353 Serra Mall
Palo Alto, CA, US
☎ *+1 (607) 379 4739*
✉ *ozansener@cs.stanford.edu*
www.ozansener.net

September 7, 2017

Dear Learning& Adaptive Systems Lab,

I am writing to apply for the Post-Doc in Machine Learning position. I have recently completed my PhD studies at Cornell University in Machine Learning under the supervision of Ashutosh Saxena. Following my advisors move, I joined Stanford University as a visiting PhD student and continued to stay on as a postdoctoral fellow collaborating with Silvio Savarese. I am particularly interested in this postdoctoral position at ETH Zurich as I believe my interest and experience in deep learning, transfer learning, active learning and graphical models make me an ideal candidate.

Following the unprecedented success of deep learning using large-scale labelled datasets, I became interested in the question "How can we develop rigorous algorithms to improve the labelling efficiency of deep learning?" My research on label efficient deep learning initially focused on understanding the geometry of the feature space learned by deep networks. Consequently, we developed a metric learning algorithm [2] to impose regularity to this space and we proposed a transfer learning mechanism making deep learning models generalize in the covariate shift scenario (ie. when the train and the test distributions do not match). We further studied the geometry of the feature space of deep networks in the active learning setting. We defined the active learning problem as choosing a set of points to label such that when they are labelled; the loss over the labelled examples and unlabelled examples are close to each other. Hence, minimizing the loss over the labelled examples would minimize the expected loss over unlabelled ones. We gave a rigorous bound between the expected loss over unlabelled examples and empirical loss over labelled examples in a realistic deep learning setting [1]. Minimizing the bound we obtained, we proposed a theoretically sound active learning algorithm which sets a new state-of-the-art by a large gap.

In addition to the geometric understanding of data landscape, it is of the utmost importance to study the uncertainty in deep learning. Explicitly handling the uncertainty over the data as well as the model, would increase labelling efficiency significantly. The expertise and the knowledge accumulated at the Learning&Adaptive Systems Lab would be highly beneficial for me in developing a rigorous understanding of uncertainty in deep learning. I believe it is crucial to transfer the knowledge from graphical models to the field of (Bayesian) deep learning. I have already started to work in this direction to complement the geometric understanding I developed. We recently studied the problem of learning under privileged information using deep neural networks. By using Bayesian neural networks and ideas from information theory such as information bottleneck, we develop an algorithm which is able to effectively use privileged information. Our algorithm has better generalization guarantees, and resulted in a significant increase in labelling efficiency when compared with state-of-the-art algorithms for various computer vision and natural language processing problems. Although our paper detailing this work is under preparation, I would be more than willing to share the details of this result in our future correspondences. Combining the geometric and probabilistic understanding of deep learning, I believe we can increase the labelling efficiency of deep learning significantly, and develop rigorous, efficient and accurate one-shot learning, semi-supervised learning and active learning algorithms.

Thank you for your consideration. I look forward to hearing from you.

Sincerely,

[redacted]

[1] Ozan. Sener and Silvio. Savarese. Active Learning for Convolutional Neural Networks: A Core-Set Approach. *ArXiv e-prints*, August 2017.

[2] Ozan Sener, Hyun Oh Song, Ashutosh Saxena, and Silvio Savarese. Learning transferrable representations for unsupervised domain adaptation. In *Advances in Neural Information Processing Systems*, pages 2110–2118, 2016.