

同行专家业内评价意见书编号：20250860041

附件1

浙江工程师学院（浙江大学工程师学院） 同行专家业内评价意见书

姓名：白晟佐

学号：22260232

申报工程师职称专业类别（领域）：生物与医药

浙江工程师学院（浙江大学工程师学院）制

2025年03月20日

填表说明

一、本报告中相关的技术或数据如涉及知识产权保护、军工项目保密等内容，请作脱密处理。

二、请用宋体小四字号撰写本报告，可另行附页或增加页数，A4纸双面打印。

三、表中所涉及的签名都必须用蓝、黑色墨水笔，亲笔签名或签字章，不可以打印代替。

四、同行专家业内评价意见书编号由工程师学院填写，编号规则为：年份4位+申报工程师职称专业类别(领域)4位+流水号3位，共11位。

一、个人申报

(一) 基本情况【围绕《浙江工程师学院(浙江大学工程师学院)工程类专业学位研究生工程师职称评审参考指标》，结合该专业类别(领域)工程师职称评审相关标准，举例说明】

1. 对本专业基础理论知识和专业技术知识掌握情况(不少于200字)

在生物与医药交叉领域的研究与实践中，我系统构建了肿瘤免疫学与人工智能的双重知识体系，形成了从生物学机制解析到计算模型开发的全链条技术能力。在免疫学理论层面，我聚焦肿瘤新抗原预测的核心生物学问题，深入研习肿瘤微环境中免疫编辑机制、抗原呈递通路(MHC-I类分子加工递呈)的动态调控原理，以及T细胞受体(TCR)与抗原肽-HLA复合物的相互作用模式，为新抗原免疫原性预测模型的构建奠定分子机制基础。在人工智能技术领域，我精通深度学习算法在生物信息学中的创新应用，熟练掌握卷积神经网络(CNN)对蛋白质空间特征的提取、循环神经网络(RNN)对肽段时序信息的建模，以及基于Transformer架构的多模态数据融合策略。通过PyTorch框架，我实现了从原始生物数据清洗(如质谱数据去噪、HLA分型数据标准化)、特征工程(肽序列的理化属性编码、三维结构特征嵌入)到端到端模型训练与部署的全流程技术闭环，并利用Biopython等工具包实现了生物数据与机器学习管道的无缝对接。

在浙江大学智能创新药物研究院的科研实践中，我主导了蛋白酶体切割位点预测这一肿瘤新抗原筛选的关键技术攻关。基于蛋白酶体对长肽前体的特异性切割是抗原递呈步骤的科学认知，我系统性整合了IEDB等数据库中40万条抗原肽质谱数据，设计了基于双向长短期记忆网络(BiLSTM)的上下文特征捕捉模块，构建模型预测抗原肽能够经过胞内处理被成功递呈的概率，测试集AUC达0.886。最终成功按计划完成模型的训练以及部署，并成功实现了对抗原递呈过程的拟合，为后续肿瘤新抗原预测算法构建打下了坚实基础。

2. 工程实践的经历(不少于200字)

在浙江大学智能创新药物研究院的科研实践中，我主导了蛋白酶体切割位点预测这一肿瘤新抗原筛选的关键技术攻关。基于蛋白酶体对长肽前体的特异性切割是抗原递呈步骤的科学认知，我系统性整合了IEDB等数据库中40万条抗原肽质谱数据，设计了基于双向长短期记忆网络(BiLSTM)的上下文特征捕捉模块，构建模型预测抗原肽能够经过胞内处理被成功递呈的概率，测试集AUC达0.886。最终成功按计划完成模型的训练以及部署，并成功实现了对抗原递呈过程的拟合，为后续肿瘤新抗原预测算法构建打下了坚实基础。

3. 在实际工作中综合运用所学知识解决复杂工程问题的案例(不少于1000字)

在本人的研究生生涯中，致力于探索肿瘤新抗原的预测问题，并围绕抗原胞内处理和免疫原性预测两个关键环节，构建了高精度的预测模型。这项研究不仅在技术层面上取得了突破，还对肿瘤免疫治疗的精准性提升做出了重要贡献。在此过程中，本人综合运用了深度学习、特征工程、数据分析和模型优化等多方面的知识，以解决复杂的生物医药工程问题。

一、问题背景

肿瘤新抗原是个性化癌症免疫治疗的重要靶点，然而，其筛选过程复杂且成本高昂。当前主流的预测方法在精度和适用性上仍存在局限性，因此，提高新抗原预测的准确性和通用性，对推动个性化免疫治疗的临床应用具有重要意义。

为了系统性地建模抗原从胞内加工到免疫识别的完整过程，本人构建了两个核心模型：抗原胞内处理预测模型ProSense以及免疫原性预测模型Multi-IMG。

二、模型开发与技术创新

1. 抗原胞内处理预测模型 ProSense

在抗原处理预测方面，本人创新性地提出了一种基于序列上下文信息的建模方案。具体来说，本人采用了特殊标记的方式，将抗原肽序列与其上下游氨基酸信息进行整合，从而充分利用局部序列环境。数据处理优化方面，本人系统探索了不同的数据预处理策略，确定了最优的上下游序列长度(15个氨基酸)和数据编码方式；在特征工程选择实验中，本人通过尝试不同的表征方式发现，简单的nn.Embedding特征学习方法在此任务上优于复杂的预训练模型表征，比如大型蛋白预训练模型ESM2得到的特征；在模型架构探索中，本人通过实验比较不同网络结构，发现双层BiLSTM网络最适用于此任务，结果优于多层Transformer或者是CNN等

其他模型，最终ProSense模型在测试集上取得了0.873的AUC和0.089的PRAUC的优异成绩。同时，本人将ProSense与现有的抗原-

MHC结合预测模型进行决策融合，也显著提高了整体预测性能。该模型的开发显著提升了抗原胞内处理预测的准确性，为后续免疫原性筛选提供了更为精准的输入数据。

2. 免疫原性预测模型 Multi-IMG

在免疫原性预测方面，本人重点针对影响免疫识别的关键特征进行了深入研究，并设计了一种多阶段学习框架。本人首先对24个免疫原性相关特征进行了系统性分析，并通过特征选择实验确定了最优特征组合，发现Prime模型打分、部分氨基酸组成特征和一系列生理过程预测得分对免疫原性预测贡献最大。接下来，本人设计了一个多阶段学习框架：在预训练阶段，利用多个物种来源的70000多条免疫原性数据进行预训练，增强模型对序列的泛化能力；之后在序列数据的基础上添加免疫原性相关特征，并在免疫原性模型训练中引入蒸馏回归学习，利用预训练模型特征重构免疫原性相关特征辅助任务提升模型的特征表示能力。最终模型在独立测试集和上进行性能评估：在GBM和TESLA两个独立测试集上，分别达到了0.7853/0.5024的AUC/PRAUC和0.7993/0.2966的AUC/PRAUC，显著优于现有方法，并且Multi-IMG在Top-k召回率方面表现优异，证明了模型在新抗原筛选中的实用价值。

三、 实际应用与工程价值

本研究的成果在肿瘤免疫治疗领域具有广泛的应用潜力，为精准新抗原筛选提供了有力工具，具有如下意义：1) 降低新抗原筛选成本，提高预测精度，减少实验验证所需的样本数，从而降低成本；2) 推动个性化免疫治疗，帮助临床医生快速筛选出更具免疫原性的肿瘤新抗原，提高治疗方案的有效性；3) 深化对免疫应答机制的理解，通过深入分析抗原处理和免疫原性特征，有助于揭示免疫系统如何识别和响应新抗原。

四、 总结与个人贡献

在该项目中，本人综合运用了深度学习、特征工程、数据挖掘和优化方法，以解决复杂的生物医药工程问题。本人的主要贡献包括：1) 提出了一种基于上下文整合的新抗原抗原胞内处理预测方法，提高了预测准确性；2) 设计了一种多阶段免疫原性学习框架，并引入蒸馏策略提升模型性能；3) 通过大规模数据分析和特征筛选，优化了新抗原预测流程，提升了实际应用价值；4) 结合已有模型进行决策融合，提高了新抗原预测的整体性能。

这项研究不仅是对生物信息学与人工智能交叉应用的一次探索，也为肿瘤免疫治疗提供了新的技术支持。希望未来能够将这些研究成果进一步推广，助力精准医学的发展。这一案例展示了本人在实际工作中综合运用人工智能、生物信息学和医学知识，解决复杂工程问题的能力，同时也证明了本人在生物医药领域中级工程师职位上的胜任能力。

（二）取得的业绩（代表作）【限填3项，须提交证明原件（包括发表的论文、出版的著作、专利证书、获奖证书、科技项目立项文件或合同、企业证明等）供核实，并提供复印件一份】

1. 公开成果代表作【论文发表、专利成果、软件著作权、标准规范与行业工法制定、著作编写、科技成果获奖、学位论文等】

成果名称	成果类别 [含论文、授权专利（含发明专利申请）、软件著作权、标准、工法、著作、获奖、学位论文等]	发表时间/ 授权或申请时间等	刊物名称/ 专利授权或申请号等	本人排名/ 总人数	备注

2. 其他代表作【主持或参与的课题研究项目、科技成果应用转化推广、企业技术难题解决方案、自主研发设计的产品或样机、技术报告、设计图纸、软课题研究报告、可行性研究报告、规划设计方案、施工或调试报告、工程实验、技术培训教材、推动行业发展中发挥的作用及取得的经济社会效益等】

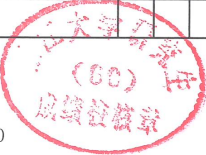
(三) 在校期间课程、专业实践训练及学位论文相关情况	
课程成绩情况	按课程学分核算的平均成绩： 86 分
专业实践训练时间及考核情况(具有三年及以上工作经历的不作要求)	累计时间： 1 年（要求1年及以上） 考核成绩： 85 分
本人承诺	
个人声明：本人上述所填资料均为真实有效，如有虚假，愿承担一切责任，特此声明！	
申报人签名：白晟佐	

浙 江 大 学 研 究 生 院
攻读硕士学位研究生成绩表

学号：22260232		姓名：白晟佐		性别：男		学院：工程师学院		专业：生物与医药				学制：2.5年			
毕业时最低应获：24.0学分				已获得：27.0学分				入学年月：2022-09				毕业年月：			
学位证书号：						毕业证书号：						授予学位：			
学习时间		课程名称		备注	学分	成绩	课程性质	学习时间		课程名称		备注	学分	成绩	课程性质
2022-2023学年秋季学期		人工智能算法与系统			2.0	100	跨专业课	2022-2023学年春季学期		数学建模			2.0	70	专业选修课
2022-2023学年秋季学期		工程技术创新前沿			1.5	83	专业学位课	2022-2023学年春季学期		自然辩证法概论			1.0	80	公共学位课
2022-2023学年秋季学期		新时代中国特色社会主义思想理论与实践			2.0	92	公共学位课	2022-2023学年夏季学期		工程师创新创业思维			2.0	93	专业选修课
2022-2023学年秋冬学期		工程伦理			2.0	92	公共学位课	2022-2023学年夏季学期		药品创制工程实例			2.0	90	专业学位课
2022-2023学年秋冬学期		高阶工程认知实践			3.0	90	专业学位课	2022-2023学年夏季学期		研究生英语基础技能			1.0	免修	公共学位课
2022-2023学年冬季学期		新药发现理论与实践			2.0	88	专业学位课	2022-2023学年夏季学期		研究生英语			2.0	免修	公共学位课
2022-2023学年冬季学期		产业技术发展前沿			1.5	90	专业学位课			硕士生读书报告			2.0	通过	
2022-2023学年秋冬学期		研究生论文写作指导			1.0	80	专业学位课								

说明：1. 研究生课程按三种方法计分：百分制，两级制（通过、不通过），五级制（优、良、中、及格、不及格）。

2. 备注中“*”表示重修课程。

学院成绩校核章：

成绩校核人：张梦依

打印日期：2025-03-20