

1. Pendahuluan

Latar Belakang

Obat adalah kombinasi zat, termasuk produk biologis yang digunakan untuk mempengaruhi atau mengubah sistem fisiologis atau kondisi patologis dalam hal diagnosis, pengobatan zat, pencegahan, peningkatan kesehatan, dan kesadaran manusia[1]. Obat-obatan memberikan kontribusi signifikan terhadap peningkatan kualitas hidup dengan menjaga kesehatan tubuh serta membantu dalam pemulihan dari kondisi patologis. Obat-obatan juga memainkan banyak peran di luar penyembuhan, mereka juga digunakan untuk pencegahan, membantu diagnosis, dan mengurangi gejala[2]. Tetapi, selain memiliki manfaat, obat juga mengandung toksisitas yang tinggi. Toksisitas adalah sifat atau kemampuan suatu zat untuk menyebabkan kerusakan atau efek merugikan pada organisme hidup. Toksisitas mempengaruhi berbagai sistem pada tubuh, mengakibatkan komplikasi serius seperti kerusakan organ dan disfungsi kognitif[3]. Dalam sebuah makalah hasil penelitian membahas tingkat gesekan obat yang tinggi dalam uji klinis, dengan sekitar 90% gagal berkembang karena masalah toksisitas tak terduga yang diidentifikasi selama tahap praklinis[4]. Masalah ini menekankan perlunya metode yang lebih aman dan akurat dalam mengevaluasi toksisitas obat.

Pada tahun 2020, perusahaan farmasi dan regulator di banyak negara mengandalkan pengujian *in vitro* dan *in vivo*[5]. Pengujian *in vitro* dan *in vivo* pada hewan coba dilakukan untuk mendapatkan data toksisitas akut. Hasil dari pengujian *in vitro* dan *in vivo* pada hewan coba dapat diimplementasikan ke manusia atau digunakan sebagai bahan prediksi toksikologi untuk hewan domestik dan ternak[6]. Meskipun pendekatan ini memberikan data yang diperlukan bagi peneliti, tetapi pendekatan ini mempunyai kelemahan tertentu seperti, kemampuan prediksi pengujian *in vitro* dan *in vivo* untuk toksisitas manusia tidak selalu akurat, tingginya biaya, waktu yang lama, dan pertanyaan etis terkait penggunaan hewan coba. Kurangnya prediktibilitas ini dapat mengakibatkan reaksi merugikan yang tidak terduga pada manusia meskipun lulus pengujian pada hewan[5].

Seiring dengan berkembangnya teknologi, metode *in silico* telah menjadi alternatif yang menarik dalam penelitian toksisitas obat. Salah satu metode *in silico* yang menjanjikan adalah implementasi *Machine Learning* (ML) untuk prediksi toksisitas[7]. Dalam sebuah penelitian, ada tujuh metode *Machine Learning*, yaitu *naive Bayes*, *support vector machine (SVM)*, *recursive partitioning*, *k-nearest neighbor*, *C4.5 decision tree*, *random forest* dan *Adaboost*, digunakan untuk membangun model klasifikasi biner untuk toksisitas perkembangan. Di antara model-model ini, pengklasifikasi SVM mewakili kinerja dan stabilitas prediksi terbaik, yang memberikan akurasi prediksi keseluruhan 91,11%, akurasi seimbang 91,50%, dan 0,818 MCC untuk set pelatihan, serta menghasilkan 83,93% konkordansi, 81,85% akurasi seimbang, dan 0,627 MCC untuk set tes. Domain aplikasi dianalisis dan hanya satu bahan kimia dalam set pengujian yang diidentifikasi berada di luar domain aplikasi[8]. Untuk mempertahankan bagian prediksi dan meningkatkan hasil dengan meningkatkan fitur dapat menggunakan algoritma yang membantu bagian prediksi obat menjadi lebih efektif, seperti algoritma *Artificial Bee Colony* (ABC). Dilakukan penelitian oleh Zarei et al. mengembangkan model *Quantitative Structure-Activity Relationship* (QSAR) untuk memprediksi toksisitas benzena tersubstitusi ke *Tetrahymena pyriformis* menggunakan algoritma ABC untuk pemilihan deskriptor[9]. Sebagai algoritma optimasi, *Artificial Bee Colony* (ABC) digunakan dalam prediksi toksisitas. Hal ini dapat membantu dalam menentukan parameter yang paling efektif untuk prediksi toksisitas suatu senyawa. seperti, sifat farmakologis, fisiologis, dan toksisitas dari suatu senyawa yang diberikan terhadap organ target. Dengan cara ini, ABC dapat membantu mengembangkan pengobatan yang lebih aman dan efektif serta menurunkan risiko terhadap kesehatan manusia dan lingkungan[10]. Meskipun demikian, penelitian di bidang ini masih memiliki ruang untuk pengembangan lebih luas. Misalnya, belum banyak penelitian yang dilakukan pada kombinasi ABC-SVM dalam penelitian toksisitas AR-LBD.

Androgen Receptor (AR) adalah jenis protein yang mengikat androgen, yang merupakan hormon pria seperti testosteron. Ini memainkan peran penting dalam berbagai proses biologis, termasuk pengembangan karakteristik pria dan regulasi penyakit tertentu, seperti kanker prostat. Sedangkan, *Ligand Binding Domain* (LBD) adalah bagian spesifik dari reseptor androgen yang berinteraksi dengan androgen. Ketika androgen mengikat LBD, lalu mengaktifkan reseptor, memungkinkannya mempengaruhi ekspresi gen dan aktivitas seluler. Memahami AR dan LBD sangat penting dalam toksikologi karena banyak bahan kimia dapat meniru atau memblokir aksi androgen. Hal ini dapat menyebabkan efek buruk pada kesehatan, terutama pada penyakit terkait hormon[11]. Dalam sebuah penelitian menggunakan *Deep Neural Network* (DNN) sebagai pendekatan pembelajaran mendalam dan *Random Forest* (RF) sebagai metode pembelajaran dangkal untuk memprediksi toksisitas kimia berdasarkan hubungan struktur-aktivitas mendapatkan hasil perbandingan kinerja DNN secara signifikan mengungguli algoritma RF dalam memprediksi toksisitas kimia. DNN menunjukkan peningkatan 22-27% di empat metrik (*precision*, *recall*, *F-measure*, dan *AUPRC*), dengan nilai p kurang dari 0,001 menunjukkan signifikansi statistik[11].

Penelitian ini memiliki tujuan memprediksi toksisitas senyawa dalam studi kasus toksisitas domain *Androgen Receptor - Ligand Binding Domain* (AR-LBD) menggunakan metode ABC-SVM. Algoritma ABC merupakan algoritma optimasi yang digunakan untuk seleksi fitur dalam klasifikasi toksisitas AR-LBD. Pada penelitian ini, SVM digunakan sebagai teknik optimasi atau klasifikasi untuk menemukan parameter optimal dalam ABC. *Support Vector Machine* (SVM) adalah pengklasifikasi yang kuat dalam pembelajaran mesin, dikenal karena akurasi dan kemampuannya untuk menangani masalah matematika yang kompleks[12], [13]. Dalam studi interaksi obat, SVM mengungguli metode lain, mencapai akurasi 93,76% dan menunjukkan potensi untuk mengungkap

jenis interaksi baru[14]. Dengan demikian, pengembangan model prediktif berbasis *machine learning* seperti ABC-SVM dapat membantu mengatasi kelemahan pengukuran toksisitas secara konvensional. Model ini dapat memprediksi toksisitas suatu zat dengan lebih akurat dan efektif, serta dapat membantu dalam pengembangan obat-obatan dan bahan kimia yang lebih aman dan efektif. Kombinasi kedua metode ini yaitu, *Artificial Bee Colony* (ABC) dan *Support Vektor Machine* (SVM) dapat memberikan prediksi toksisitas obat atau virus seperti AR-LBD.

Topik dan Batasannya

Penelitian ini berfokus pada prediksi toksisitas obat dengan menggunakan kombinasi *Artificial Bee Colony* (ABC) sebagai metode seleksi fitur dan *Support Vector Machine* (SVM) sebagai algoritma klasifikasi, dengan studi kasus toksisitas *Androgen Receptor-Ligand Binding Domain* (AR-LBD). Dalam penelitian ini, dataset yang digunakan berasal dari *Tox21 Data Challenge*, yang terdiri dari 8.599 sampel dan 883 fitur yang mencakup *fingerpint* molekuler PubchemFP. Proses seleksi fitur dilakukan menggunakan ABC, yang bertujuan untuk mengurangi jumlah fitur tanpa kehilangan informasi penting. Kemudian, model dikembangkan menggunakan SVM dengan tiga jenis kernel (Linear, Polynomial, RBF) serta dilakukan *hyperparameter tuning* menggunakan *Grid Search*.

Penelitian ini memiliki batasan, yaitu tidak membahas efek toksisitas spesifik pada organ atau mekanisme biologis lebih lanjut, hanya menggunakan ABC sebagai model seleksi fitur tanpa melakukan perbandingan dengan metode lain, serta hanya menggunakan SVM sebagai algoritma klasifikasi tanpa eksplorasi terhadap metode lain seperti *Random Forest* atau *Neural Networks*. Selain itu, penelitian ini juga tidak melakukan analisis interpretasi biologis terhadap senyawa toksik yang teridentifikasi. Dengan demikian, penelitian ini bertujuan untuk mengoptimalkan seleksi fitur dan klasifikasi toksisitas AR-LBD secara efisien menggunakan ABC-SVM, serta memberikan kontribusi dalam pengembangan model prediksi berbasis *Machine Learning* untuk bioinformatika.

Tujuan

Penelitian ini bertujuan untuk mengimplementasikan algoritma *Artificial Bee Colony* (ABC) sebagai metode seleksi fitur dalam analisis data toksisitas AR-LBD. Proses ini dilakukan untuk mengidentifikasi fitur-fitur yang paling relevan, sehingga dapat meningkatkan efisiensi dan akurasi prediksi. Selain itu, penelitian ini juga bertujuan untuk memanfaatkan algoritma *Support Vector Machine* (SVM) dengan pengoptimalan parameter guna meningkatkan kinerja model prediksi. Pengoptimalan ini melibatkan penyesuaian parameter inti SVM, seperti kernel dan regulasi, untuk mendapatkan hasil yang optimal berdasarkan metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*. Akhirnya, penelitian ini diharapkan mampu mengevaluasi efektivitas kombinasi algoritma ABC dan SVM dalam memprediksi toksisitas pada kasus AR-LBD, sehingga memberikan wawasan baru terkait penerapan metode ini dalam analisis bioinformatika.

Tabel 1. Keterkaitan antara tujuan, pengujian dan kesimpulan

No	Tujuan	Pengujian	Kesimpulan
1	Untuk mengimplementasi metode algoritma <i>Artificial Bee Colony</i> untuk seleksi fitur.	Pengujian efektivitas ABC dalam seleksi fitur berdasarkan jumlah fitur terpilih dan <i>best fitness</i> .	Algoritma ABC berhasil mengurangi jumlah fitur tanpa menurunkan performa model, dengan Kernel RBF memilih 77 fitur dan <i>best fitness</i> 0.123, menunjukkan efektivitas seleksi fitur.
2	Untuk menggunakan algoritma <i>Support Vector Machine</i> untuk pengoptimalan parameter.	Pengujian dengan <i>Grid Search</i> untuk mendapatkan kombinasi <i>hyperparameter</i> terbaik (C, gamma, degree).	Pengoptimalan parameter dengan <i>Grid Search</i> meningkatkan performa model, dengan Kernel RBF mengalami peningkatan akurasi dari 0.958 menjadi 0.971 setelah <i>tuning</i> parameter.
3	Untuk mengetahui algoritma <i>Artificial Bee Colony</i> dan <i>Support Vector Machine</i> dalam prediksi toksisitas kasus AR-LBD.	Pengujian akurasi model SVM setelah seleksi fitur dan tuning parameter menggunakan metrik evaluasi (<i>accuracy</i> , <i>precision</i> , <i>recall</i> , <i>F1-score</i>).	Kombinasi ABC dan SVM dengan Kernel RBF menunjukkan performa terbaik, dengan akurasi 0.97, <i>precision</i> 0.58, <i>recall</i> 0.69, dan <i>F1-score</i> 0.63 pada data uji, membuktikan efektivitas model dalam klasifikasi toksisitas AR-LBD.

Organisasi Tulisan

Penyusunan tulisan dimulai dengan tinjauan literatur yang mencakup analisis pengujian pada berbagai topik. Kemudian, penjelasan tentang metodologi yang digunakan dalam penelitian diberikan, dan diskusi tentang evaluasi hasil penelitian dilakukan di langkah berikutnya. Terakhir, makalah akan mencakup temuan dan saran untuk penelitian yang akan datang.