

# 人工智能与生物医学大数据

## A 医学图像

### 一、题目：医学图像的分类

二、背景：甲状腺是人体内重要的内分泌腺，其功能异常会导致许多疾病。甲状腺超声图像是目前诊断甲状腺疾病的主要影像学方法之一。为了提高甲状腺疾病的诊断准确性和效率，本次比赛旨在开发一种基于甲状腺超声图像的自动分类模型，以辅助医生进行甲状腺疾病的诊断。

### 三、任务

要求通过本竞赛所提供的数据集，设计算法对测试集中的甲状腺图像进行二分类，即判断其为良性或恶性。竞赛数据集 `thyroid_dataset`，包含三个文件夹，分别为 `malignant`、`benign` 和 `test`，`malignant` 文件夹下的图像为恶性，`benign` 文件夹下的图像为良性，`test` 文件夹下的图像为测试图像。数据集总共有 946 张，其中比赛前发放 850 张（提供良恶性标注），另有 96 张测试图像未提供良恶性标注，用于比赛模型的评判。比赛评价标准采用 F1-Score。F1-Score 越高，分类器的性能越好。

$$F1 - Score = \frac{2 * (precision * recall)}{(precision + recall)}$$

其中，`precision` 表示模型的精度，`recall` 表示模型的召回率。

### 四、评分标准：

	评分项目	分数
产品设计	设计报告：问题定义与分析，数据分析与预处理，模型设计与选择，实验测试，全部代码	50
	实际运行测试：将测试集中 96 个样本的预测结果保存在“ <code>results.txt</code> ”中，提交的预测结果文件需要包含测试集中的每个图像的预测标签。预测标签为 0 表示该图像为良性，预测标签为 1 表示该图像为恶性。提交的预测结果文件需要包含一个表头，ID 为文件名， <code>Category</code> 为预测结果。 格式如下： ID,Category 1,0 2,1 3,0 ...	50
	总分	100

**郑重声明：**此数据集仅供此次竞赛使用，参赛者不得将参赛作品以及此数据集转让给第三方使用，不得将参赛作品用于任何商业用途及其它使用，比赛完毕自行删除本次赛事相关数据。若违反规定，由此造成的一切后果自负。

## B 生物信息

一、**题目：**肾癌免疫亚型的构建及免疫相关分析

二、**背景：**肾癌是泌尿系统常见的恶性肿瘤之一，包括起源于泌尿小管不同部位的各种肾细胞癌亚型，其中，肾透明细胞癌（KIRC）是最常见的组织学亚型。晚期肾癌对化疗药物不敏感，主要以免疫治疗和靶向治疗为主。分子分型是准确认识复杂疾病进而构建精准医学体系的重要基石，已成为精准医疗的一个重要研究方向，但构建具有精确性、鲁棒性及进一步可推广性的分子分型体系是众多复杂疾病面临的极大挑战。同时，分子分型的临床价值不仅在于提示具有差异疾病表型、分子基础的患者群体，针对不同特征群体筛选有效的临床干预策略也是分子分型体系构建的重要目标。

三、**任务：**

根据所提供的免疫相关基因集，可构建肾透明细胞癌免疫相关亚型，并针对不同亚型进行免疫相关分析（如免疫浸润及免疫检查点分析等），以了解不同亚型的免疫特征，为后续治疗提供参考。本竞赛提供了肾透明细胞癌的转录组数据和相对应的临床信息以及免疫相关基因，请利用所提供的数据，完成以下任务：  
1.基于相关算法进行免疫相关分子亚型的聚类研究，选择最佳分类结果；  
2.结合临床信息进行不同亚型间的预后分析；  
3.根据构建的免疫亚型，开展免疫浸润和免疫检查点等相关分析。

四、**评分标准：**

	项目	分数
产 品 设 计	竞赛报告：包括设计思路、模型算法的具体介绍和实现，以及对实验结果的分析等。	40
	程序及实验结果：算法的完整程序及最后的实验结果。	60
	总分	100