

## 1. Pendahuluan

### Latar Belakang

Obat merupakan produk biologi yang digunakan untuk menyembuhkan dan mencegah suatu penyakit[1]. Filosofi obat konvensional ‘satu gen, satu obat & satu penyakit’ berfokus pada satu target penyakit dan menimbulkan efek samping jika berinteraksi dengan protein lain diluar target[2]. Efek samping dan kurangnya kemanjuran menyebabkan 30% obat gagal dalam uji klinis[2]. Tiap tahun terdapat 2 juta pasien di Amerika Serikat terpengaruh efek samping obat yang berakibat pada 100.000 kematian[2].

Pengembangan obat sangat mahal dan memakan waktu, dibutuhkan 1,8 miliar dolar dan waktu rata-rata 9 hingga 12 tahun[3]. Pengembangan obat dengan cara *in vitro* dan *in vivo* sangat menantang karena perlu dilakukan uji seluler dan biokimia terhadap senyawa, namun protein target dari banyak obat belum ter karakterisasi[4],[5]. Metode *in silico* sangatlah menarik karena hanya menggunakan struktur ‘virtual’ dari senyawa dapat diintegrasikan kedalam tahap awal proses penciptaan obat[6]. Dalam ilmu Farmakologi, *in silico* didefinisikan sebagai penciptaan model komputasi atau simulasi dalam pembuatan prediksi, menyarankan hipotesis dan penemuan atau kemajuan dalam kedokteran dan ilmu terapi[7]. Terdapat beberapa metode yang dapat digunakan dalam memprediksi efek samping obat seperti, *random forest*, *SVM (Support Vector Machine)*, dan *neural network* [8].

Studi pengembangan metode *in silico* dalam prediksi efek samping obat terus dilakukan dengan memanfaatkan *machine learning*. Pada tahun 2019, Xiangxiang Zeng dkk. melakukan penelitian mengenai pendekatan deep learning yang diberi nama *DI* berbasis network yang dapat digunakan untuk memprediksi efek samping suatu obat[9]. Hasil dari penelitian ini adalah *deepDr* memiliki kinerja tinggi (area di bawah kurva karakteristik operasi penerima (AUROC) 0,908) mengungguli *conventional network-based* dan *machine learning-based approaches*. Selain itu, pada tahun 2015 Wen Zhang dkk melakukan penelitian prediksi efek samping obat menggunakan *multi label-learning* dan *ensemble learning*. Metode yang digunakan adalah *feature selection-based multi-label k-nearest neighbor (FS-MLKNN)*[10]. Hasil penelitian menunjukkan bahwa *ensemble learning* yang diintegrasikan dengan FS-MLKNN merupakan alat yang menjanjikan untuk memprediksi efek samping obat dari metode lainnya.

Kemudian, pada tahun 2012 Liu M dkk melakukan penelitian tentang prediksi efek samping obat menggunakan sifat kimia, biologi, dan fenotipik. Metode yang digunakan melalui pendekatan *machine learning* untuk prediksi obat[11]. Setelah dilakukan validasi sebanyak lima kali lipat, algoritma support vector machine (SVM) mengungguli algoritma lainnya. Lalu, pada tahun 2017 Giovanna Maria Dimitri & Pietro Lió melakukan penelitian tentang *DrugClust* (Pendekatan *machine learning* untuk memprediksi efek samping obat) menggunakan metode Bayesian[12]. Hasil penelitian tersebut mengungguli penelitian sebelumnya yang dilakukan oleh Zhang[10], Liu[11], dan Mizutani[13], dimana memiliki validitas yang lebih baik.

Pada tahun 2020, Jiang Li dkk melakukan penelitian tentang metode *machine learning* untuk kombinasi prediksi suatu obat[14]. Hasil penelitian menunjukkan bahwa metode *neighbor recommender* yang dikombinasikan dengan pembelajaran algoritma *ensemble* mengungguli algoritma pembelajaran mesin tradisional.

Pada penelitian ini, akan dibangun model prediksi efek samping obat dengan menggunakan metode *Gray Wolf Optimizer (GWO)* untuk seleksi fitur dan *Support Vector Machine (SVM)* untuk membangun model prediksi. Adapun untuk contoh kasus, penulis akan memprediksi efek samping obat terhadap studi kasus *Hepatobiliary Disorders*. *Hepatobiliary disorders (HBD)* merupakan sebuah penyakit yang menyerang hati, saluran empedu, dan kantong empedu [15]. Algoritma *Gray Wolf Optimizer (GWO)* dipilih karena memiliki keseimbangan yang baik antara eksplorasi dan eksploitasi yang menghasilkan penghindaran local optima yang tinggi [16] dan implementasi GWO untuk prediksi efek samping obat belum banyak dilakukan dalam penelitian. SVM dipilih karena memiliki konsep yang lebih jelas secara matematis dan memiliki kelebihan dalam menentukan jarak menggunakan *support vector* sehingga proses komputasi menjadi cepat [17], [18] dan penerapan SVM dalam studi kasus *Hepatobiliary Disorders* belum banyak dilakukan.

### Topik dan Batasannya

Topik yang dibahas pada tugas akhir ini adalah Prediksi efek samping obat pada kasus *Hepatobiliary Disorders* menggunakan Metode *Fingerprint-based Support Vector Machine* yang Dioptimasi dengan *Grey Wolf Optimization*. Batasan masalah yang ada yaitu data teknis merupakan data Sider yang menyediakan hasil molekul untuk satu set *smiles*. Data yang ada merupakan nilai eksperimental yang dilaporkan dalam literatur ilmiah.

### Rumusan Masalah

Rumusan masalah yang diangkat pada penelitian ini yaitu:

1. Apakah metode *Support Vector Machine* yang dioptimasi menggunakan *Gray Wolf Optimization* dapat digunakan untuk klasifikasi?

2. Bagaimana pembangunan model prediksi menggunakan *Fingerprint-based Support Vector Machine* yang dioptimasi dengan *Gray Wolf Optimization*?
3. Bagaimana performa dari model yang telah dibuat?

### **Tujuan**

Tujuan yang ingin dicapai pada penelitian ini yaitu:

1. Mengimplementasikan metode *Gray Wolf Optimization* dalam melakukan optimasi arsitektur *Support Vector Machine* untuk prediksi efek samping obat *Hepatobiliary Disorders*.
2. Membangun model prediksi efek samping obat *Hepatobiliary Disorders* dengan metode *Support Vector Machine* (SVM).
3. Mendapatkan hasil performa metode *Fingerprint-based Support Vector Machine* yang dioptimasi dengan *Gray Wolf Optimization* saat melakukan klasifikasi untuk penemuan obat.