

# 🌟 优化版多模态回归模型架构文档

## ◆ 1. 分子特征提取模块（GIN + 虚节点）

输入：

- 原子特征：6维（原子类型、度数、电荷等）
- 键特征：3维（键类型、共轭、是否在环上）

网络结构：

- 5层 GIN（Graph Isomorphism Network）图卷积
- 引入 **虚节点（Virtual Node）**，增强全局结构建模能力
- 每层后接 BatchNorm + ReLU + Dropout (p=0.2)
- 最终输出分子图表示向量：**256维**

## ◆ 2. 蛋白质特征提取模块（ESM + ResNet1D + 可学习位置编码）

输入：

- 蛋白质序列，最大长度  $\leq 1024$

网络结构：

### 1. 预训练模型：ESM-1b

- 输出每个氨基酸的嵌入向量（1280维）

### 2. 可学习位置编码

- 加入每个位置的可学习向量，增强位置信息建模能力

### 3. 深层时序建模：ResNet1D（带残差连接）

- 4层残差块（Conv1D + BatchNorm + GELU + Shortcut）
- 维度变化：1280  $\rightarrow$  256  $\rightarrow$  512  $\rightarrow$  512  $\rightarrow$  512

#### 4. 特征聚合：Attention Pooling

- 用自注意力加权聚合时间维度信息
- 得到蛋白质表示向量：**256维**

### ◆ 3. 跨模态融合模块（Cross-Attention + 门控机制）

输入：

- 分子表示（256维）
- 蛋白质表示（256维）

融合结构：

#### 1. Cross-Attention 模块

- 蛋白质作为 Query，分子作为 Key/Value（和反向方向）
- 相互提取关键信息，捕捉局部对齐关系

#### 2. Transformer 编码器

- 2层标准 Transformer Encoder，8头注意力
- 前馈网络维度 1024（4×256）

#### 3. 门控融合机制（Gated Multimodal Unit）

- 对融合后的信息进行加权过滤
- 输出融合向量：**256维**

### ◆ 4. 回归预测模块（多任务）

主分支（回归预测）：

- 全连接：256 → 128 → 1
- 中间层：LayerNorm + GELU + Dropout(p=0.2)
- 输出：**标准化的 KIBA 分数**

辅助分支（多任务分类）【可选】：

- 判别 KIBA 区间（如高/中/低亲和）
- 帮助主任务提升泛化能力

# 输出处理：

- 推理阶段自动反标准化
- 支持原始 KIBA 值恢复

## ◆ 5. 训练策略（推荐）

项目	策略
优化器	AdamW
学习率调度	CosineAnnealingLR / ReduceLROnPlateau
正则化	Dropout, 权重衰减 (1e-4)
EarlyStopping	patience = 5, 监控验证 R²
梯度裁剪	max_norm = 5.0
Batch Size	32 或 64
训练轮数	最多 100 轮，结合 EarlyStopping

## ✦ 模型输出摘要

模块	输出维度	特征类型
GIN + 虚节点	256	分子结构嵌入
ESM + ResNet1D	256	蛋白质语义表示
跨模态融合	256	相互作用表示
回归预测（主分支）	1	KIBA 分数