Hi, Sheraz Ahmad. How are You. Hope you're doing great.

As we are targeting Q1 SCI Ranked journal, So kindly ensure our research project is original, novel and contribute significantly to the field.   A Noval Deep learning predictive model (Python based) which propose a new model based on the shortcomings analyzed in literature review or we can improve existing model on another similar type of dataset. We can work on project based any research area which you have good command of knowledge and it has a significant gap to fill. Basically, this will be a research paper with code which is likely to publish on GitHub later So, you can also introduce me similar research area which involves prediction task on drug. It can be Drug Target Interaction prediction (DTIs), Drug Response prediction, Drug Target binding affinity prediction (DTA), Protein-ligand interaction Prediction. etc. basically python based predictive model.

Deep learning models can be complex, so ensure methodology is clear, reproducible, and well-justified. Because Journals are often interested in new approaches to prediction models. Make sure to review recent publications to confirm our model introduces methodological innovations and/or improved accuracy over existing models/techniques.

During this project i will provide you useful comments from my senior supervisors to make our project successful and share your efforts and workload. And if you face any restriction in getting access to any tool, and resources related to this project, let me know I will try my best to get you access to that resource through my institution.

To start with, I would like you to write a research proposal, which includes following.

1. **Title**

**Title**: Propose a clear and concise title that reflects the research problem and scope of the project.

**2. Introduction**

* **Problem Statement and Motivation**
  + Describe the research domain (e.g., drug target interaction, protein-ligand interaction).
  + Identify limitations of existing methods (e.g., poor interpretability, lack of biological grounding).
  + Begin by introducing the research problem or question that the project aims to address.
  + Explain the significance of this problem in the broader context of the field (e.g., bioinformatics, computational chemistry).
  + Describe the motivation for choosing this topic, including its real-world applications.
  + State the gap or challenge in existing research that your project will address (e.g., the need for more accurate predictions, efficient computational models).
* **Research Gap**: What’s missing in current solutions?

**Objectives**: List 3–5 goals. Outline the key objectives of the research project. What do you hope to achieve? (e.g., *"Design a hybrid CNN/MLP model for Prediction……"*).

Example objectives:

* Develop a deep learning model for a specific prediction task (e.g., drug-target interactions, protein folding).
* Evaluate the performance of the proposed model using established datasets.
* Compare the new model with existing methods to assess improvements in accuracy and efficiency.
* Analyze the interpretability of the model's results.

**3. Related Work**

* Summarize 3–5 key papers (e.g., DeepDTA, TransformerCPI) and their shortcomings.
* Highlight how your approach will differ.

**4. Literature Review**:

* Provide a brief overview of existing work related to the research topic.
* Identify and summarize key studies, models, or approaches that are relevant to your project.
* Highlight the strengths and weaknesses of current methodologies and explain why a new or improved approach is needed.
* Reference recent advancements in the field to justify the use of the selected methods or technologies (e.g., deep learning, neural networks).

**5. Methodology**

* **Model Architecture design of our Proposed Model**:
  + Describe each component.
  + Use diagrams/flowcharts (e.g., model pipeline).
  + Describe the input data format and how it will be processed by the model (e.g., raw sequence data, feature extraction).
* **Innovations**: Emphasize novelty.
* **Mathematical Formulations**: Include key equations (e.g., loss functions, attention mechanisms). Specify any techniques used for model training, including loss functions, optimization methods, and regularization techniques.
* **Baselines**: List models you’ll compare against (e.g., DeepDTA, GraphDTA).

**Model Evaluation**:

* Outline the performance metrics you will use to evaluate the model's success (e.g., accuracy, precision, recall, mean squared error).
* Discuss how you will compare your model's performance against existing baseline models or methods.
* Mention any cross-validation techniques or tests for robustness (e.g., k-fold cross-validation).

**6. Datasets Preparation and Evaluation**

**Data Collection and Preprocessing**:

* Describe the datasets you will use (e.g., biological data, chemical properties, genomic sequences).
* Outline the data collection process, including any sources or databases you will reference (e.g., PubChem, UniProt).
* Explain any preprocessing steps that will be taken to clean, normalize, or format the data.
* **Metrics**: Define evaluation criteria (e.g., CI, MSE, AUPR for affinity prediction; accuracy for BRs).

**7. Implementation Plan**

* **Tools**: Frameworks (PyTorch/TensorFlow), libraries (RDKit for drugs), and hardware (GPU requirements).

**8. Resources and Tools**:

* List any software tools, libraries, and frameworks that will be used in the project (e.g., TensorFlow, Keras, PyTorch, scikit-learn).
* Mention any additional resources that might be needed, such as access to GPUs, databases, or cloud computing services.
* Specify programming languages (e.g., Python, R) and any version control systems (e.g., GitHub) that will be used.
* **Milestones**

|  |  |  |  |
| --- | --- | --- | --- |
| **Milestones** | **Task** | **Cost Breakdown** | **Timeline (Duration)** |
| Milestones 2 | **Coding which includes following.**  Data preprocessing  Model Development, Implementation and training of model  Evaluation and validation/Testing  etc |  |  |
| Milestones 3 | Scientific Paper Writing.  Submission in journal and addressing the reviewer comments.  and Successful Acceptance |  |  |

* **Deliverables**:
  + Complete Dataset (Original, preprocessed) with documentation
  + Code files + Trained model (with documentation).
  + Manuscript (A research paper detailing the methodology, experiments, results, and conclusions).
  + Visualizations and charts to support the results and analysis.
  + Documentation for the model and code, including instructions for replicating the experiments.