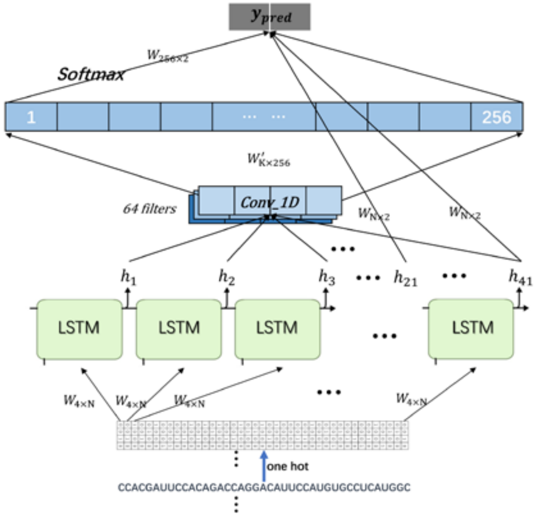
DeepM6ASeq-EL: Prediction of Human N6-Methyladenosine (m6A) Sites with LSTM and Ensemble Learning

Juntao CHEN, Jing LI, Quan ZOU

Frontiers of Computer Science, DOI: 10.1007/s11704-020-0180-0

Problems & Ideas

- Problems of improving the accuracy of m6A site prediction
  - Most of the existing methods are based on traditional machine learning methods, which lack the exploration of deep learning methods
- Ideas: Develop a m6A site predictor based on deep learning
  - compare three RNA sequence encoding methods (one-hot, distributed and word2vec) and discuss the effect of different lengths of encoding on model performance
  - combine the LSTM and CNN networks to develop a prediction model



Main Contributions

- **Comparisons in the experiment indicated that the optimal encoding length is the single nucleotide and the optimal encoding method is one-hot encoding.**

- **Develop an m6A site predictor named DeepM6ASeq-EL, which integrates an ensemble of five LSTM and CNN classifiers with the combined strategy of hard voting.**