

## 1. Pendahuluan

### Latar Belakang

Obat-obatan sering kali memiliki peran ganda, baik untuk tujuan pengobatan maupun non-terapi. Semua obat memiliki efek samping. Efek samping obat adalah dampak farmakologis yang melewati batas dari tujuan terapeutik yang terjadi setelah penggunaan obat[1]. Efek samping yang tidak disengaja adalah salah satu penyebab kekecewaan yang paling umum dalam penanganan perbaikan pengobatan. Sebagai gambaran, efek samping obat yang serius dinilai sebagai penyebab keempat kematian di Amerika Serikat yang terjadi pada 100.000 kematian per tahun[2],[3]. Obat-obatan dapat mengikat protein "tidak tepat sasaran", yang berpotensi menyebabkan efek samping yang tidak diinginkan atau ADR, mulai dari kemalasan yang ringan hingga kardiotoksitas yang berbahaya[4]. Dalam dekade terakhir, kekhawatiran akan toksisitas obat pada dasarnya telah meluas (dari 10% menjadi 20% dalam satu dekade).

Studi sistematis dan kuantitatif terhadap dampak efek samping obat bersifat antagonis menjadi semakin penting karena meningkatnya kekhawatiran mengenai sitotoksitas obat yang sedang dikembangkan[5]. Kesulitan utama dalam menemukan efek samping obat melalui eksperimen adalah bahwa merancang uji klinis untuk menemukan efek samping biasanya tidak efisien dalam hal biaya[6], maupun waktu, dan yang terpenting adalah sangat mahal dan memakan waktu. Untuk mempercepat waktu dan mengurangi biaya dalam proses penemuan obat, sangat penting untuk menggunakan pembelajaran mesin untuk mengidentifikasi potensi efek samping sejak dini[3],[4]. Beberapa peneliti sebelumnya telah melakukan penelitian terkait.

Beberapa peneliti sebelumnya telah melakukan penelitian terkait. Pada tahun 2017, Dimitri dan tim melakukan penelitian yang mengusulkan algoritma DrugClust yang mengkategorikan obat berdasarkan fitur-fiturnya dan memprediksi efek samping dengan menggunakan skor Bayesian. Hasil penelitian menunjukkan kinerja yang lebih baik dalam hal metrik Area Under Curve (AUC) dan Area Under Precision-Recall Curve (AUPR) dengan nilai rata-rata 0,3 pada dataset Zhang[6]. Pada tahun 2019, Xiangxiang Zeng dan tim melakukan eksplorasi dengan menggunakan metode machine learning yang disebut DeepDR untuk mempelajari efek samping obat. Hasil penelitian menunjukkan bahwa metode yang digunakan berhasil mencapai performa yang lebih unggul dengan nilai 0.908, dibandingkan dengan metode konvensional[7].

Kurniawan dan tim pada tahun 2021 melakukan penelitian yang menghasilkan model *Quantitative Structure-Activity Relationship*(QSAR) untuk memprediksi aktivitas protease inhibitor HIV-1 dengan menggunakan *Neural Network*(NN) dan *Gravitational Search Algorithm*(GSA). Hasilnya menunjukkan bahwa model ketiga dengan tujuh deskriptor menunjukkan hasil terbaik dengan koefisien kepastian ( $r^2$ ) sebesar 0,84 dan 0,82 untuk informasi pengujian[8]. Pada tahun 2021, Jaganathan dan tim membangun model hubungan struktur-aktivitas kuantitatif untuk memprediksi efek samping obat yang menyebabkan toksisitas hati, khususnya dengan memanfaatkan metode klasifikasi berbasis *Support Vector Machine*(SVM). Pada validasi internal, temuan penelitian ini mencapai akurasi sebesar 0,811 dan koefisien korelasi Mathew sebesar 0,623[9].

Pada tahun 2023, Mazumdar dan kawan-kawan mengembangkan model pembelajaran mesin *Deep Neural Network*(DNN) untuk memprediksi bagaimana senyawa tertentu dapat memengaruhi ginjal. Hal ini sangat penting dalam pengembangan obat. Untuk memastikan keakuratan model, mereka menggunakan lima kali validasi silang untuk validasi model, ada 52 kali lipat dengan skor *Receiver Operating Characteristic - Area Under Curve*(ROC-AUC) di atas rata-rata (0,75). Hasilnya menunjukkan nilai generalisasi yang lebih baik dalam validasi silang[10]. Pada tahun 2023, Wiliatama dan tim mengembangkan model prediksi efek samping obat menggunakan *Gravitational Search Algorithm*(GSA) dan beberapa metode ensemble yang berfokus pada studi kasus gangguan hepatobilier. Penelitian ini menemukan bahwa model *Random Forest* memiliki performa terbaik, dengan akurasi 0.68 dan nilai F1 sebesar 0.77[11].

Berdasarkan beberapa penelitian, proses beberapa metode komputasi telah digunakan untuk memprediksi hubungan efek samping obat, namun masih ada ruang untuk peningkatan efektivitasnya[1],[4]. Oleh karena itu, metode heuristik diperlukan. *Gravitational Search Algorithm*(GSA) adalah salah satu strategi heuristik, kelebihan dari teknik heuristik adalah kemampuan untuk menemukan pengaturan ideal di seluruh dunia, hal ini dapat bermanfaat dalam penentuan sorotan, yang membantu memilih informasi utama untuk diperiksa. Hal ini menjauhkan diri dari batasan yang terkait dengan mengulur-ulur pengaturan sub-parameter[11].

Penelitian ini bertujuan untuk mengembangkan model prediksi dengan menggunakan metode *Gravitational Search Algorithm*(GSA) dan *Support Vector Machine*(SVM) untuk efek samping obat. Sebuah studi kasus gangguan mata akan menjadi subjek penelitian ini. Berdasarkan hukum kedua Newton tentang gerak, metode *Gravitational Search Algorithm*(GSA) mengikuti gerak partikel dengan massa tertentu di seluruh ruang pencarian. Metode GSA dapat sangat membantu dalam menghindari masalah pada *local optima* dengan menggunakan ambang batas volume, yang memungkinkan konvergensi yang lebih cepat ke solusi[12]. Komunitas penelitian statistik dan pembelajaran mesin berkolaborasi untuk menciptakan metode *Support Vector Machine*(SVM). Manfaat mendasar dari teknik SVM adalah kemampuannya untuk menaklukkan 'masalah berlapis tinggi' dalam karakterisasi[13].

### **Rumusan Masalah**

Berdasarkan dengan penjelasan pada latar belakang, maka dapat diangkat beberapa rumusan masalah sebagai berikut:

1. Bagaimana efektifitas seleksi fitur menggunakan metode *Gravitational Search Algorithm*(GSA)?
2. Bagaimana efektifitas optimasi *hyperparameter tuning* menggunakan metode *Support Vector Machine*(SVM)?
3. Bagaimana hasil performa *Gravitational Search Algorithm - Support Vector Machine*(GSA-SVM) dalam memprediksi efek samping obat pada kasus gangguan mata?

### **Batasan Masalah**

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Seleksi fitur dilakukan menggunakan metode *Gravitational Search Algorithm*(GSA)
2. Optimasi hyperparameter tuning dilakukan menggunakan metode *Support Vector Machine*(SVM)
3. Evaluasi performa dilakukan dengan menggunakan metrik seperti akurasi, presisi, recall, dan F1-score.

Penelitian ini hanya berfokus pada kombinasi metode *Search Algorithm*(GSA) dan *Support Vector Machine*(SVM) untuk memprediksi efek samping obat dengan studi kasus gangguan mata, tidak membandingkan dengan metode lainnya.

## Tujuan

Tujuan penelitian ini adalah, sebagai berikut:

1. Mengimplementasikan metode *Gravitational Search Algorithm*(GSA) untuk melakukan seleksi fitur.
2. Menerapkan metode *Support Vector Machine*(SVM) untuk melakukan optimasi.
3. Mendapatkan hasil performa prediksi efek samping obat pada studi kasus gangguan mata menggunakan *Gravitational Search Algorithm - Support Vector Machine*(GSA-SVM).

## Organisasi Tulisan

Organisasi tulisan pada laporan penelitian ini adalah sebagai berikut:

1. Pada sub studi terkait peneliti melakukan tinjauan pustaka terhadap beberapa paper yang menjadi referensi. Kemudian, mencari tinjauan untuk teori mengenai, perkembangan penggunaan machine learning dalam prediksi efek samping obat, penggunaan metode *Gravitational Search Algorithm*(GSA) dalam prediksi efek samping obat, dan metode *Support Vector Machine*(SVM) dalam prediksi efek samping obat.
2. Pada sub sistem yang dibangun berisikan perancangan model prediksi yang akan dibangun. Metode heuristik yaitu *Gravitational Search Algorithm*(GSA) akan digunakan sebagai metode untuk seleksi fitur. Kemudian model akan dibangun menggunakan metode *Support Vector Machine*(SVM) dengan tiga kernel yaitu *Linear*, *Radial Basis Function*(RBF), dan *Polynomial*. Kemudian pada sub-bagian akan mendeskripsikan dataset yang digunakan, proses seleksi fitur, pengembangan, hingga evaluasi pada model.
3. Pada sub evaluasi dijelaskan hasil pengujian dan analisis dari model prediksi yang dibangun. Pengujian dilakukan dengan hyperparameter tuning untuk meningkatkan akurasi dan menentukan parameter terbaik. Adapun parameter yang digunakan pada ketiga kernel adalah sebagai berikut: a.  $C = \{0.1, 1, 10, 100, 1000\}$ , b.  $\text{Gamma} = \{\text{'scale'}, \text{'auto'}\}$ , dan c.  $\text{Degree} = \{1, 2, 3, 4, 5\}$ . Dari ketiga parameter tersebut, didapatkan nilai best parameter yang lebih baik dari nilai default parameter. Evaluasi masing-masing hasil set pelatihan dan set pengujian dari ketiga kernel dilakukan menggunakan *confusion matrix*. Setelah itu, dilakukan analisis terhadap hasil performa set pelatihan dan set pengujian dari ketiga kernel berdasarkan beberapa metrik, yaitu *Accuracy*, *Recall*, *Precision*, dan *F1-Score*.

Pada sub kesimpulan berisikan kesimpulan dari penelitian yang telah dilakukan, dan saran untuk penelitian yang akan dilakukan pada masa mendatang.