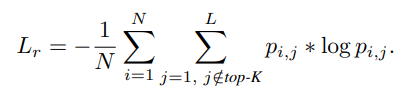
Residue loss



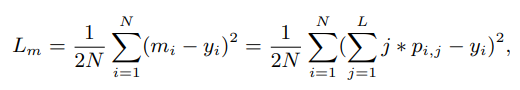
N: number of instances, in the case of traffic prediction, it is the number of predicted values for sensors at different times

L: the number of classes, in the case of traffic, it is the classes of speed = 0, 1, 2, …

Top-K: the K classes with highest probabilities; if K is flexible, classes with probabilities greater or equal to the probabilities of the ground truth are in top-K; if K is fixed, the k classes with highest probabilities are top-K

P: probabilities

Mean loss



N: same as above

Mi: the predicted value

Yi: ground truth

L: same as above

P: probabilities

Mean and Residue loss combined with cross entropy loss, with weights lambda1 and 2

