

Machine Learning Based QSAR Model for Screening PPAR-gamma Stimulator

Edwin Nirwana^{1,2}, Gaga I. Nugraha^{3*}, Eko F. Ariyanto⁴, Ardo Sanjaya⁵

¹Magister Program of Biomedical Science, Faculty of Medicine, Universitas Padjadjaran, Sumedang, Indonesia

²Department of Biomedical Science, Faculty of Medicine, Parahyangan Catholic University, Bandung, Indonesia

³Department of Biomedical Science, Faculty of Medicine, Universitas Padjadjaran, Sumedang, Indonesia

⁴Division of Biochemistry and Molecular Biology, Department of Biomedical Science, Faculty of Medicine, Universitas Padjadjaran, Sumedang, Indonesia

⁵Department of Anatomy, Faculty of Medicine, Maranatha Christian University, Bandung, Indonesia

Abstract

Type-2 diabetes (DM2) is still growing into a global problem. Decreased insulin sensitivity is the main mechanism causing DM2. PPAR-gamma has been shown to have a direct relationship with insulin sensitivity. This study aimed to develop a machine-learning model to evaluate phytochemical compounds' activity as a PPAR-gamma agonist. The dataset used for modeling was procured from the ChEMBL database. A total of 3668 substances were retained after null and duplicate values were removed. Canonical SMILES were collected and described to generate an extended connectivity fingerprint, which was then used to predict the pEC₅₀ value. The machine learning-generated model was then used to predict 2846 phytochemical compounds from the KEGG database. This study uses the Light gradient-based machine learning method. The performance of the model was evaluated using Root Mean Square Error (RMSE) and R-square (R²) of 0.63 and 0.73, respectively. Afterward, the model was used to predict phytochemical compounds. Five compounds with the lowest EC₅₀ were obtained: Pedilstatin (60.05 nM), Gnididin (64.29 nM), Paclitaxel (98.93 nM), Vincristine (113.9 nM), and Camellidin II (129.72 nM). Further research regarding the direct effect of PPAR-gamma needs to be tested on these five compounds. Machine learning models will save time and resources for drug discovery or repurposing previously used drugs.

Keywords: drug discovery, insulin sensitivity, machine learning, PPAR-gamma, type-2 diabetes.

Model QSAR Berbasis Pembelajaran Mesin untuk Melakukan Skrining Stimulator PPAR-gamma

Abstrak

Diabetes tipe-2 (DM2) masih menjadi masalah secara global. Penurunan sensitivitas insulin merupakan mekanisme utama yang menjadi penyebab terjadinya DM2. PPAR-gamma telah terbukti memiliki hubungan langsung dengan sensitivitas insulin. Studi ini bertujuan mengembangkan model pembelajaran mesin untuk mengevaluasi aktivitas senyawa fitokimia sebagai agonis PPAR-gamma. Set data yang digunakan untuk pembuatan model diambil dari basis data ChEMBL. Sebanyak 3668 zat dipertahankan setelah nilai kosong dan duplikat dihilangkan. *Canonical SMILES* dikumpulkan dan dideskripsikan menghasilkan *extended connectivity fingerprint* dan digunakan untuk prediksi nilai pEC₅₀. Model yang dihasilkan pembelajaran mesin lalu digunakan untuk memprediksi 2846 senyawa fitokimia dari basis data KEGG. Studi ini menggunakan metode *Light gradient-based machine* untuk model pembelajaran mesin. Performa dari model dievaluasi menggunakan *Root Mean Square Error* (RMSE) dan *R-square* (R²) yaitu 0,63 dan 0,73. Selanjutnya, model tersebut digunakan untuk melakukan prediksi senyawa fitokimia. Akhirnya didapatkan lima senyawa dengan EC₅₀ terendah: Pedilstatin (60,05 nM), Gnididin (64,29 nM), Paclitaxel (98,93 nM), Vincristine (113,9 nM), Camellidin II (129,72 nM). Penelitian lanjut mengenai efek langsung terhadap PPAR-gamma perlu dilakukan terhadap kelima senyawa tersebut. Penggunaan model pembelajaran mesin akan menghemat waktu dan sumber daya untuk penemuan senyawa baru sebagai obat atau memungkinkan kita untuk melakukan alih fungsi dari obat yang telah digunakan sebelumnya.

Kata Kunci: penemuan obat, sensitivitas insulin, pembelajaran mesin, PPAR-gamma, diabetes tipe-2

Article History:

Submitted 19 May 2023

Revised 19 March 2024

Accepted 27 March 2024

Published 31 October 2025

*Corresponding author:
gaga@unpad.ac.id

Citation:

Nirwana, E.; Nugraha, G.I.; Ariyanto, E.F.; Sanjaya, A. Machine Learning Based QSAR Model for Screening PPAR-gamma Stimulator. Indonesian Journal of Pharmaceutical Science and Technology. 2025; 12 (3), 304-310.

1. Pendahuluan

Menurut sebuah studi menggunakan regresi logistik, populasi diabetes melitus (DM) global diprediksi berkembang pesat dari 536 juta pada tahun 2021 menjadi sekitar 783 juta pada tahun 2045. Laporan Riset Kesehatan Dasar (Riskesdas) tahun 2018 menunjukkan prevalensi DM di Indonesia meningkat sebesar 1,6%. Situasi ini memberikan beban sosial-ekonomi yang besar dalam kesehatan masyarakat, terutama pada negara berkembang seperti Indonesia.^{1,2} Diabetes melitus tipe-2 (DM2) adalah kelainan metabolisme yang disebabkan oleh kombinasi dua faktor utama: defek sekresi insulin oleh sel β pankreas dan ketidakmampuan jaringan untuk merespons insulin dengan baik. Faktor risiko untuk DM2 meliputi genetik, obesitas, aktivitas fisik rendah, gaya hidup menetap, dan diet tidak sehat. DM2 memiliki kaitan erat dengan insufisiensi dan resistensi insulin.^{3,4}

Insulin merupakan hormon hasil produksi sel- β pankreas yang memainkan peran utama dalam homeostasis glukosa karena merupakan satu-satunya hormon yang mampu menurunkan konsentrasi glukosa darah. Karena alasan inilah insufisiensi insulin mengakibatkan gangguan metabolisme terkait dengan DM2.⁵ Mekanisme yang terjadi pada DM2 adalah sekresi insulin tidak dapat lagi mengkompensasi terjadinya resistensi insulin.⁶ Sensitivitas insulin telah terbukti memiliki hubungan langsung dengan PPAR-gamma.⁷

Peroxisome proliferator-activated receptor (PPAR) adalah sensor lipid yang memodulasi metabolisme energi pada seluruh tubuh. PPAR-gamma adalah regulator utama adipogenesis, yang dapat meningkatkan penyimpanan lipid, sensitivitas insulin, dan metabolisme glukosa.^{8,9} Stimulasi PPAR-gamma oleh *Thiazolidinedione* (TZD) telah terbukti meningkatkan sensitivitas insulin perifer, di hati dan otot rangka, dan mendorong adipogenesis yang menyebabkan penurunan produksi glukosa endogen dan glukoneogenesis postprandial, penurunan glukosa puasa dan postprandial, dan penurunan kadar glukosa darah dan insulin.^{10,11}

Teknik pembelajaran mesin (*machine learning*) secara modern dapat digunakan untuk membuat model *Quantitative structure-activity relationship* (QSAR) dan mengembangkan program kecerdasan artifisial yang secara akurat dapat memprediksi secara *in silico* bagaimana modifikasi kimia dapat mempengaruhi perubahan biologis. Banyak karakteristik fisiokimia obat, seperti toksisitas, metabolisme, interaksi obat dan karsinogenesis telah dimodelkan secara efektif oleh teknik QSAR.¹² Beberapa penelitian menunjukkan bahwa agonis PPAR-gamma memiliki efek biologis

yang menguntungkan: peningkatan toleransi glukosa pada pasien dengan DM2, perbaikan glukosa dan profil lipid pada pasien dengan dislipidemia, penurunan tingkat keparahan penyakit *non-alcoholic fatty liver* pada tikus, perbaikan profil lipid dan penurunan resistensi insulin pada tikus obesitas.^{10,13-15}

Untuk mencari agonis PPAR-gamma dengan struktur kimia yang baru, dapat digunakan model QSAR berbasis pembelajaran mesin. Model pembelajaran mesin ini menggunakan deskriptor molekuler dua dimensi yang disebut *extended connectivity fingerprint* 6 (ECFP6). Deskriptor ini dipilih karena merupakan deskriptor dua dimensi terbaik yang ada dan telah sering digunakan oleh beberapa penelitian sebelumnya.¹⁶

Menggunakan fingerprint tersebut, model ini dilatih untuk memprediksi konsentrasi EC₅₀ mereka terhadap PPAR-gamma. Setelah model dibangun, dilakukan pengambilan molekul fitokimia dari basis data fitokimia *Kyoto Encyclopedia of Genes and Genomes* (KEGG) lalu kemudian prediksi dijalankan pada model tersebut dan melaporkan molekul-molekul yang memiliki potensi sebagai alternatif TZD.

2. Bahan dan Metode

2.1. Pengumpulan Data untuk Pembuatan Model

Untuk membuat prediksi ini, digunakan PPAR-gamma (ChEMBL ID: ChEMBL235) sebagai target dari sumber web ChEMBL.¹⁷ Hasil pencarian dibatasi pada data laporan konsentrasi EC₅₀ sebagai pengukuran standar aktivitas. *Canonical Simplified Molecular-input line-entry System* (SMILES) dari setiap zat lalu digunakan untuk menentukan *molecular fingerprints* zat tersebut agar dapat dibaca oleh perangkat lunak komputer. Penelitian ini menggunakan ECFP6, yaitu kelas baru dari topologi *fingerprint* yang dikembangkan untuk pencarian substruktur dan kesamaan karakteristik molekuler.¹⁸ *Features*, yaitu informasi data yang digunakan untuk melatih model pembelajaran mesin ini berasal dari ECFP6 yang disediakan oleh *Python Package RDKit*.¹⁹

Sebanyak 1024 bit data dibentuk mewakili *fingerprint* ECFP6 dan digunakan sebagai *features* saat pembuatan model. Selanjutnya seleksi dilakukan dengan cara menghilangkan *features* yang memiliki variansi nol atau berkorelasi tinggi satu sama lain ($r > 0.95$). Untuk memastikan bahwa tidak ada data EC₅₀ yang menghasilkan nilai negatif, data diperiksa dan dilakukan normalisasi nilai untuk konsentrasi di atas 108 nM. Kemudian set data dibagi dengan rasio 8:2 (*training* : *test*) untuk evaluasi akhir. Perangkat komputer yang digunakan pada penelitian ini memiliki

spesifikasi pada Tabel 1.

Kondisi dan konfigurasi lain yang diperlukan untuk mereplikasi penelitian disediakan pada material tambahan.

2.2. Validasi Metode

Model pembelajaran mesin ini berdasar pada model regresi, sehingga digunakan *Root Mean Square Error* (RMSE) dan *R-square* (R²) sebagai metrik untuk evaluasi. Strategi *K-fold cross-validation* juga digunakan untuk mengevaluasi model dengan sepuluh lipat *cross-validation*. Beberapa model dibandingkan menggunakan *Python's PyCaret package*,²⁰ dan model regresi yang memiliki RMSE terendah dan R² tertinggi akan digunakan. Dalam pembelajaran mesin, *hyperparameter* adalah parameter yang nilainya digunakan untuk konfigurasi proses pembuatan model. Model yang terpilih dilakukan penyetelan secara otomatis menggunakan *PyCaret package* untuk menghasilkan *R-square* maksimum yang dapat diperoleh saat *cross-validation* sebelum lalu digunakan untuk memprediksi *test set*.

2.3. Kumpulan Data untuk Kesimpulan

Zat yang diklasifikasikan sebagai fitokimia dari basis data senyawa KEGG dikumpulkan.^{20,21} Lalu kumpulan data tersebut melalui langkah-langkah tertentu dibentuk menjadi features menggunakan fingerprint ECFP6 sebelum kemudian dimasukkan ke model pembelajaran mesin. Hasilnya didapatkan beberapa zat dengan pEC₅₀ tinggi (berkaitan dengan EC₅₀ rendah).

3. Hasil

3.1. Proses Data

Pencarian pada basis data ChEMBL didapatkan 4866 zat dengan laporan nilai EC₅₀ terhadap PPAR- γ manusia. Dari data tersebut, dilakukan penyaringan

data yang kehilangan canonical SMILES dan EC₅₀ dan tersisa sejumlah 3688 zat. Canonical SMILES dikumpulkan dan dideskripsikan menggunakan RDKit untuk menghasilkan fingerprint ECFP6 sebagai features. Nilai EC₅₀ dikonversi menjadi pEC₅₀ untuk digunakan sebagai labels, yaitu subjek yang akan diprediksi. Total seluruhnya terbentuk 1024 bit fingerprint, dan 1024 features dipertahankan setelah variansi nol atau korelasi tinggi dihilangkan. Kemudian 1024 features dan pEC₅₀ labels yang dihasilkan diacak dan dibagi dengan rasio 8:2 untuk set training dan test. Alur kerja dapat dilihat pada Gambar 1.

3.2. Hasil Validasi Metode

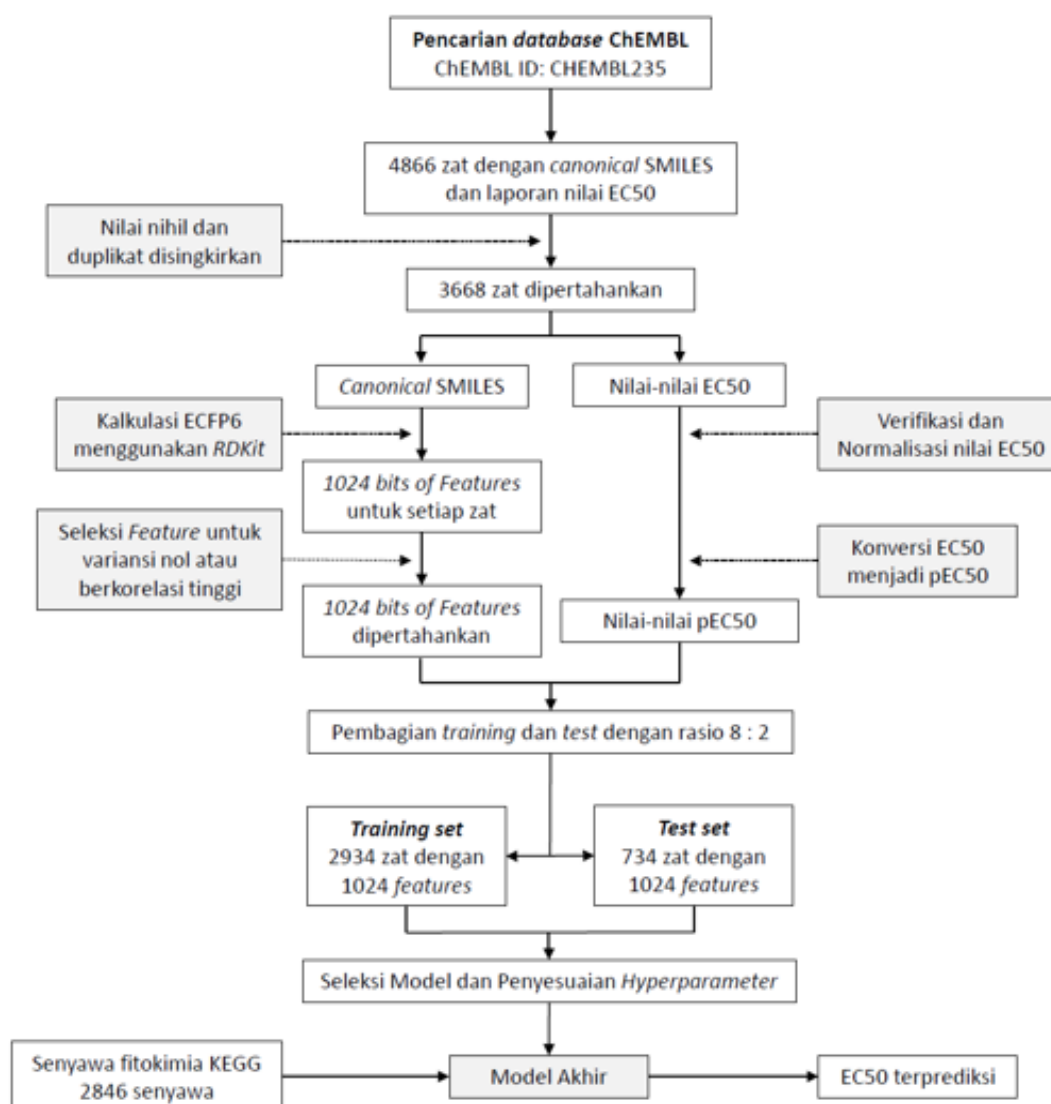
Model paling optimal dipilih dengan RMSE terendah dan R² tertinggi menggunakan data *training*. Studi ini menggunakan metode *Light gradient-based machine* dengan nilai RMSE 0,74 dan R² 0,66 sebelum dilakukan penyetelan. Metode tersebut juga digunakan sebagai algoritma pembelajaran mesin ini karena keakuratan dan kecepatannya. Penyesuaian model dilakukan menggunakan *PyCaret package* dan dilakukan pengulangan sebanyak 50 kali. Akurasi dari model uji lainnya dan pengaturan *hyperparameter* model akhir disediakan pada material tambahan. Data set *test* kemudian digunakan untuk memastikan kemampuan dari model pembelajaran mesin. Hasil yang didapatkan cukup akurat dengan RMSE 0,63 dan R² 0,73. Hasil evaluasi model dapat dilihat pada Gambar 2.

3.3. Prediksi Senyawa Fitokimia

Kesimpulan dibuat menggunakan model untuk memprediksi beberapa senyawa alami yang menstimulasi PPAR- γ . Total 2846 senyawa dari basis data senyawa KEGG digunakan dan selanjutnya diprediksi menggunakan model. Lima senyawa dengan EC₅₀ rendah yang ditemukan antara lain: Pedilstatin (60,05 nM), Gnididin (64,29 nM), Paclitaxel (98,93 nM), Vincristine (113,9 nM), Camellidin II (129,72 nM). Senyawa lainnya yang dengan pEC₅₀ terprediksi disediakan pada material tambahan.

Tabel 1. Spesifikasi Komputer untuk Studi *in silico*

Jenis Perangkat	Spesifikasi
Processor	AMD Ryzen 9 5900X 12-Core @ 3,70 GHz
Installed RAM	32.0 GB
Graphics Card	GeForce RTX 3070
System type	64-bit operating system, x64-based processor
Edition	Windows 10 Pro
Version	20H2
OS build	19042.1288
Experience	Windows Feature Experience Pack 120.2212.3920.0



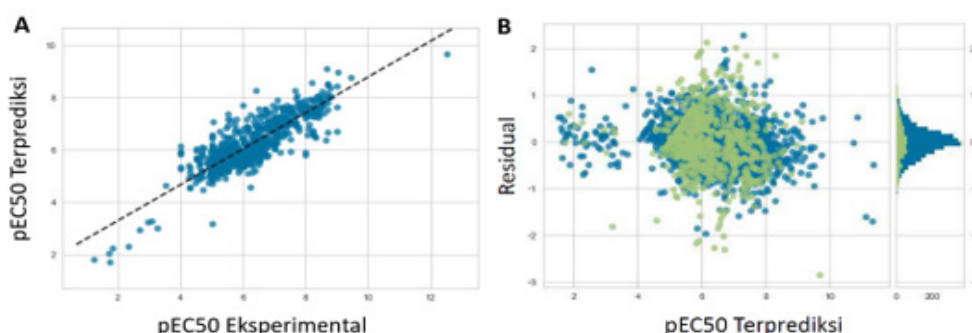
Gambar 1. Alur kerja penelitian yang digunakan dalam penelitian ini. Model difinalisasi dengan seluruh set data sebelum digunakan untuk memprediksi senyawa fitokimia KEGG.

4. Pembahasan

Sejauh pengetahuan penulis, artikel ini merupakan artikel pertama yang mendeskripsikan aktivitas dari senyawa fitokimia terhadap PPAR-gamma menggunakan pembelajaran mesin. Model ini berdasar pada metode *light gradient boosting machine*, yaitu sebuah metode gradient-boosted berbasis algoritma dengan struktur seperti pohon untuk membuat keputusan yang mirip dengan model *random forest* atau XGBoost yang umum diterapkan pada pembelajaran mesin.²² Setelah dilakukan penyetelan, model ini didapatkan cukup akurat dengan *R-square* 0,73 ketika divalidasi menggunakan data yang sebelumnya tidak tampak. Akurasi model untuk prediksi PPAR-gamma pada studi ini berada pada rerata nilai *R-square* yang dilaporkan dalam studi Quantitative structure–activity relationship lainnya.²³ Akurasi dari hasil prediksi oleh model ini dapat dipercaya karena berdasar pada

diagram residual tampak sebagian besar dari set data test, residual atau kesalahan pada prediksi untuk PPAR-gamma adalah <1 , dengan *root mean square error* 0,63. Model ini digunakan untuk memprediksi pEC₅₀ dari 2846 senyawa menggunakan set data dengan sumber dari basis data KEGG. Hasilnya didapatkan beberapa senyawa dengan pEC₅₀ cukup tinggi dari seluruh senyawa yang dilakukan pengujian. Lima senyawa dengan pEC₅₀ tertinggi yang memiliki cukup referensi akan dilakukan pembahasan.

Pedilstatin dengan EC₅₀ 60.05 nM merupakan senyawa pertama yang diprediksi memiliki pEC₅₀ tertinggi terhadap PPAR-gamma dari seluruh senyawa yang diuji. Pedilstatin merupakan senyawa yang diisolasi dari tumbuhan *Pedilanthus* sp. Pada studi yang lalu, senyawa ini ditemukan dapat menghambat pertumbuhan lini sel leukemia limfositik P388 secara signifikan dengan ED₅₀ 0.28 µg/mL



Gambar 2. Evaluasi model menggunakan set data test. **A.** Diagram pencar menunjukkan nilai pEC50 yang terprediksi terhadap nilai pEC50 tervalidasi oleh set data yang ditunjukkan oleh titik biru. **B.** Diagram residual menunjukkan kesalahan prediksi dari model. Titik biru menggambarkan kesalahan prediksi dari set data training. Titik hijau menggambarkan kesalahan prediksi dari kumpulan set data test yang sebelumnya tidak tampak.

untuk mencapai konsentrasi 2-5 μM . Selain itu juga, Pedilstatin menunjukkan efek proteksi hingga 80% dari infeksi terhadap sel turunan manusia CEM-SS dan pembunuhan sel oleh HIV-1. Pedilstatin bekerja sebagai ligan untuk protein kinase C.²⁴ Terlepas dari banyaknya informasi mengenai mekanisme yang mengontrol aktivasi dan fungsi PKC dalam model seluler, relevansi PKC dalam perkembangan kanker manusia masih menjadi kontroversi.²⁵ Selain itu, belum ditemukan studi mengenai efek pedilstatin terhadap PPAR-gamma.

Senyawa lainnya yang ditemukan yaitu Gnididin yang diprediksi memiliki aktivitas biologis terhadap PPAR-gamma dengan EC_{50} 64.29 nM. Gnididin berasal dari *Gnidia lamprantha* Gilg, merupakan ester diterpenoid anggota pertama dari kelas langka daphnetoxin yang telah terbukti memiliki aktivitas anti-tumor. Pada sebuah studi, ekstrak etanol *Gnidia lamprantha* Gilg menunjukkan aktivitas anti-leukemik poten secara *in vivo* terhadap P-388 leukemia pada tikus.²⁶ Gugus ester diduga berfungsi sebagai gugus pembawa dalam proses yang berkaitan dengan penetrasi sel atau pembentukan kompleks molekul selektif.²⁷ Saat ini belum didapatkan studi mengenai efek gnididin terhadap PPAR-gamma.

Paclitaxel (Taxol) dengan EC_{50} 98.93 nM adalah senyawa yang berasal dari kulit batang yew barat *Taxus brevifolia*. Taxol ditemukan sebagai obat anti-kanker yang memiliki aktivitas sitotoksik terhadap banyak jenis tumor, misalnya kanker ovarium dan payudara.²⁸ Pada sebuah studi didapatkan paclitaxel bekerja dengan cara meningkatkan pemberhentian siklus sel G2/M dan apoptosis, hal ini juga didapatkan pada studi lain dengan lini sel sarkoma dan adenokarsinoma payudara.²⁹ Akan tetapi, studi mengenai efek paclitaxel terhadap PPAR-gamma belum dapat ditemukan.

Vincristine dengan EC_{50} 113.9 nM awalnya ditemukan pada program skrining untuk mencari potensi anti-

diabetes dari ekstrak tanaman *Cantharanthus roseus* (sebelumnya dikenal sebagai *Vinca rosea* Linn). Meskipun tidak efektif sebagai pengobatan antidiabetes oral, ekstrak ini memiliki aktivitas antileukemia yang signifikan dengan menghambat produksi dan pematangan leukosit.³⁰⁻³² Pada sebuah studi didapatkan efek sinergis anti-proliferasi rosiglitazone dan vincristine terkait dengan penghambatan polimerisasi tubulin, penghentian siklus sel dalam fase G2/M dan induksi apoptosis sel. Kombinasi tersebut meregulasi kelangsungan jalur PTEN/PI3K/AKT dengan meningkatkan regulasi PTEN dan menurunkan regulasi p-AKT.³³ Saat ini juga belum ditemukan studi mengenai efek vincristine terhadap PPAR-gamma secara langsung.

Camellidin II merupakan triterpenoid saponin yang diisolasi dari air atau ekstrak metanol daun kamelia (*Camellia japonica*) dengan EC_{50} 129.72 nM. Dalam sebuah studi, camellidin II dengan formula $\text{C}_{53}\text{H}_{84}\text{O}_{24}$ memiliki aktivitas antifungal. Mekanisme kerja camellidin II belum diketahui secara lengkap, akan tetapi asetilasi dari gugus 18 β -hidroksil meningkatkan aktivitas antifungal pada camellidin.³⁴ Mengenai efek langsung camellidin II terhadap PPAR-gamma masih belum diketahui.

Model yang dihasilkan penelitian ini mampu mendapatkan nilai R^2 yang sebanding dan bahkan lebih tinggi dari beberapa model QSAR yang telah dipublikasikan.³⁵ Oleh karena itu dapat disimpulkan bahwa model ini mampu menghasilkan prediksi yang akurat. Pada sebuah studi *in vitro*, Rosiglitazone yang merupakan TZD memiliki nilai EC_{50} 80 nM terhadap PPAR-gamma.³⁶ Dari lima senyawa di atas yang telah dibahas, Pedilstatin dan Gnididin secara teori memiliki potensi lebih kuat dari Rosiglitazone sebagai agonis PPAR-gamma. Akan tetapi, sebagian besar mekanisme kerja dari senyawa tersebut masih belum diteliti terhadap PPAR-gamma dan diperlukan penelitian lanjutan untuk mengkonfirmasi hasil yang

didapat dari model pembelajaran mesin. Selain itu, dari referensi yang didapatkan, ternyata kelima senyawa ini sebagian besar memiliki manfaat sebagai anti kanker. Beberapa penelitian mengungkapkan bahwa aktivasi PPAR-gamma oleh agonisnya menyebabkan penghentian siklus sel³⁷, apoptosis³⁸, angiogenesis³⁹, inhibisi⁴⁰, dan rediferensiasi⁴¹, yang merupakan kunci dari proses molekuler terkait dengan pencegahan pertumbuhan dan progresivitas tumor.⁴² Eksperimen *in vivo* menunjukkan bahwa PPAR-gamma berinteraksi dengan jalur pensinyalan hilir dari aksis *insulin-like growth factor* (IGF) yang melibatkan kaskade *mitogen-activated protein kinase* (MAPK), *phosphatidylinositol 3-kinase* (PI3K), dan *mammalian target of rapamycin* (mTOR).^{43,44}

Potensi dari senyawa tersebut menarik untuk diteliti lebih lanjut, yaitu efek serta mekanisme kerja senyawa tersebut terhadap PPAR-gamma dan juga korelasi antara PPAR-gamma agonis dengan berbagai macam sel kanker termasuk leukemia.

5. Simpulan

Berdasarkan penelitian ini, dapat disimpulkan bahwa akurasi dari hasil prediksi oleh model ini dapat dipercaya berdasar statistik yang tampak pada data dalam percobaan ini. Dari percobaan ini, setidaknya yaitu Pedilstatin, Gnididin, Paclitaxel, Vincristine, dan Camellidin II ditemukan dengan potensi sebagai agonis PPAR-gamma. Penggunaan model ini akan dapat membantu menghemat waktu dan sumber daya untuk penemuan senyawa baru sebagai obat atau memungkinkan kita untuk melakukan alih fungsi dari obat yang telah digunakan sebelumnya. Selanjutnya, model pembelajaran mesin ini akan lebih baik apabila dilanjutkan dengan studi *in silico* seperti *virtual screening* seperti *molecular docking*, *molecular dynamic*, *screening pharmacophore*, dan penelitian secara *in vitro* serta *in vivo*.

Konflik Kepentingan

Penulis menyatakan bahwa data yang dipublikasikan pada naskah ini tidak ada konflik kepentingan terhadap pihak manapun

Daftar Pustaka

1. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res Clin Pract*. 2022 Jan;183:109119.
2. Kemenkes RI. Hasil Riset Kesehatan Dasar Tahun 2018. Kementerian Kesehatan RI. 2018;53(9):1689–99.
3. Galicia-Garcia U, Benito-Vicente A, Jebari S, Larrea-Sebal A, Siddiqi H, Uribe KB, et al. Pathophysiology of Type 2 Diabetes Mellitus. *Int J Mol Sci*. 2020 Aug 30;21(17):6275.
4. Czech MP. Insulin action and resistance in obesity and type 2 diabetes. *Nat Med*. 2017 Jul 1;23(7):804–14.
5. Rorsman P, Ashcroft FM. PANCREATIC-CELL ELECTRICAL ACTIVITY AND INSULIN SECRETION: OF MICE AND MEN. *Physiol Rev* [Internet]. 2018;98:117–214. Available from: www.prv.org
6. Yang Q, Vijayakumar A, Kahn BB. Metabolites as regulators of insulin sensitivity and metabolism. Vol. 19, *Nature Reviews Molecular Cell Biology*. Nature Publishing Group; 2018. p. 654–72.
7. Gharge S, Alegaon SG, Ranade SD, Kavalapure RS, Prashantha Kumar BR, Mhaske PC. Expression of PPAR- γ TF by newly synthesized thiazolidine-2,4-diones to manage glycemic control: Insights from *in silico*, *in vitro* and experimental pharmacology in wistar rats. *Bioorg Chem*. 2024 Dec;153:107966.
8. Moutagne D, Butruille L, Staels B. PPAR control of metabolism and cardiovascular functions. *Nat Rev Cardiol*. 2021 Dec 14;18(12):809–23.
9. Jain N, Bhansali S, Kurpad A V., Hawkins M, Sharma A, Kaur S, et al. Effect of a Dual PPAR α/γ agonist on Insulin Sensitivity in Patients of Type 2 Diabetes with Hypertriglyceridemia- Randomized double-blind placebo-controlled trial. *Sci Rep*. 2019 Dec 12;9(1):19017.
10. Basak S, Murmu A, Matore BW, Roy PP, Singh J. Thiazolidinedione an auspicious scaffold as PPAR- γ agonist: its possible mechanism to Manoeuvre against insulin resistant diabetes mellitus. *European Journal of Medicinal Chemistry Reports*. 2024 Aug;11:100160.
11. Davidson MA, Mattison DR, Azoulay L, Krewski D. Thiazolidinedione drugs in the treatment of type 2 diabetes mellitus: past, present and future. Vol. 48, *Critical Reviews in Toxicology*. Taylor and Francis Ltd; 2018. p. 52–108.
12. Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M, et al. QSAR Modeling: Where Have You Been? Where Are You Going To? *J Med Chem*. 2014 Jun 26;57(12):4977–5010.
13. Rastogi A, Dunbar RL, Thacker HP, Bhatt J, Parmar K, Parmar D v. Abrogation of postprandial triglyceridemia with dual PPAR α/γ agonist in type 2 diabetes mellitus: a randomized, placebo-controlled study. *Acta Diabetol*. 2020 Jul 6;57(7):809–18.
14. Lefere S, Puengel T, Hundertmark J, Penners C, Frank AK, Guillot A, et al. Differential effects of selective- and pan-PPAR agonists on experimental steatohepatitis and hepatic macrophages-. *J Hepatol*. 2020 Oct;73(4):757–70.
15. Chen T, Zhang Y, Liu Y, Zhu D, Yu J, Li G, et al. MiR-27a promotes insulin resistance and mediates glucose metabolism by targeting PPAR- γ -mediated PI3K/AKT signaling. *Aging*. 2019 Sep 28;11(18):7510–24.
16. Carracedo-Reboredo P, Liñares-Blanco J, Rodríguez-Fernández N, Cedrón F, Novoa FJ, Carballal A, et al. A review on machine learning approaches and trends in drug discovery. *Comput Struct Biotechnol J*. 2021;19:4538–58.
17. Davies M, Nowotka M, Papadatos G, Dedman N, Gaulton A, Atkinson F, et al. ChEMBL web services: Streamlining access to drug discovery data and utilities.

- Nucleic Acids Res. 2015;43(W1):W612–20.
18. Rogers D, Hahn M. Extended-Connectivity Fingerprints. *J Chem Inf Model*. 2010 May 24;50(5):742–54.
 19. Landrum G, Tosco P, Kelley B, Ric, sriniker, gedeck, et al. rdkit/rdkit: 2021_03_5 (Q1 2021) Release. 2021 Aug 24 [cited 2023 Feb 6]; Available from: <https://zenodo.org/record/5242603>
 20. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res*. 2016;44(D1):D457–62.
 21. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes [Internet]. Vol. 28, *Nucleic Acids Research*. 2000. Available from: <http://www.genome.ad.jp/kegg/>
 22. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree [Internet]. Available from: <https://github.com/Microsoft/LightGBM>.
 23. Wu Z, Zhu M, Kang Y, Leung ELH, Lei T, Shen C, et al. Do we need different machine learning algorithms for QSAR modeling? A comprehensive assessment of 16 machine learning algorithms on 14 QSAR data sets. *Brief Bioinform*. 2021 Jul 1;22(4).
 24. Pettit GR, Ducki S, Tan R, Gardella RS, McMahon JB, Boyd MR, et al. Isolation and structure of pedilstatin from a Republic of Maldives *Pedilanthus* sp. *J Nat Prod*. 2002 Sep 1;65(9):1262–5.
 25. Isakov N. Protein kinase C (PKC) isoforms in cancer, tumor promotion and tumor suppression. *Semin Cancer Biol*. 2018 Feb;48:36–52.
 26. Kupchan SM, Sweeny JG, Baxter RL, Murae T, Zimmerly VA, Sickles BR. Tumor inhibitors. 102. Gnididin, gniditrin, gnidicin, novel potent antileukemic diterpenoid esters from *Gnidia lamprantha*. *J Am Chem Soc*. 1975 Feb 1;97(3):672–3.
 27. Yoshimitsu Nagao, Masahito Ochiai. Design and Study of New Tumor Inhibitors. *Bull Inst Chem Res Kyoto Univ* [Internet]. 1985 Aug 24 [cited 2023 Apr 22];63(2):132–55. Available from: <http://hdl.handle.net/2433/77099>
 28. Kuriakose GC, Arathi BP, Divya Lakshmanan M, Jiby M V., Gudde RS, Jayabhaskaran C. Sub-acute Toxicity Assessment of Taxol Isolated From *Fusarium Solani*, an Endophytic Fungus of *Taxus Brevifolia*, in Wistar Rats and Analyzing Its Cytotoxicity and Apoptotic Potential in Lung Cancer Cells. *Front Oncol*. 2020 Oct 6;10.
 29. Abu Samaan TM, Samec M, Liskova A, Kubatka P, Büsselberg D. Paclitaxel's Mechanistic and Clinical Effects on Breast Cancer. *Biomolecules*. 2019 Nov 27;9(12):789.
 30. Silverman JA, Deitcher SR. Marqibo® (vincristine sulfate liposome injection) improves the pharmacokinetics and pharmacodynamics of vincristine. *Cancer Chemother Pharmacol*. 2013 Mar 5;71(3):555–64.
 31. Škubník J, Pavlíčková VS, Ruml T, Rimpelová S. Vincristine in Combination Therapy of Cancer: Emerging Trends in Clinics. *Biology (Basel)*. 2021 Aug 31;10(9):849.
 32. Jin X, Liu K, Jiao B, Wang X, Huang S, Ren W, et al. Vincristine promotes migration and invasion of colorectal cancer HCT116 cells through RhoA/ROCK/ Myosin light chain pathway. *Cell Mol Biol (Noisy-le-grand)*. 2016 Oct 31;62(12):91–6.
 33. Wang HY, Zhang Y, Zhou Y, Lu YY, Wang WF, Xin M, et al. Rosiglitazone elevates sensitization of drug-resistant oral epidermoid carcinoma cells to vincristine by G2/M-phase arrest, independent of PPAR-γ pathway. *Biomedicine and Pharmacotherapy*. 2016 Oct 1;83:349–61.
 34. NAGATA T, TSUSHIDA T, HAMAYA E, ENOKI N, MANABE S, NISHINO C. Camellidins, antifungal saponins isolated from *Camellia japonica*. *Agric Biol Chem*. 1985;49(4):1181–6.
 35. Qin Z, Yan A. QSAR studies on hepatitis C virus NS5A protein tetracyclic inhibitors in wild type and mutants by CoMFA and CoMSIA. *SAR QSAR Environ Res*. 2020 Apr 2;31(4):281–311.
 36. Huiying Z, Guangying C, Shiyang Z. Design, synthesis and biological activity evaluation of a new class of 2,4-thiazolidinedione compounds as insulin enhancers. *J Enzyme Inhib Med Chem*. 2019 Jan 1;34(1):981–9.
 37. Augimeri G, Bonofiglio D. PPARgamma: A Potential Intrinsic and Extrinsic Molecular Target for Breast Cancer Therapy. *Biomedicines*. 2021 May 13;9(5):543.
 38. Wagner N, Wagner KD. Peroxisome Proliferator-Activated Receptors and the Hallmarks of Cancer. *Cells*. 2022 Aug 5;11(15):2432.
 39. Wagner N, Wagner KD. PPARs and Angiogenesis—Implications in Pathology. *Int J Mol Sci*. 2020 Aug 10;21(16):5723.
 40. Basilotta R, Lanza M, Casili G, Chisari G, Munao S, Colarossi L, et al. Potential Therapeutic Effects of PPAR Ligands in Glioblastoma. *Cells*. 2022 Feb 10;11(4):621.
 41. Cheng HS, Yip YS, Lim EKY, Wahli W, Tan NS. PPARs and Tumor Microenvironment: The Emerging Roles of the Metabolic Master Regulators in Tumor Stromal–Epithelial Crosstalk and Carcinogenesis. *Cancers (Basel)*. 2021 Apr 29;13(9):2153.
 42. Ballav S, Biswas B, Sahu VK, Ranjan A, Basu S. PPAR-γ Partial Agonists in Disease-Fate Decision with Special Reference to Cancer. Vol. 11, *Cells*. MDPI; 2022.
 43. Vella V, Nicolosi ML, Giuliano S, Bellomo M, Belfiore A, Malaguarnera R. PPAR-γ Agonists As Antineoplastic Agents in Cancers with Dysregulated IGF Axis. *Front Endocrinol (Lausanne)*. 2017 Feb 22;8.
 44. Chi T, Wang M, Wang X, Yang K, Xie F, Liao Z, et al. PPAR-γ Modulators as Current and Potential Cancer Treatments. *Front Oncol*. 2021 Sep 23;11.