

## KLASIFIKASI KATEGORI OBAT MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE

### DRUG CATEGORY CLASSIFICATION USING SUPPORT VECTOR MACHINE ALGORITHM

**Mardewi<sup>1</sup>**

STMIK Kreatindo

Manokwari<sup>1</sup>

email:

[mardewi0004@gmail.com](mailto:mardewi0004@gmail.com)

**Nuru Yarkuran<sup>2</sup>**

Institut Teknologi dan

Bisnis Muhammadiyah

Wakatobi<sup>2</sup>

email:

[nuruyarkuran01@gmail.com](mailto:nuruyarkuran01@gmail.com)

**Sofyan<sup>3</sup>**

STMIK Kreatindo

Manokwari<sup>3</sup>

email:

[sofyanarifin018@gmail.com](mailto:sofyanarifin018@gmail.com)

**Firman Aziz<sup>4</sup>**

Universitas Pancasakti

Makassar<sup>4</sup>

email:

[firman.aziz@unpacti.ac.id](mailto:firman.aziz@unpacti.ac.id)

**Abstrak:** Dalam mencegah penyakit dan menjaga Kesehatan, setiap orang biasanya mengomsumsi obat. Namun banyak yang belum mengetahui kategori/jenis obat sebelum mengomsumsi obat tersebut. Masyarakat sebaiknya mengetahui kategori/jenis obat sebelum mengonsumsinya. Hal ini akan menimbulkan dampak negatif bagi kesehatan jika obat tersebut tergolong kedalam kategori/jenis obat keras. maka diperlukan klasifikasi kategori/jenis obat untuk membantu masyarakat agar dapat mendeteksi kategori/jenis obat. Penelitian ini mengusulkan untuk mengklasifikasi kategori/jenis obat menggunakan algoritma Support Vector Machine dengan tiga kernel yaitu linear, polynomial, dan RBF. Data terdiri dari lima variabel prediktor yaitu Usia, Jenis Kelamin, Tingkat Tekanan Darah, Tingkat Kolesterol, dan Na to K. jumlah data sebanyak dua ratus record. Hasil menunjukkan bahwa akurasi terbaik dari ketiga kernel SVM didapatkan ketika menggunakan kernel linear dan polinomial. berbeda dengan kernel RBF yang memiliki akurasi yang lebih rendah dari kedua kernel tersebut. kernel linear dan polinomial menghasilkan akurasi sebesar 95.0% sedangkan RBF menghasilkan akurasi sebesar 94.5%.

**Kata Kunci:** klasifikasi, kategori obat, SVM, linear, polynomial, RBF.

**Abstract:** In preventing disease and maintaining health, everyone usually takes medicine. However, many do not know the category/type of drug before taking the drug. Society should know the category/type of drug before consuming it. This will have a negative impact on health if the drug belongs to the category/type of hard drug. it is necessary to classify categories/types of drugs to help the public to be able detect categories/types of drugs. This study proposes to classify categories/types of drugs using the Support Vector Machine algorithm with three kernels, namely linear, polynomial, and RBF. The data consisted of five predictor variables, namely Age, Gender, Blood Pressure Level, Cholesterol Level, and Na to K. The total data was two hundred records. The results show that the best accuracy of the three SVM kernels is obtained when using linear and polynomial kernels. different from the RBF kernel which has lower accuracy than the two kernels. the linear and polynomial kernels produce an accuracy of 95.0% while the RBF produces an accuracy of 94.5%.

**Keywords:** classification, drug category, SVM, linear, polynomial, RBF.

JOPACS

E-ISSN: 2985-8593

Vol. 1, No. 1, Februari,

2023

## PENDAHULUAN

Untuk mencegah penyakit serta menjaga Kesehatan setiap orang biasanya mengonsumsi obat yang dijual di warung, toko, dan apotik. Namun masih banyak yang belum mengetahui kategori/jenis obat. Masyarakat sebaiknya mengetahui kategori/jenis obat secara nasional dan internasional (Nurmala et al., n.d.; Sidoretno et al., 2018). Terdapat beberapa kategori/jenis obat seperti obat herbal, obat keras, obat terbatas, obat bebas, obat bius/narkotika, obat tradisional dan lainnya (Obat & 2022, n.d.).

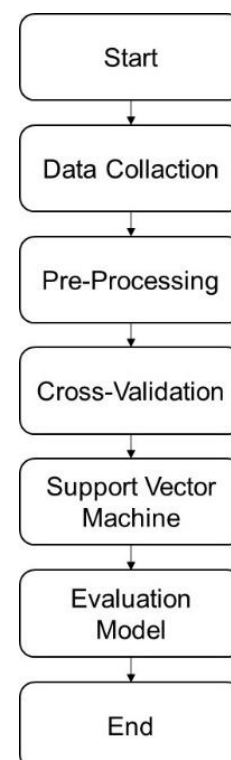
Secara umum, masyarakat kurang memperhatikan kategori/jenis obat sebelum mengonsumsinya. Hal ini akan menimbulkan dampak negatif bagi kesehatan jika obat tersebut tergolong kedalam kategori/jenis obat keras yang memerlukan resep dari dokter (Jenis et al., 2016). maka diperlukan klasifikasi kategori/jenis obat untuk membantu masyarakat agar dapat mendeteksi kategori/jenis obat.

Klasifikasi merupakan satu diantara beberapa teknik data mining untuk melakukan prediksi, dimana prediksi nilai berupa label (variable target). Terdapat beberapa penelitian terkait klasifikasi kategori/jenis obat. Diantaranya penelitian (Nursyahfitri et al., n.d.) yang mengusulkan decision tree untuk mengklasifikasi jenis obat. Hasil menunjukkan bahwa kinerja algoritma decision tree dinilai sangat baik dalam memprediksi jenis obat. Penelitian (Jenis et al., 2016) mengusulkan

metode naïve bayes untuk mengklasifikasi jenis obat berdasarkan gambar logo pada kemasan. Hasil menunjukkan bahwa beberapa klasifikasi menghasilkan akurasi yang kurang tepat karena disebabkan gambar logo yang terlalu rumit atau gambar yang kurang jelas.

Berdasarkan permasalahan diatas, kami mengusulkan untuk mengklasifikasi kategori/jenis obat menggunakan algoritma Support Vector Machine. Penerapan algoritma Support Vector Machine karena efektif dalam mengklasifikasi kelas dengan mengukur hyperplane terbaik dan membagi kedalam dua kelas (Indra Borman et al., n.d.).

## METODE



Gambar 1. Tahapan Penelitian

#### A. Data Collection

Data Collection merupakan suatu kegiatan yaitu mengumpulkan data yang digunakan (people & 2009, n.d.). Pada paper ini, data didapatkan melalui situs uci repository dataset. Data yang digunakan dalam penelitian ini yaitu data jenis obat.

#### B. Preprocessing

Data Preprocessing merupakan kegiatan untuk mempersiapkan data sebelum diproses menggunakan algoritma klasifikasi (Breiman, 2001; García et al., 2016). Tahapannya adalah cleaning data, transformasi data dan normalisasi data.

#### C. K-Fold

Cross-validasi merupakan sebuah metode validasi model untuk menilai bagaimana hasil dari sebuah statistik analisis akan menyamaratakan kumpulan data independent (Wager, 2019).

#### D. Support Vector Machine

Merupakan teknik yang melakukan prediksi untuk klasifikasi maupun regresi (Neneng et al., 2020; Rizal et al., 2019). prinsip dasar klasifikasi Support Vector Machine yaitu secara linier memisahkan, namun SVM telah berkembang sehingga dapat bekerja pada problem non-linier dengan mengoptimasi fungsi kernel dengan hyperplane (hyperplane) yang dapat memaksimalkan jarak (margin) antara kelas data. Hyperplane SVM di notasikan:

$$f(x) = w^2 x + b \quad (1)$$

#### E. Performance Evaluation

untuk menghitung nilai kesalahan metode klasifikasi digunakan matriks konfusi agar kinerja metode klasifikasi dapat dievaluasi seperti pada tabel 1 berikut (Provost et al., n.d.).

Tabel 1 Confusion Matrix

Actual	Prediction				
	Obat A	Obat B	Obat C	Obat X	Obat Y
Obat A	T <sub>0</sub>	F <sub>01</sub>	F <sub>02</sub>	F <sub>03</sub>	F <sub>04</sub>
Obat B	F <sub>10</sub>	T <sub>1</sub>	F <sub>12</sub>	F <sub>13</sub>	F <sub>14</sub>
Obat C	F <sub>20</sub>	F <sub>21</sub>	T <sub>2</sub>	F <sub>23</sub>	F <sub>24</sub>
Obat X	F <sub>30</sub>	F <sub>31</sub>	F <sub>32</sub>	T <sub>3</sub>	F <sub>34</sub>
Obat Y	F <sub>40</sub>	F <sub>41</sub>	F <sub>42</sub>	F <sub>43</sub>	T <sub>4</sub>

Akurasi menunjukkan seberapa dekat hasil pengukuran dengan nilai sebenarnya. Menghitung akurasi menggunakan Persamaan 2 berikut:

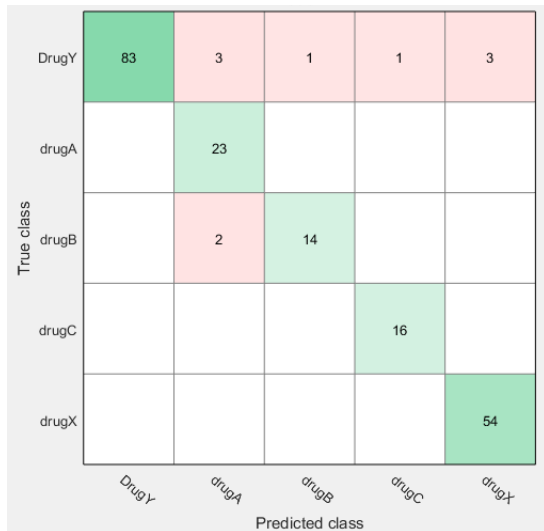
$$accuracy = \frac{\sum_j T_i}{N} \quad (2)$$

### HASIL DAN DISKUSI

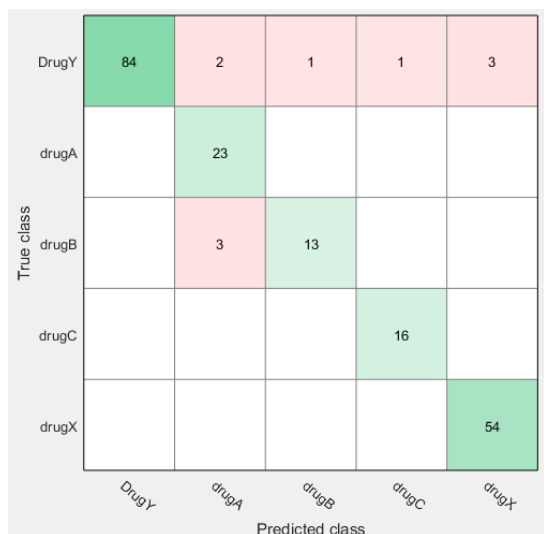
Klasifikasi dibagi kedalam lima kategori/jenis obat yang akan dijadikan data yaitu obat A, obat B, obat C, obat X dan obat Y. Data yang digunakan merupakan data benchmarking yang diakses melalui laman <https://www.kaggle.com/prathamtripathi/drug-classification>. Data terdiri dari lima variabel prediktor yaitu Usia, Jenis Kelamin, Tingkat

Tekanan Darah, Tingkat Kolesterol, dan Na to K. jumlah data sebanyak dua ratus record.

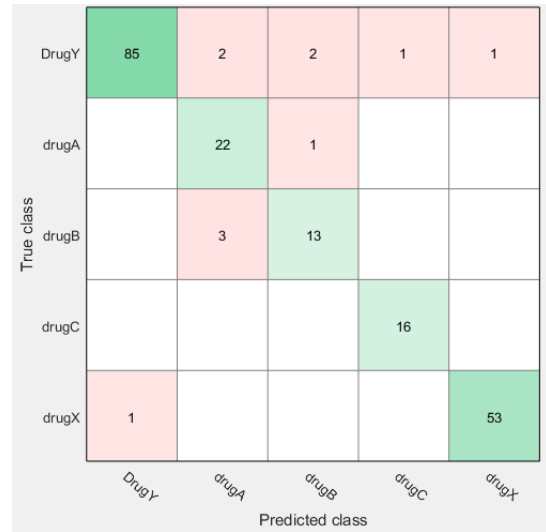
Kemudian data kategori/jenis obat akan diproses cross-validation untuk mendapatkan data pelatihan dan data pengujian. Proses klasifikasi menggunakan algoritma Support Vector Machine. Untuk mendapatkan hasil klasifikasi pada penelitian ini digunakan matlab.



Gambar 2. Hasil Confusion Matrix SVM Kernel Linear



Gambar 3. Hasil Confusion Matrix SVM Kernel Polinomial

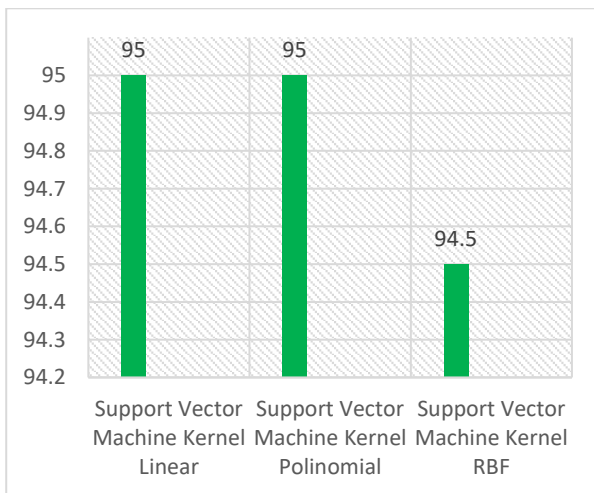


Gambar 4. Hasil Confusion Matrix SVM Kernel RBF

Berdasarkan hasil confusion matrix dari Support Vector Machine dengan tiga fungsi kernel seperti terlihat pada gambar 2 sampai 4, dilakukan perhitungan kinerja menggunakan persamaan 3. Hasil perhitungan kinerja Support Vector Machine dengan tiga fungsi kernel dapat dilihat pada tabel 2 berikut:

Tabel 2 Hasil Klasifikasi

Classification	Accuracy
Support Vector Machine Kernel Linear	95.0 %
Support Vector Machine Kernel Polinomial	95.0 %
Support Vector Machine Kernel RBF	94.5%



Gambar 4. Hasil Confusion Matrix SVM Kernel RBF

Berdasarkan tabel 2, akurasi terbaik dari ketiga kernel SVM didapatkan ketika menggunakan kernel linear dan polinomial. berbeda dengan kernel RBF yang memiliki akurasi yang lebih rendah dari kedua kernel tersebut. kernel linear dan polinomial menghasilkan akurasi sebesar 95.0% sedangkan RBF menghasilkan akurasi sebesar 94.5%. Berdasarkan hasil yang diperoleh, ketiga kernel SVM berhasil mengklasifikasi kategori/jenis obat dengan performa yang baik. hasil performa ini, didukung dengan jumlah data melimpah.

## KESIMPULAN

Kesuksesan pengusulan algoritma Support Vector Machine pada penelitian ini, diukur berdasarkan nilai akurasi. berdasarkan hasil yang didapatkan, jumlah data mempengaruhi kinerja dari algoritma Support Vector Machine. dari ketiga kernel SVM, yang memiliki kinerja terbaik adalah kernel linear dan polinomial

menghasilkan akurasi sebesar 95.0% sedangkan RBF menghasilkan akurasi sebesar 94.5%. tetapi penelitian ini menemukan permasalahan ketika nilai cross-validation semakin besar mengakibatkan performa menurun.

## REFERENSI

- Breiman, L. (2001). Random Forests. *Machine Learning* 2001 45:1, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M., & Herrera, F. (2016). Big data preprocessing: methods and prospects. *Big Data Analytics*, 1(1). <https://doi.org/10.1186/S41044-016-0014-0>
- Indra Borman, R., Aizzuddin Abd Rahni, A., Rossi, F., Dwi Putra, S., Jusman, Y., & Herdiansah, A. (n.d.). Identification of Herbal Leaf Types Based on Their Image Using First Order Feature Extraction and Multiclass SVM Algorithm. *Ieeexplore.Ieee.Org*. Retrieved February 3, 2023.
- Jenis, I., Berdasarkan, O., Logo, G., Kemasan, P., Metode, M., Bayes, N., Gede, I., & Rahayuda, S. (2016). Identifikasi Jenis Obat Berdasarkan Gambar Logo Pada Kemasan Menggunakan Metode Naive Bayes. *Scholar.Archive.Org*, 06(01).
- Neneng, N., Putri, N., CYBERNETICS, E. S.-, & 2021, undefined. (2020). Klasifikasi Jenis Kayu Menggunakan Support Vector Machine Berdasarkan Ciri Tekstur Local Binary Pattern. *Academia.Edu*, 4(02), 93–100.
- Nurmala, S., Farmasi, D. G.-F. J. I., & 2020, undefined. (n.d.). Pengetahuan penggunaan obat antibiotik pada masyarakat yang tinggal di kelurahan Babakan Madang. *Scholar.Archive.Org*. <https://doi.org/10.33751/jf.v10i1.1728>
- Nursyahfitri, R., Novebrian Maharadja, A., Farissa, R. A., Umaidah, Y., Informatika, T., Komputer, I., & Karawang, S. (n.d.).

Klasifikasi Penentuan Jenis Obat Menggunakan Algoritma Decision Tree. *Jip.Polinema.Ac.Id*. Retrieved February 3, 2023.

Obat, O. M.-P., & 2022, undefined. (n.d.). BAB V PSIKOTROPIKA. *Books.Google.Com*. Retrieved February 3, 2023.

people, M. G.-R. with children and young, & 2009, undefined. (n.d.). Data collection and analysis. *Books.Google.Com*. Retrieved January 29, 2023.

Provost, F., learning, R. K.-M., & 1998, undefined. (n.d.). Guest editors' introduction: On applied research in machine learning. *Academia.Edu*. Retrieved March 8, 2022.

Rizal, R. achmad, ... I. G.-R. R. dan E.-J., & 2019, undefined. (2019). Klasifikasi Wajah Menggunakan Support Vector Machine (SVM). *Polgan.Ac.Id*, 3(2).

Sidoretno, W., Masyarakat, I. R.-J. P., & 2018, undefined. (2018). Edukasi Bahaya Bahan Kimia Obat Yang Terdapat Didalam Obat Tradisional. *Jurnal.Univrab.Ac.Id*, 1(2), 117–123.  
<https://doi.org/10.36341/jpm.v1i2.453>

Wager, S. (2019). *Cross-Validation, Risk Estimation, and Model Selection*.