Discrete I-vector Representations for Modeling DNN Hidden Layer Posteriors

Najim Dehak[†1]

Abstract: The I-vector representation became the state of the art method for several audio classification tasks such as speaker and language recognition. This approach consists of modeling and capturing all the different variability in the Gaussian Mixture Model (GMM) mean components between several audio recordings. More recently a discrete versions of the I-vector representation has been proposed to model the variability between the GMM weights rather than the GMM means. These last techniques such as Non-negative Factor Analysis, Subspace Multinomial Model and Non Negative Matrix Factorization had to deal with the constraint that the GMM weights are always positive and they should sum to one. In this talk, we will show how these discrete I-vector representations can be also used to model deep neural network hidden layer posteriors for sequential data recognition task such as language and dialect recognition.

†1 Johns Hopkins University