Natural Language Processing (NLP)
Laboratory

Lausanne, September 4, 2024

Dear Sir or Madam,

I am writing to you to in support of Kélyan Théo Hangard's application to internships at your company.

Kélyan Théo took my EPFL Masters course *CS-552: Modern Natural Language Processing* in Spring 2024. The goal of this course is to cover the foundations of modern methods for natural language processing, such as word embeddings, recurrent neural networks, transformers, and pretraining, and how they can be applied to important tasks in the field, such as machine translation and text classification. We also cover issues with these state-of-the-art approaches (such as robustness, interpretability, calibration), identify their failure modes in different NLP applications, and discuss analysis and mitigation techniques for these issues.

Students are expected to complete assignments, review papers, collect and evaluate data quality, implement state of the art deep learning models, and perform all of these tasks within the context of a large research and development project. Specifically, in the last six weeks of the semester, students must fine-tune a pretrained LLM (e.g., using DPO or RLHF) to respond to course questions from our curriculum at EPFL, and then either augment this trained system through quantization or retrieval agumentation (i.e., RAG). Over the course of the semester, they work with common libraries such as PyTorch and HuggingFace, and use large-scale GPU compute resources (e.g., a cluster of 100+ V100 GPUs) to implement their solutions. They also submit progress reports and a final report in the style of an academic research paper. This course is one of the most intensive in our masters program, but prepares students effectively for the real-world rigors of LLM implementation and evaluation in applied settings.

Kélyan Théo received a final grade of 6.0 / 6.0 in the class, a high score in a course of this nature. As a result, I **recommend** him for an internship position related to ML, NLP, and LLMs at your company.

Please feel free to reach out if I can provide any further information for his application.

Best Regards,

Antoine Bosselut[1]

[1] I am a tenure-track assistant professor in the School of Computer and Communication Sciences at the Swiss Federal Institute of Technology in Lausanne (EPFL). Prior to my current appointment, I was a postdoctoral researcher with Stanford University and a young investigator at the Allen Institute for AI (AI2). I received a Ph.D. in computer science from the University of Washington and a B.Eng. in electrical engineering from McGill University. Throughout my career, my research has focused on natural language processing and machine commonsense reasoning.

School of Computer and Communication Sciences

Prof. Antoine Bosselut
EPFL IC IINFCOM NLP
Station 14
CH–1015 Lausanne
Switzerland

+41 79 218 7442
antoine.bosselut@epfl.ch
atcbosselut.github.io