AI Assignment Library   Schools, Colleges, and Departments

8-30-2023

Generative AI-Assisted Pathway Analysis and Interpretation of RNA-Seq Experiment Data

Junguk Hur
University of North Dakota, junguk.hur@med.und.edu

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: https://commons.und.edu/ai-assignment-library

Part of the Bioinformatics Commons

Recommended Citation

BIMD514: Foundations of Bioinformatics
(3-credit graduate-level course)
Dept. of Biomedical Sciences, School of Medicine and Health Sciences

Course overview: The course targets biomedical or biology graduate students without previous computational training. Students will learn fundamental concepts and methods in bioinformatics, a field at the intersection of biology and computing. The course surveys a wide range of topics, including Linux command line environment, bioinformatics web resources, R data handling, next-generation sequencing analysis (RNA-Seq), network/pathway analysis, and machine learning application. The course will familiarize students with the computational tools to process and analyze large biological sequencing data.

Potential Homework Assignment

Title: Generative AI-Assisted Pathway Analysis and Interpretation of RNA-Seq Experiment Data

Objective: Utilize an artificial intelligence (AI)-powered conversational platform, such as ChatGPT by OpenAI or Claude by Anthropic, to assist in the analysis of RNA-Seq high-throughput gene expression profiling experiment at the pathway level, thereby gaining an understanding of the potential and limitations of AI in bioinformatics.

Task:
Data Preparation:
- Download publicly available RNA-Seq raw data, the details of which will be disclosed later.
- Process the data using a standard RNA-Seq pipeline to identify differentially expressed genes (DEGs).
- Conduct a pathway enrichment analysis using tools such as GSEA, DAVID, or richR to identify the top 5 most significantly affected pathways in your dataset. You may use the Kyoto Encyclopedia of Genes and Genomes (KEGG) database or the Reactome pathway database for the pathway annotation.

Generative AI Interaction:
- Write a brief description of each of the top 5 pathways based on primary literature or known databases with respect to the biological condition of the dataset. Properly cite references.
- Interact with the AI tools to ask about the significance or functions of these pathways in relation to the biological condition of the dataset. Record the AI's response.
- Compare the AI's interpretation with your initial understanding and primary literature. Note any discrepancies, insights, or novel interpretations and briefly summarize them (maximally two paragraphs)

Reflection: Reflect on the experience of using a generative AI model to assist in RNA-Seq pathway interpretation. Answer the following questions:
- Did the AI provide information that was consistent with primary literature?

- Were there any novel insights or interpretations provided by the AI?
- What are the advantages of integrating AI into the RNA-Seq analysis and interpretation process?
- What are the potential limitations or pitfalls of relying on AI for this task?

Submission:

- The completed assignment must be submitted by November 21st at 11 AM.
- Ensure the following items are included:
    - Detailed description of the pipeline used for the DEG and pathway analysis.
    - The total number of DEGs.
    - A table containing the top 5 most significant pathways, including the official pathway ID, title, the number of DEGs in this pathway, and statistical significance represented by adjusted p-value.
    - Your description and interpretation of the top 5 pathways.
    - Responses from AI.
    - Your answers to the reflection questions.

Scoring/Evaluation Criteria for the Revised Assignment:

Criteria	Points	Description/Notes
1. Data Preparation	**25**	
Processing and Identification of DEGs	15	Correctness and clarity of RNA-Seq pipeline used.
Pathway Enrichment Analysis	10	Quality and relevance of the top 5 pathways chosen.
2. Generative AI Interaction	**40**	
Description of Pathway Function	10	Clarity, accuracy, and depth of initial pathway descriptions.
AI Interaction Documentation	20	Clear presentation of questions posed to AI and its responses for each pathway.
Comparison and Analysis	10	Depth of comparison between AI's responses and primary literature.
3. Reflection	**30**	
Consistency with Primary Literature	7.5	Depth and clarity in discussing AI's alignment with primary literature.
Novel Insights or Interpretations	7.5	Discussion on novel or unexpected insights provided by the AI.
Advantages of AI Integration	7.5	Insightfulness in recognizing potential benefits of AI in pathway analysis.
Limitations of AI	7.5	Critical understanding of potential pitfalls/limitations of AI reliance.
4. Overall Report Presentation	**5**	
Organization and Clarity	5	Logical structure, clear headings, and coherence in presenting analysis/findings.

Disclaimer: This assignment has been generated with the support of ChatGPT 4.0.