-318

5. G. Doddington, "Automatic evaluation of machine translation quality using n-gram co-occurrence statistics", in Proc. of the Human Lang. Tech. Conf., San Diego, CA, USA, Mar 2002

6. A. LavieandM.J. Denkowski, "The Meteor metric for automatic evaluation of machine translation", Machine Translation, vol. 23, pp. 105-115, May 2009

7. M. Snoveret al., "Study of Translation Edit Rate with Targeted Human Annotation", in Proc. of the 7th Biennial Conf. of the Assoc. for Mach. Translation in the Amer., Cambridge, MA, USA, Aug 2006

8. C.K. Lo and D. Wu, "MEANT: An inexpensive, high-accuracy, semi-automatic metric for evaluating translation utility based on seman-tic roles",in Proc. ofthe49th Annu. Meeting oftheAssoc. forComputation-alLinguistics, Portland, OR, USA, Jun 2011, pp. 220-229