Chapter 2

Deep neural networks for predicting DNA methylation

2.1 Motivation

2.2 Existing methods and limitations

2.3 DeepCpG model architecture

2.3.1 DNA module

$$a_{nfi} = \text{ReLU}(\sum_{l=1}^{L} \sum_{d=1}^{D} w_{fld} s_{n,i+l,d}) \tag{2.1}$$

$$p_{nfi} = \max_{|k|<P/2} (a_{nf,i+k}) \tag{2.2}$$

2.3.2 CpG module

$$\bar{x}_t = \text{ReLU}(W x_t + b) \tag{2.3}$$

$$r_t = \text{sigmoid}(W_{lr}\bar{x}_t + W_{hr}h_{t-1} + b_r) \tag{2.4}$$

$$u_t = \text{sigmoid}(W_{lu}\bar{x}_t + W_{hu}h_{t-1} + b_u) \tag{2.5}$$

$$\tilde{h}_t = \tanh\left(W_{l\tilde{h}}\bar{x}_t + W_{h\tilde{h}}(r_t \odot h_{t-1}) + b_{\tilde{h}}\right) \tag{2.6}$$

$$h_t = (1 - u_t) \odot h_{t-1} + u_t \odot \tilde{h}_t \tag{2.7}$$

2.3.3 Fusion module

$$\hat{y}_{nt} = \text{sigmoid}(x) = \frac{1}{1 + e^{-x}} \tag{2.8}$$

2.3.4 Model training

$$L(w) = \sum_{n=1}^{N}\sum_{t=1}^{T} o_{nt}\left[y_{nt}\log(\hat{y}_{nt}) + (1 - y_{nt})\log(1 - \hat{y}_{nt})\right] \tag{2.9}$$

2.4 Prediction performance evaluation

References

[1] Abramovich, Y. A., Aliprantis, C. D., and Burkinshaw, O. (1995). Another characterization of the invariant subspace problem. *Operator Theory in Function Spaces and Banach Lattices.* The A.C. Zaanen Anniversary Volume*, Operator Theory: Advances and Applications*, 75:15–31. Birkhäuser Verlag.

[2] Aupetit, B. (1991). *A Primer on Spectral Theory.* Springer-Verlag, New York.

[3] Conway, J. B. (1990). *A Course in Functional Analysis.* Springer-Verlag, New York, second edition.

[4] Ljubič, J. I. and Macaev, V. I. (1965). On operators with a separable spectrum. *Amer. Math. Soc. Transl. (2)*, 47:89–129.