**TUGAS RESUME**

**DATA MINING**

Tugas ini disusun untuk Memenuhi tugas mata kuliah metodologi pembelajaran

**Dosen pengampu: Dr. Wahyudi Setiawan, M. Pd.**

****

**Disusun Oleh**

**Giraldo Nainggolan (220441100064)**

**Wisnu Ary Swadana (220441100121)**

**Abib MaulanaAan Nafudi (220441100118)**

**Fairuz Abdullah (220441100070)**

**Program Studi Sistem Informasi**

**Fakultas Teknik**

**Universitas Truojoyo Madura**

**Tahun Ajaran 2024 / 2025**

**The AlphaFold Database of Protein Structures: A Biologist’s Guide**

**Sumber:** Centre for Integrative System Biology and Bioinformatics, Department of Life Sciences, Imperial College London, London SW7 2AZ, UK

https://www.sciencedirect.com/science/article/pii/S0022283621005738

**Penulis:** Alessia David, Sheena E. Radford

**Abstrak**

AlphaFold, algoritme pembelajaran mendalam yang dikembangkan oleh DeepMind, baru-baru ini merilis model tiga dimensi dari seluruh proteom manusia kepada komunitas ilmiah. Di sini kami membahas keunggulan, keterbatasan, dan tantangan yang masih belum terpecahkan dari model AlphaFold dari sudut pandang seorang ahli biologi, yang mungkin bukan ahli dalam biologi struktural

**Pendahuluan**

Klasifikasi data adalah proses pengkategorian data ke dalam kelompok-kelompok yang bermakna. Klasifikasi data banyak digunakan di berbagai bidang, seperti pendidikan, asuransi, media sosial, dan pemasaran.

Pada Juli 2021, model tiga dimensi yang diprediksi untuk seluruh proteom manusia yang dihasilkan menggunakan AlphaFold, algoritme pembelajaran mendalam yang dikembangkan oleh DeepMind, tersedia untuk umum, seperti yang baru-baru ini dilaporkan di Nature.1 Dengan tidak adanya struktur eksperimental, metode komputasi telah digunakan selama beberapa dekade untuk memprediksi model protein tiga dimensi. Sebelum munculnya algoritma AlphaFold, pendekatan utama yang digunakan adalah pemodelan homologi dan abinitio. Dalam pemodelan homologi (atau pendekatan berbasis templat), yang merupakan pendekatan yang paling sukses dan banyak digunakan, model dibangun berdasarkan struktur eksperimental homolog, yang berfungsi sebagai templat struktural. Dalam metode ab-initio (atau pendekatan tanpa templat), model dibangun dengan menggunakan fungsi energi berbasis fisika dan/atau berbasis pengetahuan, dikombinasikan dengan informasi evolusi, yang digunakan untuk menghasilkan peta jarak (atau kontak).

**Kajian Teori**

**Klasifikasi Data**

Klasifikasi data adalah proses pengkategorian data ke dalam kelompok-kelompok yang bermakna. Ada banyak metode klasifikasi data, seperti machine learning, deep learning, dan statistik. Machine learning adalah algoritma yang dapat belajar dari data dan meningkatkan kinerjanya tanpa diprogram secara eksplisit.expand\_more Machine learning dapat digunakan untuk klasifikasi data dengan menganalisis data dan menemukan pola yang dapat digunakan untuk memprediksi kategori data baru.expand\_more

**Pembelajaran Mendalam**

Deep learning adalah subset dari machine learning yang menggunakan jaringan saraf tiruan untuk belajar dari data.expand\_more Deep learning dapat digunakan untuk klasifikasi data dengan menganalisis data dan menemukan pola yang kompleks yang tidak dapat ditemukan dengan metode machine learning tradisional.expand\_more

**Statistik**

Statistik dapat digunakan untuk klasifikasi data dengan menganalisis data dan menemukan distribusi data. Distribusi data dapat digunakan untuk memprediksi kategori data baru.

**Pembahasan**

**https://alphafold.ebi.ac.uk/download**

Tantangan prediksi struktur protein. A) Model AlphaFold dari reseptor hormon pertumbuhan (GHR, UniProt P10912).Ekor intraseluler yang panjang dan tidak terstruktur dari reseptor hormon pertumbuhan (residu 289-638) disajikan dalam warna magenta sebagai filamen panjang dan ditempatkan secara tidak tepat di sebelah domain ekstraseluler.Domain ekstraseluler (residu 19-264) disajikan dengan warna biru dan domain transmembran (residu 265-288) dengan warna cyan. B) Di sebelah kiri, model AlphaFold dari protein PIK3R1 (dalam warna magenta, UniProt P27986). Domain utama PIK3R1 disorot dengan garis putus-putus. Di sebelah kanan, model AlphaFold PIK3R1 (dalam warna magenta) ditumpangkan pada struktur eksperimental PIK3R1 (dalam warna cyan) dalam kompleks dengan PIK3CD (dalam warna hijau; PDB 5M6U). Penempatan interdomain PIK3R1 akan menghasilkan benturan sterik dengan PIK3CD. PI3K-P85-iSH2, subunit pengatur fosfatidilinositol 3-kinase P85 domain inter-SH2.

**Simpulan**

Klasifikasi data dapat membantu industri asuransi dalam memerangi penipuan asuransi. Perusahaan asuransi dapat menggunakan hasil klasifikasi data untuk:

* Cakupan struktural proteom manusia tidak seragam seperti menunjukkan bahwa beberapa kelas protein, seperti target obat,
* algoritme AlphaFold pantas disebut sebagai “pengubah permainan” di bidang biologi struktural dan telah menunjukkan salah satu dari banyak penerapan algoritme pembelajaran mendalam dalam [biomedis](https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/medical-biology)

**Saran**

Klasifikasi data dapat membantu industri asuransi dalam memerangi penipuan asuransi. Perusahaan asuransi dapat menggunakan hasil klasifikasi data untuk:

* AlphaFold harusnya bisa dengan sepenuhnya memecahkan “masalah pelipatan protein” dengan menggunakan perhitungan dari penungurutan klasifikasi pembagian data yang terstruktur.
* Memprediksi posisi relatif domain dalam suatu rantai, bagaimana domain mengubah konformasi relatifnya sebagai respons terhadap rangsangan harus lebih akurat dan lebih detail karna ini mencakup sistem kesehatan seluruh manusia
* Bomain bertransisi dari harus lebih beraturan untuk memesan dan urutkan sesuai klasifikati data numerik.