

## 作业 5

本次作业数据集随附件提供。数据背景：基于随机库合成了一批 5UTR 序列，测定其平均核糖体载量 MRL。

- 1) 阅读参考资料给出的数据来源原始论文，进行 400 字左右总结，重点解释原文中模型如何搭建(3 分)
- 2) 利用本次作业的附件数据集，自选一种合适的序列 AI 模型(CNN, RNN, LSTM, Transformer 等)实现基于序列特征预测核糖体的任务。简要介绍模型 (2 分)，用 pytorch 代码实现并训练 (3 分)
- 3) 合理评估模型 (2 分)

参考资料：

原始论文:<https://www.nature.com/articles/s41587-019-0164-5>

数据处理流程:[https://github.com/pjsample/human\\_5utr\\_modeling](https://github.com/pjsample/human_5utr_modeling)