

# 论文大纲

---

## 绪论

---

### 研究背景与意义

- 骨关节炎患者群体庞大
- 需求融合基因型与表现型的可解释模型
- 图神经网络广泛应用

### 现状研究综述

#### 骨关节炎

- 骨关节炎表型分析
- 骨关节炎GWAS分析

#### 图神经网络

- 图神经网络被广泛用于处理图数据
- 图神经网络可以用于处理SNP网络

#### 图结构特征估计

- 对于无特征数据需要通过不同方法完成图特征估计

### 本文工作

- 获取UKBank中关于骨关节炎的表型与基因型数据
- 设计统计学习算法根据基因型数据建图
- 构建神经网络用于骨关节炎患病风险的预测
- 通过模型分析可能与骨关节炎相关的通路与其分型

### 数据与预处理

---

- UKB介绍
- 本文研究性状
- 基因型数据筛选
- 样本筛选
- 特征筛选

### 预测模型构建

---

#### 图神经网络设计

- 选择谱图神经网络
- 谱图神经网络中以Cheby卷积层为核心，构建网络

## 图的生成

- 选择全连接法与相关系数法生成图的邻接矩阵
- 以上方法不能很好反应图结构特征，需要进一步改进

## 结合表型

- 选择报道中与骨关节炎相关的性状进行分析
- 通过中间融合法结合基因型数据与表型数据

## 图结构估计

---

### 问题描述

- 目标：通过观测值估计图节点之间关联
- 流程：
  - 图构建：knn法与距离法
  - 图结构建模：神经网络法；损失函数法；直接法

### 期望最大化算法

基于网络输出  $X'$ ，通过knn法构建观测值  $O$ ，计算邻接矩阵相对于观测值的后验概率分布

### 变分期望最大化算法

解决期望最大化算法中的问题

## 分析与讨论

---

### 传统风险评估模型与普通机器学习性能比较

### 特征筛选方法对性能的影响

### 邻接矩阵方案对模型效果的影响

### 图结构估计器性能表现

- 在Mnist数据集上的表现
- 本项目数据集上的表现

### 估计器产生图结构分析

### 融合表型模型综合风险预测结果

## 总结

---

