SciFetch Report

Generated on: 2025-06-28 22:47 UTC

Request: Recent developments in generative AI for protein structure prediction and drug discovery

Summary

Recent developments in generative AI have significantly impacted the fields of protein structure prediction and drug discovery, offering transformative approaches to longstanding challenges. Several key papers highlight these advancements: 1. **Challenge for Deep Learning: Protein Structure Prediction of Ligand-Induced Conformational Changes at Allosteric and Orthosteric Sites**: This paper discusses the use of deep learning algorithms to predict protein structures, particularly focusing on ligand-induced conformational changes. These changes are crucial for understanding protein function and interactions with drugs, which can lead to more effective drug design. 2. **Integrating Artificial Intelligence in Drug Discovery and Early Drug Development: A Transformative Approach**: This review outlines how AI can be integrated into drug discovery processes to enhance target identification, drug design, and early clinical development. By leveraging multiomics data and network-based approaches, AI can address inefficiencies in traditional drug discovery methods, reducing costs and timelines. 3. **Revolutionizing Molecular Design for Innovative Therapeutic Applications through Artificial Intelligence**: This paper highlights the role of AI and machine learning in molecular design, enabling the creation of proteins with enhanced stability, activity, and specificity. Techniques such as deep learning and reinforcement learning are

emphasized for their ability to transform protein engineering and therapeutic applications. 4. **Integrating Computational Design and Experimental Approaches for Next-Generation Biologics**: This review examines recent advances in computational and experimental methods for engineering therapeutic proteins. It focuses on areas like antibody engineering and enzyme replacement therapies, showcasing how computational tools, including machine learning, are crucial for designing improved protein therapeutics. 5. **Development of a Machine Learning-Based Target-Specific Scoring Function for Structure-Based Binding Affinity Prediction for Human Dihydroorotate Dehydrogenase Inhibitors**: This study presents a machine learning approach to predict binding affinities for inhibitors of human dihydroorotate dehydrogenase, a target for autoimmune disorders and cancer. By integrating interaction and ligand features, the study demonstrates the potential of AI in enhancing drug discovery processes. These papers collectively illustrate the profound impact of generative AI on protein structure prediction and drug discovery, offering new methodologies and insights that promise to accelerate and improve the development of therapeutic agents.

Relevant Articles

1. Whole-Body Conditioned Egocentric Video Prediction

Date: 2025-06-26

Source: arXiv

URL: http://arxiv.org/abs/2506.21552v1

Abstract: We train models to Predict Ego-centric Video from human Actions (PEVA), given the past video and an action represented by the relative 3D body pose. By conditioning on kinematic pose trajectories, structured by the joint hierarchy of the body, our model learns to simulate how physical human actions shape the environment from a first-person point of view. We train an auto-regressive conditional diffusion transformer on Nymeria, a large-scale dataset of real-world egocentric video and body pose ca...

2. Where to find Grokking in LLM Pretraining? Monitor Memorization-to-Generalization without Test

Date: 2025-06-26

Source: arXiv

URL: http://arxiv.org/abs/2506.21551v1

Abstract: Grokking, i.e., test performance keeps improving long after training loss converged, has been recently witnessed in neural network training, making the mechanism of generalization and other emerging capabilities such as reasoning mysterious. While prior studies usually train small models on a few toy or highly-specific tasks for thousands of epochs, we conduct the first study of grokking on checkpoints during one-pass pretraining of a 7B large language model (LLM), i.e., OLMoE. We compute the tr...

3. Natal kick by early-asymmetrical pairs of jets to the neutron star of supernova remnant S147

Date: 2025-06-26

Source: arXiv

URL: http://arxiv.org/abs/2506.21548v1

Abstract: We analyze the bipolar morphology of the jet-shaped core-collapse supernova (CCSN) remnant (CCSNR) S147 and its neutron star (NS) kick velocity, and suggest that two pairs of unequal, opposite jets contributed to the NS kick velocity. This kick by early asymmetrical pairs (kick-BEAP) of jets mechanism operates within the framework of the jittering jets explosion mechanism (JJEM). We examine the prominent pair of large ears and, based on their flat structure rather than the more common conical st...

4. SAM4D: Segment Anything in Camera and LiDAR Streams

Date: 2025-06-26

Source: arXiv

URL: http://arxiv.org/abs/2506.21547v1

Abstract: We present SAM4D, a multi-modal and temporal foundation model designed for promptable segmentation across camera and LiDAR streams. Unified Multi-modal Positional Encoding (UMPE) is introduced to align camera

and LiDAR features in a shared 3D space, enabling seamless cross-modal prompting and interaction. Additionally, we propose Motion-aware Cross-modal Memory Attention (MCMA), which leverages ego-motion compensation to enhance temporal consistency and long-horizon feature retrieval, ensuring r...

5. Data Efficacy for Language Model Training

Date: 2025-06-26

Source: arXiv

URL: http://arxiv.org/abs/2506.21545v1

Abstract: Data is fundamental to the training of language models (LM). Recent research has been dedicated to data efficiency, which aims to maximize performance by selecting a minimal or optimal subset of training data. Techniques such as data filtering, sampling, and selection play a crucial role in this area. To complement it, we define Data Efficacy, which focuses on maximizing performance by optimizing the organization of training data and remains relatively underexplored. This work introduces a gener...

6. DeOcc-1-to-3: 3D De-Occlusion from a Single Image via Self-Supervised Multi-View Diffusion

Date: 2025-06-26

Source: arXiv

URL: http://arxiv.org/abs/2506.21544v1

Abstract: Reconstructing 3D objects from a single image is a long-standing challenge, especially under real-world occlusions. While recent diffusion-based view synthesis models can generate consistent novel views from a single RGB image, they generally assume fully visible inputs and fail when parts of the object are occluded. This leads to inconsistent views and degraded 3D reconstruction quality. To overcome this limitation, we propose an end-to-end framework for occlusion-aware multi-view generation. O...

7. Detecting weighted hidden cliques

Date: 2025-06-26

Source: arXiv

URL: http://arxiv.org/abs/2506.21543v1

Abstract: We study a generalization of the classical hidden clique problem to graphs with real-valued edge weights. Formally, we define a hypothesis testing problem. Under the null hypothesis, edges of a complete graph on n vertices are associated with independent and identically distributed edge weights from a distribution P. Under the alternate hypothesis, k vertices are chosen at random and the edge weights between them are drawn from a distribution Q, while the remaining are sampled from P. ...

8. StruMamba3D: Exploring Structural Mamba for Self-supervised Point Cloud Representation Learning

Date: 2025-06-26

Source: arXiv

URL: http://arxiv.org/abs/2506.21541v1

Abstract: Recently, Mamba-based methods have demonstrated impressive performance in point cloud representation learning by leveraging State Space Model (SSM) with the efficient context modeling ability and linear complexity. However, these methods still face two key issues that limit the potential of SSM: Destroying the adjacency of 3D points during SSM processing and failing to retain long-sequence memory as the input length increases in downstream tasks. To address these issues, we propose StruMamba3D, ...

9. Continuous symmetry breaking in 1D spin chains and 1+1D field theory

Date: 2025-06-26

Source: arXiv

URL: http://arxiv.org/abs/2506.21540v1

Abstract: We argue that ground states of 1D spin chains can spontaneously break U(1) ``easy-plane'' spin rotation symmetry, via true long-range order of (S^x, S^y), at the phase transition between two quasi-long-range-ordered

phases. The critical point can be reached by tuning a single parameter in a Hamiltonian with the same symmetry as the XXZ model, without further fine-tuning. Equivalently, it can arise in systems of bosons with particle-hole symmetry, as a long-range-ordered transition point betwee...

10. WorldVLA: Towards Autoregressive Action World Model

Date: 2025-06-26

Source: arXiv

URL: http://arxiv.org/abs/2506.21539v1

Abstract: We present WorldVLA, an autoregressive action world model that unifies action and image understanding and generation. Our WorldVLA intergrates Vision-Language-Action (VLA) model and world model in one single framework. The world model predicts future images by leveraging both action and image understanding, with the purpose of learning the underlying physics of the environment to improve action generation. Meanwhile, the action model generates the subsequent actions based on image observations, ...

11. Integrating artificial intelligence in drug discovery and early drug development: a transformative approach.

Date: 2025-03-14

Source: PubMed

DOI: 10.1186/s40364-025-00758-2

URL: https://pubmed.ncbi.nlm.nih.gov/40087789

Abstract: Artificial intelligence (AI) can transform drug discovery and early drug development by addressing inefficiencies in traditional methods, which often face high costs, long timelines, and low success rates. In this review we provide an overview of how to integrate AI to the current drug discovery and development process, as it can enhance activities like target identification, drug discovery, and early clinical development. Through multiomics data analysis and network-based approaches, AI can hel...

12. Machine learning prediction of tau-PET in Alzheimer's disease using plasma, MRI, and clinical data.

Date: 2025-05-01

Source: PubMed

DOI: 10.1002/alz.14600

URL: https://pubmed.ncbi.nlm.nih.gov/39985487

Abstract: Tau positron emission tomography (PET) is a reliable neuroimaging technique for assessing regional load of tau pathology in the brain, but its routine clinical use is limited by cost and accessibility barriers....

13. Challenge for Deep Learning: Protein Structure Prediction of Ligand-Induced Conformational Changes at Allosteric and Orthosteric Sites.

Date: 2024-11-01

Source: PubMed

DOI: 10.1021/acs.jcim.4c01475

URL: https://pubmed.ncbi.nlm.nih.gov/39484820

Abstract: In the realm of biomedical research, understanding the intricate structure of proteins is crucial, as these structures determine how proteins function within our bodies and interact with potential drugs. Traditionally, methods like X-ray crystallography and cryo-electron microscopy have been used to unravel these structures, but they are often challenging, time-consuming and costly. Recently, a breakthrough in computational biology has emerged with the development of deep learning algorithms cap...

14. Determination of Potential Lead Compound fromMagnolia officinalisfor Alzheimer's Disease through Pharmacokinetic Prediction, Molecular Docking, Dynamic Simulation, and Experimental Validation.

Date: 2024-09-29

Source: PubMed

DOI: 10.3390/ijms251910507

URL: https://pubmed.ncbi.nlm.nih.gov/39408835

Abstract: Amyloid β protein (Aβ) deposition has been implicated as the molecular driver of Alzheimer's disease (AD) progression. The modulation of the formation of abnormal aggregates and their post-translational modification is strongly suggested as the most effective approach to anti-AD. Beta-site APP-cleaving enzyme 1 (BACE1) acts upstream in amyloidogenic processing to generate Aβ, which rapidly aggregates alone or in combination with acetylcholinesterase (AChE) to form fibrils. Accumulated Aβ promote...

15. Revolutionizing Molecular Design for Innovative Therapeutic Applications through Artificial Intelligence.

Date: 2024-09-29

Source: PubMed

DOI: 10.3390/molecules29194626

URL: https://pubmed.ncbi.nlm.nih.gov/39407556

Abstract: The field of computational protein engineering has been transformed by recent advancements in machine learning, artificial intelligence, and molecular modeling, enabling the design of proteins with unprecedented precision and functionality. Computational methods now play a crucial role in enhancing the stability, activity, and specificity of proteins for diverse applications in biotechnology and medicine. Techniques such as deep learning, reinforcement learning, and transfer learning have dramat...

16. Integration of 3D-QSAR, molecular docking, and machine learning techniques for rational design of nicotinamide-based SIRT2 inhibitors.

Date: 2024-10-10

Source: PubMed

DOI: 10.1016/j.compbiolchem.2024.108242

URL: https://pubmed.ncbi.nlm.nih.gov/39405774

Abstract: Selective inhibitors of sirtuin-2 (SIRT2) are increasingly recognized as potential therapeutics for cancer and neurodegenerative diseases. Derivatives of 5-((3-amidobenzyl)oxy)nicotinamides have been identified as some of the most potent and selective SIRT2 inhibitors reported to date (Ai et al., 2016; Ai et al., 2023, Baroni et al., 2007). In this study, a 3D-QSAR (3D-Quantitative Structure-Activity Relationship) model was developed using a dataset of 86 nicotinamide-based SIRT2 inhibitor...

17. Integrating Computational Design and Experimental Approaches for Next-Generation Biologics.

Date: 2024-08-27

Source: PubMed

DOI: 10.3390/biom14091073

URL: https://pubmed.ncbi.nlm.nih.gov/39334841

Abstract: Therapeutic protein engineering has revolutionized medicine by enabling the development of highly specific and potent treatments for a wide range of diseases. This review examines recent advances in computational and experimental approaches for engineering improved protein therapeutics. Key areas of focus include antibody engineering, enzyme replacement therapies, and cytokine-based drugs. Computational methods like structure-based design, machine learning integration, and protein language model...

18. Development of a machine learning-based target-specific scoring function for structure-based binding affinity prediction for human dihydroorotate dehydrogenase inhibitors.

Date: 2024-09-26

Source: PubMed

DOI: 10.1002/jcc.27510

URL: https://pubmed.ncbi.nlm.nih.gov/39325045

Abstract: Human dihydroorotate dehydrogenase (hDHODH) is a flavin mononucleotide-dependent enzyme that can limit de novo pyrimidine synthesis, making it a therapeutic target for diseases such as autoimmune disorders and cancer. In this study, using the docking structures of complexes generated by AutoDock Vina, we integrate interaction features and ligand features, and employ support vector regression to develop a target-specific scoring function for hDHODH (TSSF-hDHODH). The Pearson correlation coefficie...

19. Discovery of Ureido-Substituted 4-Phenylthiazole Derivatives as IGF1R Inhibitors with Potent Antiproliferative Properties.

Date: 2024-06-04

Source: PubMed

DOI: 10.3390/molecules29112653

URL: https://pubmed.ncbi.nlm.nih.gov/38893528

Abstract: The existing kinase inhibitors for hepatocellular carcinoma (HCC) have conferred survival benefits but are hampered by adverse effects and drug resistance, necessitating the development of novel agents targeting distinct pathways. To discover potent new anti-HCC compounds, we leveraged scaffold hopping from Sorafenib and introduced morpholine/piperidine moieties to develop ureido-substituted 4-phenylthiazole analogs with optimized physicochemical properties and binding interactions. Notably, com...

20. In silicoapproach revealsN-(5-phenoxythiophen-2-yl)-2-(arylthio)acetamides as promising selective SIRT2 inhibitors: the case of structural optimization of virtual screening-derived hits.

Date: 2023-12-19

Source: PubMed

DOI: 10.1080/07391102.2023.2293252

URL: https://pubmed.ncbi.nlm.nih.gov/38112299

Abstract: Epigenetic modifications play an essential role in tumor suppression and promotion. Among the diverse range of epigenetic regulators, SIRT2, a member of NAD+-dependent protein deacetylates, has emerged as a crucial regulator of cellular processes, including cell cycle progression, DNA repair, and metabolism, impacting tumor growth and survival. In the present work, a series ofN-(5-phenoxythiophen-2-yl)-2-(arylthio)acetamide derivatives were identified following a structural optimization of previ...
