

# IMPLEMENTASI RANDOM FOREST UNTUK KLASIFIKASI DAN MENGANALISIS AKTIVASI AMPK (AMP-ACTIVATED PROTEIN KINASE) PADA METABOLISME DAN PENYAKIT METABOLIK

M. Gilang Martiansyah M, Ghozi Alvin Karim, Lia Alyani, Nadilla Andhara Putri, Anisa Dini Amalia, M. Faqih

Sains Data, Fakultas Sains, Institut Teknologi Sumatera, Lampung Selatan, Indonesia

## PENDAHULUAN

Peningkatan prevalensi penyakit metabolik seperti diabetes tipe 2 dan obesitas mendorong perlunya analisis regulasi AMPK sebagai target terapeutik potensial, di mana algoritma Random Forest menawarkan solusi efektif untuk mengidentifikasi pola regulasi dari data biologis yang kompleks.

## TUJUAN

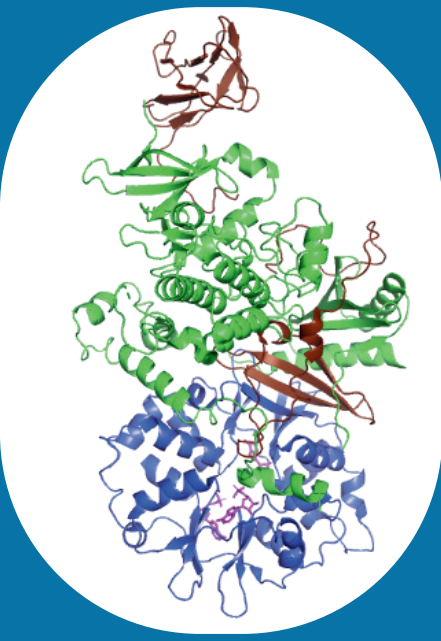
- Menganalisis aktivasi AMPK dalam konteks metabolisme dan penyakit metabolik menggunakan algoritma Random Forest.
- Mengidentifikasi pola regulasi AMPK berdasarkan data biologis yang relevan.
- Mengevaluasi efektivitas algoritma Random Forest dalam seleksi fitur dan klasifikasi data ekspresi genetik yang terkait dengan AMPK.

## DATASET

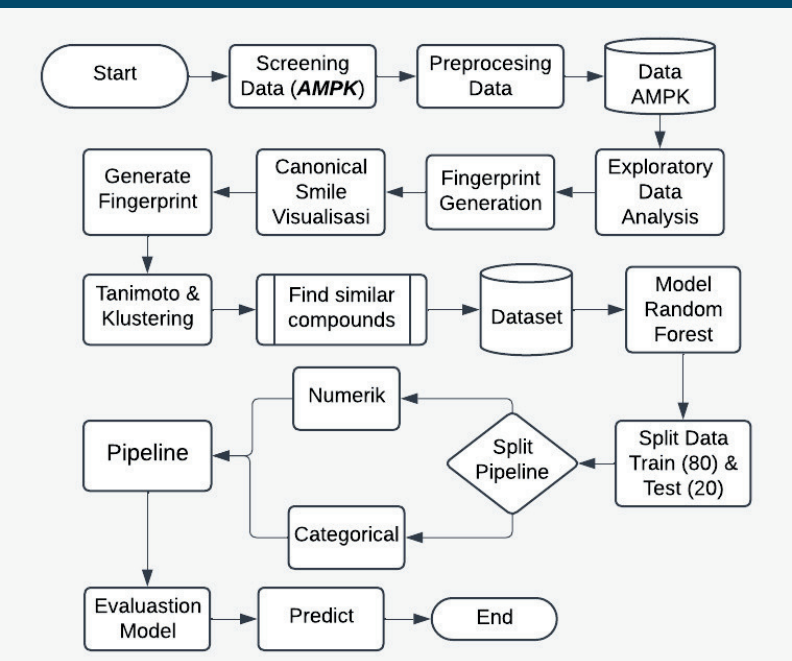
Sumber Data : ChEMBL

Data Target search :

Jumlah Molekul Didapat :

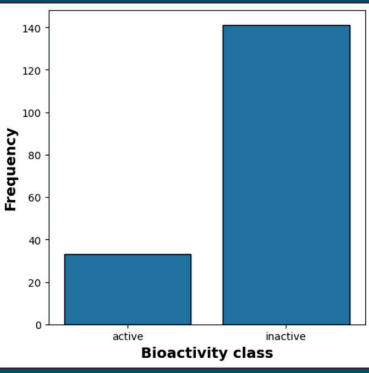


## METODOLOGI



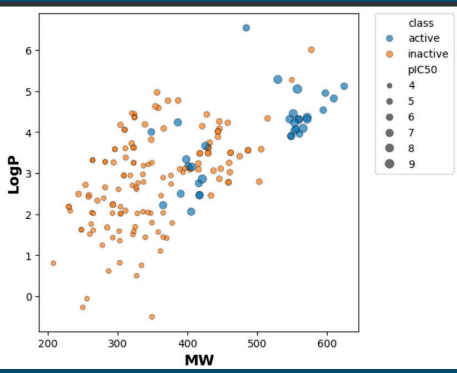
## EXPLORATORY DATA ANALYSIS (EDA)

### Frequency plot



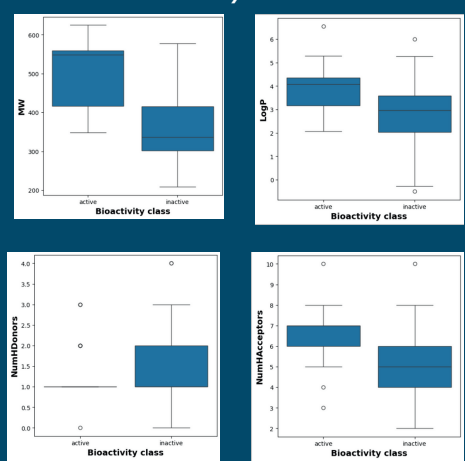
Dapat dilihat pada grafik diatas, molekul yang tidak aktif lebih banyak dibandingkan molekul yang aktif. Hal ini menunjukkan ketidakseimbangan data, dimana kelas "inactive" mendominasi.

### Scatter plot of MW versus LogP

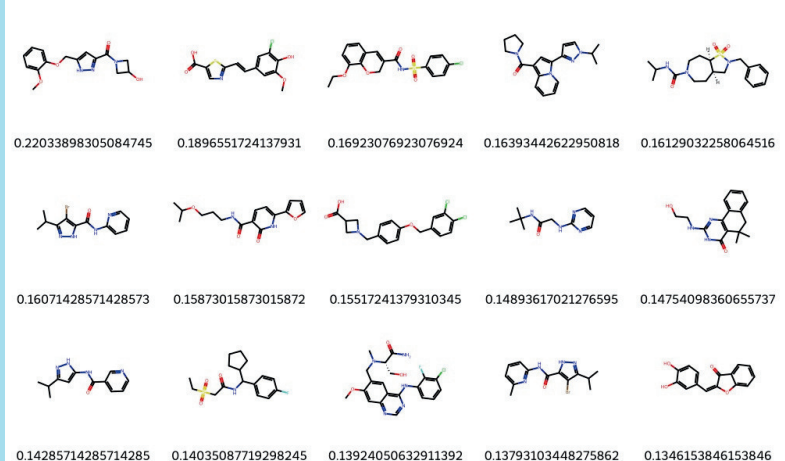


Dapat dilihat pada scatter plot, molekul aktif cenderung memiliki nilai MW dan LogP yang lebih beragam dibandingkan molekul tidak aktif. Nilai pIC50 yang lebih tinggi menunjukkan bahwa molekul aktif memiliki potensi penghambatan yang lebih besar.

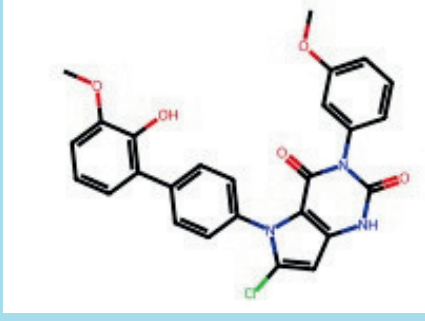
### Statistical Analysis



## HASIL DAN PEMBAHASAN

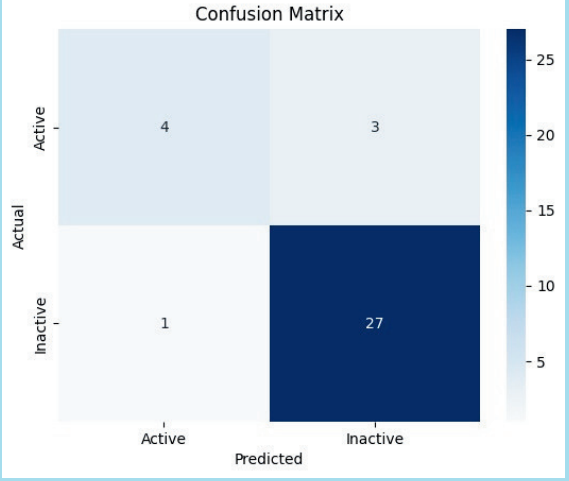


Hasil Tanimoto senyawa yang mirip dengan Resveratrol

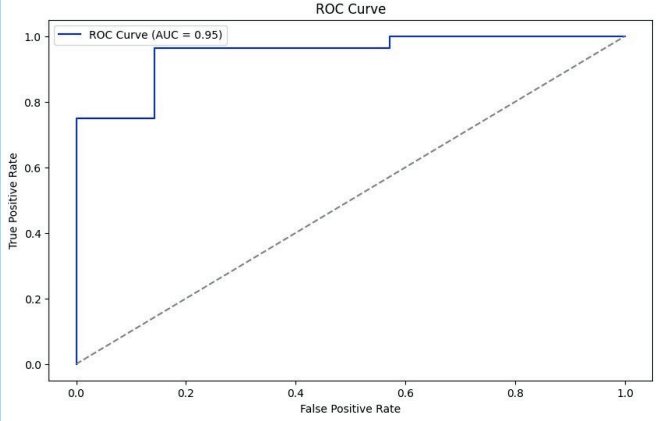


Resveratrol

Resveratrol memiliki struktur dengan cincin aromatik dan gugus hidroksil yang penting untuk aktivitas biologisnya. Hasil Tanimoto menunjukkan beberapa senyawa dengan kemiripan struktural (skor hingga 0.62) terutama berbagi fitur seperti cincin aromatik dan gugus fungsi polar. Senyawa dengan skor tinggi berpotensi memiliki aktivitas biologis serupa terhadap AMPK dan dapat menjadi kandidat untuk pengembangan obat lebih lanjut.



Hasil Evaluasi Confussion Matrix Model RandomForest



Hasil Kurva ROC Model RandomForest

ROC dengan nilai AUC 0.95 menunjukkan kemampuan model yang sangat baik dalam membedakan antara kelas aktif dan tidak aktif, dengan tingkat false positive yang rendah dan true positive rate yang tinggi.