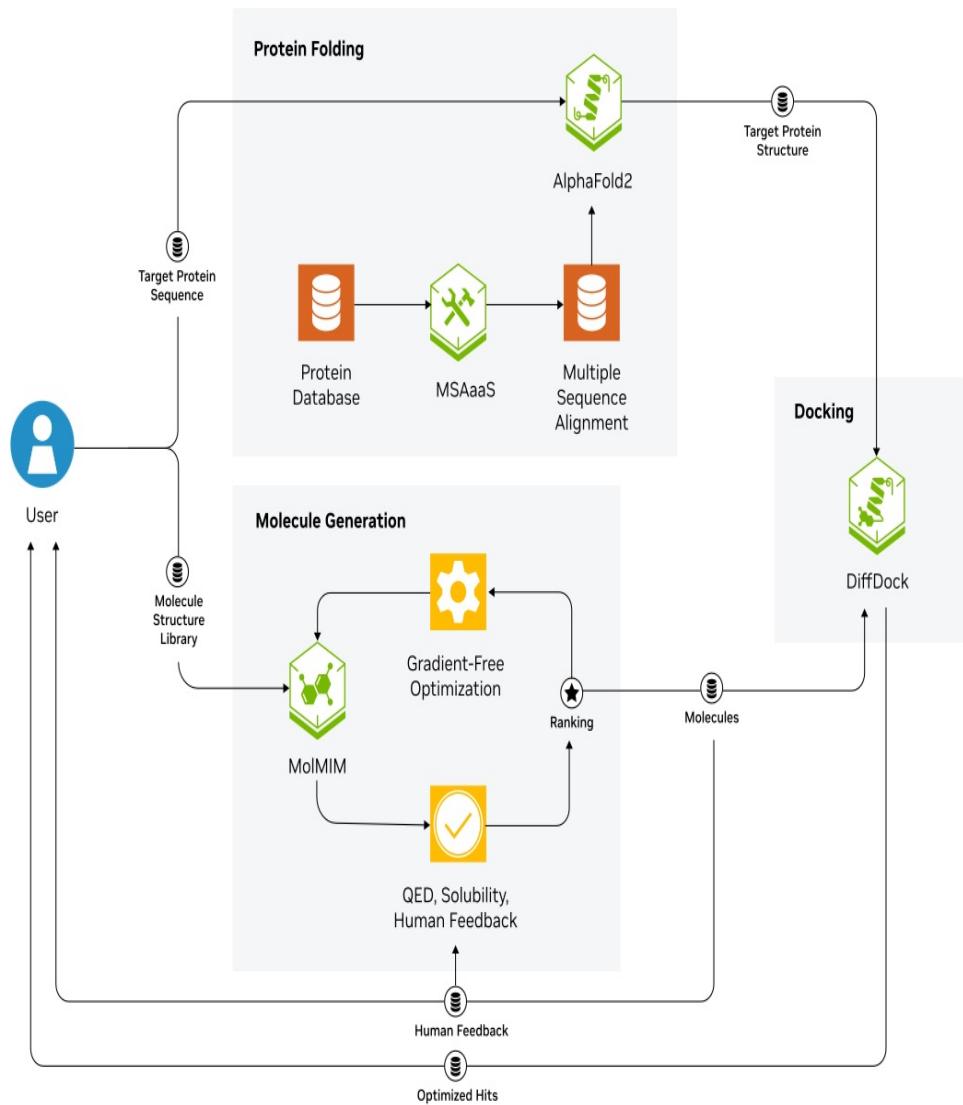




Plan is to create an AI-Driven Drug Discovery Pipeline



Pipeline flow chart

Objective of the pipeline

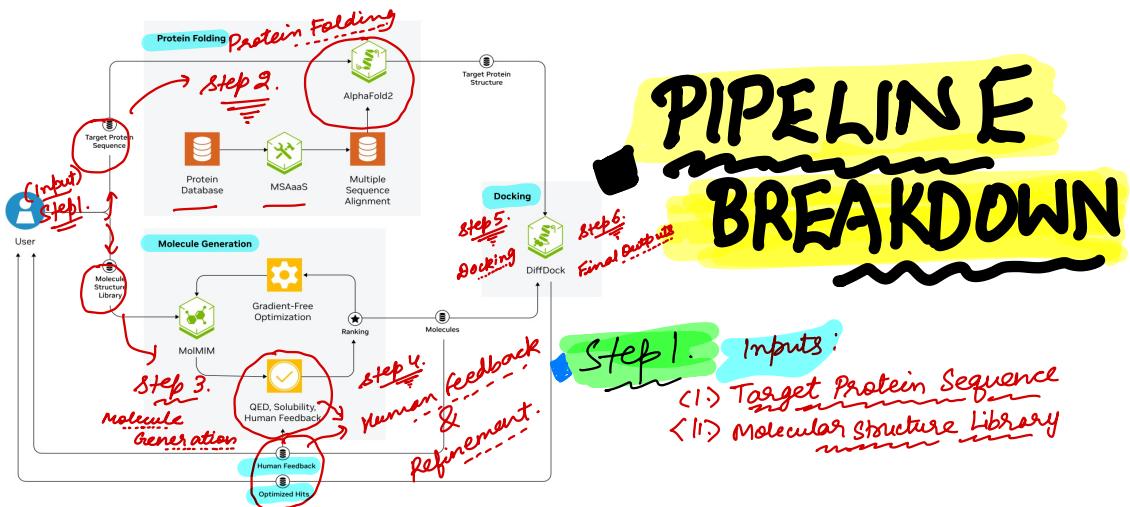
Accelerate drug discovery by leveraging AI-based protein-folding (AlphaFold 2), molecule generation (MolMIM), and docking (DiffDOCK) to find optimal drug candidate faster and cheaper than traditional method.

Why this project?

- ① AI Driven & Novel
- ② Real-world Impact
- ③ Interdisciplinary
- ④ Scalable & Deployable

... || Impact | ... ?

- ★ Can be used for cancer, neurogen. diseases, and Antibiotic Resistance.



Step 2. Protein folding with **AlphaFold 2**
 Goal would be to predict the 3D structure of the target protein from its Amino Acid Sequence ...
 Now? → Multiple Sequence Alignment (MSAaS)
 → AlphaFold2 Generate a 3D structure...

Step 3. Molecule generation with **MolMIM**
 Goal would be to create new small molecules that can potentially bind to the target protein ...
 Now? → MolMIM (Molecular Machine Intelligence)
 → Gradient-Free Opti...!! model ...
 → Ranking & Filtering !!

Step 4. **Human feedback & Refinement** ...
 Goal would be to improve the generated molecule using Expert insights and Additional filters ...
 → Human Experts Review the Molecule
 → Final optimization

Step 5 Docking with DiffDock

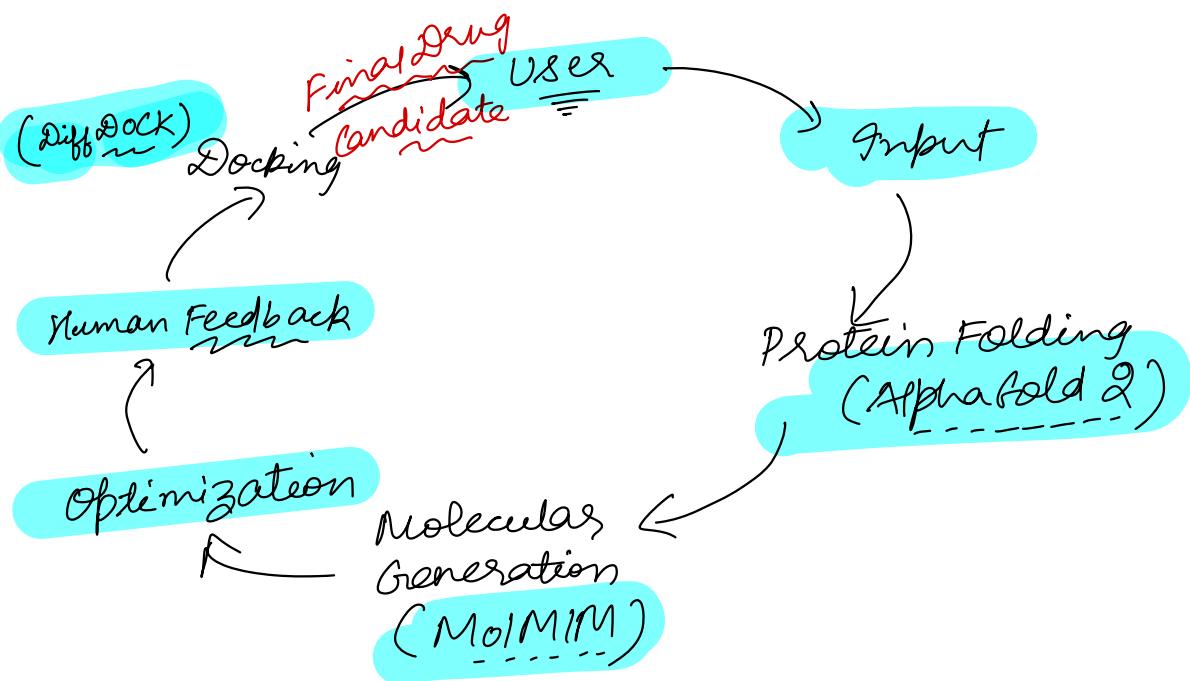
Goal would be to predict how well the optimized will bind to the protein ...

Molecules are input into DiffDOCK
DiffDOCK Predicts Binding Poses --

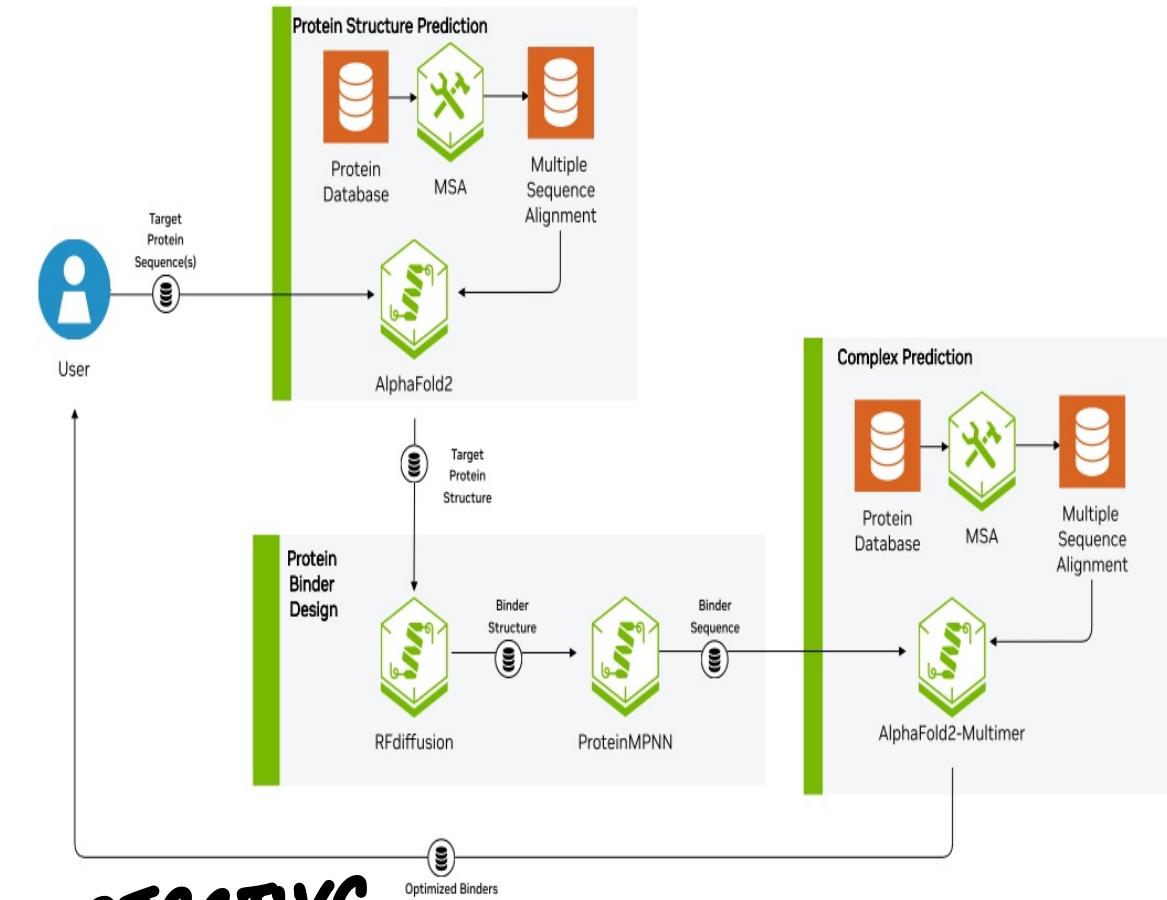
Step 6 Final Output & Optimized Drug Candidate

Goal would be to provide the best drug candidate for further testing ...

... FLOWCHART ...



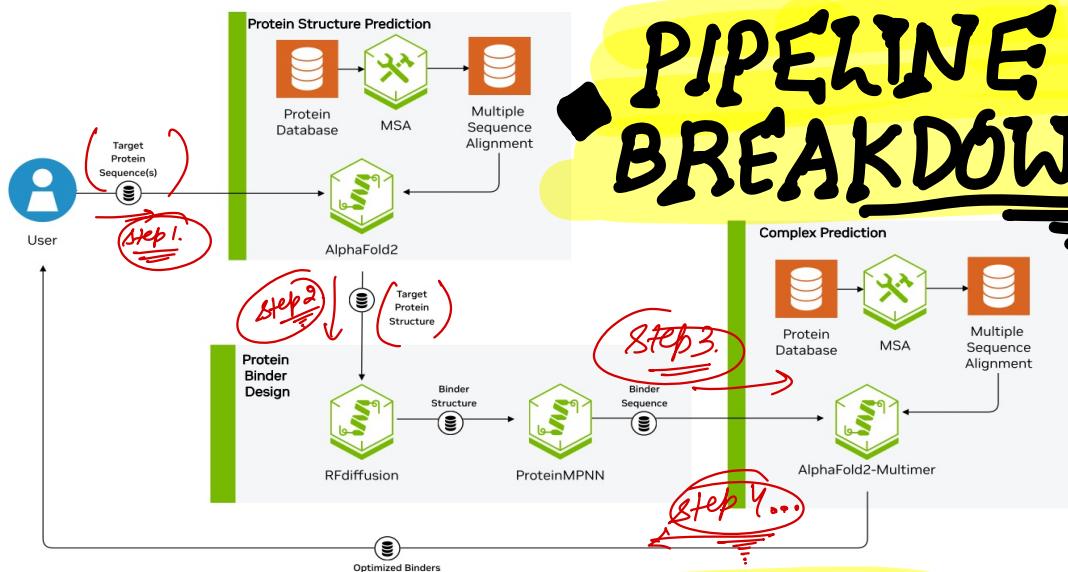
❖ Plan is to create a computational pipeline for Designing Protein Binders...



• OBJECTIVE

Goal is to predict, refine, and optimize protein structures to create stable protein complexes for therapeutic applications...

PIPELINE BREAKDOWN



Step 1. Input :> Target Protein Sequence

- ① The users provides the amino acid sequence of the target protein...
- ② This sequence is passed to AlphaFold2, which predict the 3D structure of the protein...

Step 2. Protein Structure prediction (AlphaFold2)

- ③ The protein Sequence is matched against a protein database to find similar sequences...
- ④ MSAas (Multiple Sequence Alignment as a service) is used to align similar sequences to improve str. prediction...
- ⑤ AlphaFold2 then predicts the 3D structure of the protein, which serve as the base for designing binders...

Step 3. Protein Binder Design (RFDiffusion & proteinMPNN)

- ① The predicted 3D protein structure is then used as input for RFDiffusion, an AI Driven protein design tool...
- ② RFDiffusion explores various binding conformations and optimizes the shape of the protein Binder...
- ③ The generated binder structures is then passed to ProteinMPNN...
- ④ ProteinMPNN optimizes the amino acid sequence of the Binder so that it has the right biochemical properties (e.g., stability, specificity)...

Step 4. complex Stability prediction (AlphaFold2 Multimer)

- ① Once the optimised binder sequence is generated, it is tested for Interaction Stability
- ② The target protein and the designed binders undergo complex prediction using AlphaFold2 Multimer...
- ③ This step Evaluates How well the designed Binders interact with the target protein

Step 5. Outputs: Optimized Binders

- ⑤ The final output is an optimized protein binder that has a stable and effective interaction with the target protein ...
- ⑥ These optimized binders can be used in therapeutic applications (e.g., antibody design, enzyme inhibitors.)

What we will Build ?

Build a Web-App that integrates both the drug-discovery pipeline and the protein - Binder design pipeline into a single AI-powered molecularity modelity platform.

Expected Features

<1> User Input - Panel

- ① User can either :
 - * upload a target protein sequence for protein - Binder design ...
 - * upload a protein target and small molecules for Drug Discovery ...

<11> Backend - AI pipeline

- * Protein - Binder Design pipeline
 - Uses AlphaFold 2 → RFDiffusion → protein MPNN
 - AlphaFold 2 - MultiMer ...
- * Drug Discovery Pipeline
 - Uses AlphaFold 2 → MolMIM → Gradient - free optimization → DiffDOCK ...

<III> Visualization

① 3D-molecular structures...

use PyMOL / NGLView / Py3DMOL to display predicted protein & drug-Molecules...

② Binding Scores & Docking Results ...

Show stability scores, docking interactions, and binding affinity ...

<IV> Downloadable - Output

① Users can download the optimized binders/drug Molecules in PDB / SDF Formats ..

<V> Deployment

① Streamlit / FastAPI for the frontend

② Run the ML Models on a cloud GPU (GCP / Render)...

③ Host it on Render ...

Database (optional) for storing result...

(MongoDB / PostgreSQL)

