Carnegie Mellon University
Language Technologies Institute

Alon Lavie

Unbabel / Carnegie Mellon University

Alon Lavie is currently the VP of Language Technologies at Unbabel, where he leads and manages Unbabel's US AI lab based in Pittsburgh, and provides AI strategic leadership company-wide. Prior to joining Unbabel, Alon was a senior manager at Amazon, where he led and managed the Amazon Machine Translation R&D group in Pittsburgh. Prior to that, Alon was the co-founder, President and CTO of Safaba, an MT technology company that was acquired by Amazon in 2015. For almost 20 years (1996-2015) Alon was a Research Professor at the Language Technologies Institute at Carnegie Mellon University, where he currently continues to serve as an adjunct Consulting Professor. Alon was President of the International Association for Machine Translation (IAMT) (2013-2015). Prior to that, he was president of the Association for Machine Translation in the Americas (AMTA) (2008-2012), and was General Chair of the AMTA 2010 and 2012 and MT Summit 2013 conferences.

COMET - a Neural Framework for State-of-the-Art MT Evaluation

Delivery of high-quality Machine Translation (MT) is only possible with reliable evaluation metrics to inform modelling and system development. The translation workflows we develop at Unbabel require highly-adapted MT systems which are regularly retrained to continuously deliver customer-specific, accurate translations. Unfortunately, with current state-of-the-art neural MT systems, traditional metrics such as BLEU and METEOR have been shown to no longer correlate well with human judgments, and in particular, they poorly distinguish between fine-grained accuracy distinctions of top performing MT models. This can result in misinformed MT development decisions that affect the quality of translations for our customers. To address this challenge, we recently developed COMET - a new neural-based framework for training automated MT evaluation models that are demonstrated to exhibit new state-of-the-art levels of correlation with human judgments. Our framework leverages recent breakthroughs in cross-lingual pretrained language modeling resulting in highly multilingual and adaptable MT evaluation models that exploit information from both the source input and a target-language reference translation in order to more accurately predict MT quality. We showcase our framework by training and evaluating COMET models for three different types of human judgments: Direct Assessments, Human-mediated Translation Edit Rate (HTER) and Multidimensional Quality Metrics (MQM). Our models achieve new state-of-the-art performance on the WMT 2019 and 2020 Metrics shared tasks and are sensitive to fine distinctions typical of high-performing MT systems. The COMET framework and our top-performing pretrained evaluation models are freely available open-source.

In this presentation we present an overview of the COMET framework and highlight its capabilities through assessments of the COMET models we have trained, their correlation with human judgments of translation quality, and their utility in practice for evaluating and contrasting MT models developed at Unbabel.

Friday, February 19, 2021

2:20 - 3:40 PM EST

Join the meeting on Zoom

Meeting ID 935 3287 1380 Passcode 546823

LTI Colloquium 2020-21
is generously sponsored by abridge