

国家自然科学基金资助项目批准通知

梁艳春 先生/女士：

根据《国家自然科学基金条例》和专家评审意见，国家自然科学基金委员会（以下简称自然科学基金委）决定批准资助您的申请项目。项目批准号：61972174，项目名称：基于深度学习的个性化基因组与药物相互作用研究，直接费用：59.00万元，项目起止年月：2020年01月至2023年12月，有关项目的评审意见及修改意见附后。

请尽早登录科学基金网络信息系统（<https://isisn.nsfc.gov.cn>），获取《国家自然科学基金资助项目计划书》（以下简称计划书）并按要求填写。对于有修改意见的项目，请按修改意见及时调整计划书相关内容；如对修改意见有异议，须在电子版计划书报送截止日期前向相关科学处提出。

电子版计划书通过科学基金网络信息系统（<https://isisn.nsfc.gov.cn>）上传，依托单位审核后提交至自然科学基金委进行审核。审核未通过者，返回修改后再行提交；审核通过者，打印纸质版计划书（一式两份，双面打印），依托单位审核并加盖单位公章后报送至自然科学基金委项目材料接收工作组。电子版和纸质版计划书内容应当保证一致。向自然科学基金委提交和报送计划书截止时间节点如下：

- 1、提交电子版计划书截止时间为**2019年9月11日16点**（视为计划书正式提交时间）；
- 2、提交电子修改版计划书截止时间为**2019年9月18日16点**；
- 3、报送纸质版计划书截止时间为**2019年9月26日16点**。

请按照以上规定及时提交电子版计划书，并报送纸质版计划书，未说明理由且逾期不报计划书者，视为自动放弃接受资助。

附件：项目评审意见及修改意见表

国家自然科学基金委员会
2019年8月16日

附件：项目评审意见及修改意见表

项目批准号	61972174	项目负责人	梁艳春	申请代码1	F020506
项目名称	基于深度学习的个性化基因组与药物相互作用研究				
资助类别	面上项目		亚类说明		
附注说明					
依托单位	吉林大学				
直接费用	59.00 万元		起止年月	2020年01月 至 2023年12月	
通讯评审意见： <1>具体评价意见： 一、请针对创新点详细评述申请项目的创新性、科学价值以及对相关领域的潜在影响。 本项目拟集成全基因组重测序数据，药物数据和电子病历数据，通过深度学习整合先验知识，从计算角度提升不同基因型对不同药物，不同反应的预测精度，弥补精准医学中用药缺陷。本项目研究内容具体，通过总结现有的药物反应算法，工具及公共数据，分析药物反应预测中的挑战性问题，提出了个体基因组中多组学和先验知识的融合，基因组与环境互作关系的多数据源融合的深度学习方法等科学问题。项目创新性强，研发的算法和工具可以为复杂疾病的分子分型，存活率预测提供支持，为病人寻找更有效的精准治疗药物。 二、请结合申请项目的研究方案与申请人的研究基础评述项目的可行性。 项目研究方案与技术路线合理，在异种数据预处理与融合，通用深度学习模型构建与预测，基于系统生物学的多信息融合与注释等方面都制定了可行的技术方案。在研究基础方面，项目组在数据挖掘，机器学习，统计方法和生物信息学方面，取得了一些研究成果。在本项目的前期工作中，已经收集了实验相关的数据，搭建了计算框架模型，为项目的顺利实施打下良好基础。建议优先资助。 三、其他建议 无。 <2>具体评价意见： 一、请针对创新点详细评述申请项目的创新性、科学价值以及对相关领域的潜在影响。 本课题通过融合多组学数据、个体先验知识、药物特征等信息，整合不同来源的异种数据，构建基于数据融合的异构深度学习计算模型，预测个体对不同药物的不同反应。该项目创新性较好，可行性高。 二、请结合申请项目的研究方案与申请人的研究基础评述项目的可行性。 然而，申请人应当与当前针对个体药物反应预测的方法进行比较。此外该标书没有具体阐明自主产生的样本验证模型正确性的过程，也没有给出具体的生物实验研究方案。申请人在相关领域具有一定的工作基础，经验丰富。 三、其他建议 综上，建议资助。 <3>具体评价意见： 一、请针对创新点详细评述申请项目的创新性、科学价值以及对相关领域的潜在影响。 利用深度学习算法构建复杂疾病中基因与药物的相互作用关系，寻找有效的针对病人个体的精确药物是计算生物学和系统生物学中的重要研究课题。该申请具有很强的创新性，能为医学的相关研究提供很好的理论支持。 二、请结合申请项目的研究方案与申请人的研究基础评述项目的可行性。 研究的内容充实，研究方法和研究路线具有可行性，申请人在本课题相关方面已做出过很好的					

研究工作，具有良好的研究基础。

三、其他建议
建议优先资助。

<4>具体评价意见：

一、请针对创新点详细评述申请项目的创新性、科学价值以及对相关领域的潜在影响。
该项目在多源异种多组学数据数据处理、基于异种数据的模型训练方法，以及全基因组规模的精准医疗药物预测模型等方面具有很好的创新性。主要表现在：提出了采用基于流形和迁移学习的先验知识与多源异种多组学数据的融合与表示方法，利用生成模型归一化数据，使异种数据可映射到统一的低维度流形中，通过迁移学习的方式融合不同来源的训练数据，可保留异种数据对数据模型的整体信息贡献，有利于实现全尺度的数据融合。提出了以胶囊网络为基础研究针对异种数据的模型训练方法，提出了解释胶囊网络及其它深度学习策略在个性化基因组与药物相互作用的预测模型，有利于获得与药物相互作用有关的基因组的新知识，在深度学习模型设计 and 应用角度同时进行了具有开创性的探索。通过生成模型对总体的异种高维数据进行低维流形映射，利用深度学习模型中的逐层抽象表征学习，有利于实现在全基因组规模上的精准医疗药物预测。从系统生物学的高度，针对不同基因型对不同药物的相互作用结果进行辨识、阐释和验证，将序列上的单核苷酸多态性信息与现有其它生物学信息结合在一起，从多层次多角度对获得的结果进行分析，有利于实现系统生物学尺度的生物学注释。
该项目充分利用大量的全基因组重测序数据、药物数据和公开的电子病历数据，通过深度学习工具整合先验知识，从计算的角度提升不同基因型对不同药物、不同反应的预测精度，有助于从系统论的层面理解复杂生物体局部与整体的关系，从进化的视角分析导致生物不同表型的内在动力和基因型层面的原因。项目成果能够弥补当前精准医学研究中对不同疾病分型具体用药的缺陷，有助于寻找更有效的针对病人个体的精确药物，能够为人类复杂疾病的研究提供理论参考和分析工具。

二、请结合申请项目的研究方案与申请人的研究基础评述项目的可行性。
该项目用胶囊网络、注意力机制等方法构建异构的深度学习计算框架对药物反应进行预测。通过构建蛋白相互作用网络建立药物反应关系映射的系统生物学解释，并研究其在人类复杂疾病药物治疗中的潜在应用。项目申请人近年来一直在计算机科学和生物信息学领域从事交叉研究，在生物网络以及Pathway预测、复合基因预测和操纵子预测、基因功能注释等生物信息相关领域取得了一系列高水平研究成果。这些研究工作的积累为该项目奠定了良好的前期工作基础。项目研究方案可行，研究方法和路线能够保障完成项目的研究内容，实现项目的研究目标。

三、其他建议
建议优先资助。

<5>具体评价意见：

一、请针对创新点详细评述申请项目的创新性、科学价值以及对相关领域的潜在影响。
多组学数据为刻画病人-疾病-药物作用关系提供了相互补充的多个视角。如何系统的建立计算模型整合多平台、多组学异构数据来进行疾病相关的机制研究与精准药物治疗仍然是一个开放性问题。本项目申请提出的基于深度学习的异构数据分析方法在多组学癌症研究等领域具有一定应用前景。此外，提出的基于胶囊网络的预测模型具有一定创新性和应用价值。

二、请结合申请项目的研究方案与申请人的研究基础评述项目的可行性。
3.1.1和3.1.2异构数据融合与预测模型设计方面，申请人有着详细的方法论述和现有工作基础，可行性有保障。

3.1.3深度学习可解释性部分：
不同于计算机视觉等领域，生物数据具有特征维度极高而样本量极少的特点。基于注意力机制的模型通常并不能得到稀疏解，且在样本量极少 的情况下很有可能拟合噪声。因此该方面的可行性需要进一步验证。
3.1.4部分如何融合异构的系统生物学信息，需要进一步详细论述，提供更多的可行性支持。

三、其他建议
修改意见：

信息科学部

2019年8月16日