

第 3 章 第三次作业

本次作业需要进行代码实现，并写**简短**的文档介绍自己的实现方式和实验结果。请提交文档和代码文件（请放到一个压缩包中）至**人大云盘** (<https://pan.ruc.edu.cn:443/link/0B49F4BC05437CD67E1011F69263F711>，有效期限：2024-11-01 23:59，访问密码：8Pve)

文件命名为“多模态机器学习 24 秋学期-第三次作业-[学号]-[姓名]”，截止时间为 2024 年 11 月 1 日 23:59 (UTC+8)。

3.1 多模态特征学习方式：Fusion & Coordination

本次作业要求使用代码实现课上所讲的两多模态特征的学习方式：Fusion 和 Coordination (如图 3.1 所示)。需要注意的是，Fusion 方式测试时用两个模态来实现预测，而 Coordination 方式测试时只用**单一**模态。实验采用 CMU-MOSI 数据集，实现一个情感分析（回归）任务，其标注在 $[-3, +3]$ 范围内。我们在实例代码中给出了一个简单的 Fusion 的样例，请自选 Backbone 提取 feature（可选择 LSTM, Transformer 等架构），实现这两种特征学习方式。一些要求和建议如下：

- 修改 src/models 中两个模型，并适当调节其他部分代码以实现两种特征学习方式。对于实例程序，安装 torch 后，采用 `python main.py` 即可运行。
- 模型分类效果的好坏并不十分重要，本次作业的目的是锻炼大家对于两种学习方式的理解。无需花费大量时间提升性能，只需实现正确实现代码，并在文档中讲清楚即可。

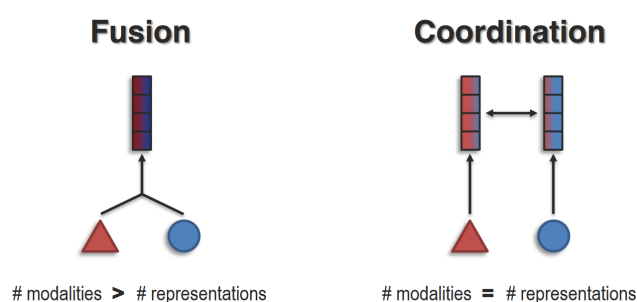


图 3.1: Examples of two multi-modal representation learning methods。