

Komparasi Algoritma Random Forest Dan Decision Tree Untuk Memprediksi Keberhasilan Immunotherapy

Fandi Yulian Pamuji¹, Viry Puspaning Ramadhan²

^{1,2} Fakultas Teknik Informatika, Universitas Dian Nuswantoro Semarang

Info Artikel

Riwayat Artikel

Diterima: 27-06-2021

Direvisi: 01-07-2021

Disetujui: 05-07-2021

Kata Kunci

Data Mining;

Penyakit Kutil;

Immunotherapy;

Random Forest;

Decision Tree;

✉ Corresponding Author

Fandi Yulian Pamuji

Universitas Dian Nuswantoro

Semarang

Tel. +6289603458246

fandiyulian6@gmail.com

ABSTRAK

Menjaga kesehatan merupakan tugas pada diri kita sendiri sebagai manusia, dengan tubuh yang sehat kita dapat melakukan aktifitas dan produktifitas kita sehari-hari. Kesehatan adalah keadaan seimbang yang dinamis, dipengaruhi faktor genetik, lingkungan dan pola hidup sehari-hari seperti makan, minum, kerja, dan istirahat. *Immunotherapy* merupakan metode pengobatan untuk mengobati penyakit kanker kulit yang dimana metode ini meningkatkan sistem kekebalan tubuh untuk mengatasi penyakit kutil. *Immunotherapy* merupakan pengobatan efektif untuk penyakit kutil. Metode ini memiliki respon yang lebih baik membutuhkan lebih sedikit sesi, dan mampu mengobati kutil yang sudah lama. Berdasarkan hasil pengujian metode *Random Forest* dan *Decision Tree* terhadap prediksi keberhasilan pengobatan imunoterapi untuk penyakit kutil menggunakan bahasa *Python*, maka dapat diambil kesimpulan bahwa hasil pengujian metode *Decision Tree* masih terdapat prediksi yang tidak tepat dengan tingkat akurasi 84,4 % kemudian metode *Random Forest* prediksinya tepat dengan tingkat akurasi 85,5 %. Hasil pengujian tersebut menunjukkan bahwa metode *Random Forest* merupakan metode yang lebih unggul dibandingkan dengan metode *Decision Tree*.

PENDAHULUAN

Menjaga kesehatan merupakan tugas pada diri kita sendiri sebagai manusia, dengan tubuh yang sehat kita dapat melakukan aktifitas dan produktifitas kita sehari-hari. Kesehatan adalah keadaan seimbang yang dinamis, dipengaruhi faktor genetik, lingkungan dan pola hidup sehari-hari seperti makan, minum, kerja, dan istirahat [4]. Kanker merupakan penyakit yang berbahaya sebagai penyebab kematian. Hal ini mempengaruhi perkembangan pada penelitian terbaru dalam menemukan obat-obat baru dengan bahan alamipun kini banyak diteliti untuk pen-gobatan penyakit kanker yang sekarang banyak dialami [5]. Oleh karena itu menjaga kesehatan sangat penting untuk diri kita sendiri. Proses pengobatan yang membutuhkan waktu lama serta biaya pengobatan yang mahal menjadi beban bagi pasien penderita kanker. Dengan mahalnya dan timbulnya kecemasan dari pengobatan kanker tersebut, maka diperlukan sebuah teknologi pengobatan yang tepat dan tingkat keberhasilannya cukup tinggi untuk proses penyembuhan penyakit kanker [8].

Immunotherapy merupakan metode pengobatan untuk mengobati penyakit kanker kulit yang dimana metode ini meningkatkan sistem kekebalan tubuh untuk mengatasi penyakit kutil. *Immunotherapy* merupakan pengobatan efektif untuk penyakit kutil. Metode ini memiliki respon yang lebih baik membutuhkan lebih sedikit sesi, dan mampu mengobati kutil yang sudah lama [2].

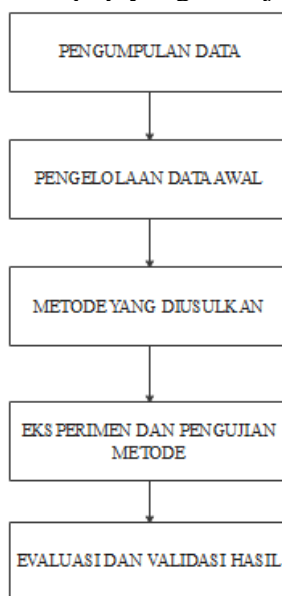
Pada penelitian oleh toni arifin dan siti syalwah dengan judul prediksi keberhasilan *immunotherapy* pada penyakit kutil dengan menggunakan algoritma *Naive Bayes*, da-lam

proses penelitian diperoleh metode data mining menggunakan algoritma *Naive Bayes* dengan *10-Fold Cross Validation* menghasilkan nilai akurasi klasifikasi sebesar 81,11% [1]. Pada penelitian selanjutnya oleh Jordy Lasmana Putra dan Mugi Raharjo dengan judul penerapan *Neural Network* dalam menentukan tingkat keberhasilan *immunotherapy* dalam proses penelitian ini teknologi pengobatan yang diyakini sebagai pengobatan kanker yaitu *immunotherapy* menjadi solusi yang tepat dalam menangani kasus kanker kulit. Penelitian ini bertujuan untuk menentukan tingkat keberhasilan *immunotherapy* dalam menyembuhkan penyakit kanker kulit, yang dimana penelitian ini dapat menghasilkan klasifikasi menggunakan algoritma *Neural Network* dengan nilai akurasi sebesar 80,00% [7].

Random Forest dan *Decision Tree* merupakan metode algoritma data mining yang digunakan untuk melakukan klasifikasi sebuah dataset. Penelitian ini dilakukan dengan tujuan untuk melakukan perbandingan antara algoritma *Random Forest* dengan *Decision Tree* dalam memprediksi keberhasilan metode pengobatan *immunotherapy* pada pengobatan penyakit kutil.

METODE

Penelitian ini dilakukan dengan memprediksi keberhasilan metode pengobatan *immunotherapy* pada pengobatan penyakit kutil. Tujuan menggunakan dua metode tersebut untuk mengetahui perbandingan antara metode *Random Forest* dan *Decision Tree* dalam pengobatan penyakit. Adapun alur kerja sistem yang dibangun secara umum dalam melakukan klasifikasi dataset *immunotherapy* yang ditunjukkan pada gambar 1 dibawah ini.



Gambar 1. Desain Penelitian

Pengumpulan Data

Pada tahapan awal dilakukan kegiatan pengumpulan data dengan cara melakukan pencarian dataset yang dapat digunakan untuk klasifikasi. Dataset yang digunakan dalam penelitian ini berupa data public bersumber dari situs *UCI Machine Learning Repository* yaitu *Immunotherapy Dataset*. Dataset tersebut berisi data-data keberhasilan pengobatan imunoterapi dalam pengobatan penyakit kutil. Data tersebut sebanyak 90 record yang terdiri dari delapan atribut dan satu atribut label [6].

Pengelolaan Data Awal

Dalam pengelolaan data awal penulis melakukan beberapa tahapan sebagai berikut:

- **Data Validation**

Pada tahap ini penulis melakukan identifikasi dan validasi data pada dataset *immunotherapy* apakah data tersebut ada yang tidak lengkap. Dalam penelitian sebanyak 90 record data telah lulus validasi.

- **Data Integration dan Transformation**

Dalam penelitian ini beberapa data pada dataset *immunotherapy* diubah menjadi data kategorikal yang mampu mewakili data.

Metode yang diusulkan

- **Algoritma Random Forest**

Random Forest (RF) adalah algoritma yang menggunakan metode pemisahan biner rekursif untuk mencapai node akhir dalam struktur pohon berdasarkan pohon klasifikasi dan regresi. Algoritma *Random Forest* menunjukkan beberapa kelebihan diantaranya mampu menghasilkan *error* yang *relative* rendah, performa yang baik dalam klasifikasi, dapat mengatasi data pelatihan dalam jumlah besar secara efisien, serta metode yang efektif untuk mengestimasi *missing* data. *Random Forest* menghasilkan banyak pohon independen dengan subset yang dipilih secara acak melalui *bootstrap* dari sampel pelatihan dan dari variable *input* disetiap node [3].

- **Algoritma Decision Tree**

Decision Tree merupakan algoritma yang umum digunakan untuk pengambilan keputusan. *Decision Tree* akan mencari solusi permasalahan dengan menjadikan kriteria sebagai node yang saling berhubungan membentuk seperti struktur pohon. *Decision tree* adalah model prediksi terhadap suatu keputusan menggunakan struktur hirarki atau pohon. Setiap pohon memiliki cabang, cabang mewakili suatu atribut yang harus dipenuhi untuk menuju cabang selanjutnya hingga berakhir di daun. Konsep data dalam *Decision Tree* adalah data dinyatakan dalam bentuk *table* yang terdiri dari atribut dan *record* [9].

- **Rule Model**

Aturan *rule* model merupakan penjelasan yang merepresentasikan sebuah algoritma *Random Forest* dan *Decision Tree*.

- **Validasi dan Pengujian**

Validasi dan pengujian merupakan pengujian yang dilakukan untuk mengetahui semua fungsi bekerja dengan baik atau tidak. Validasi dilakukan dengan *10-fold Cross Validation*. *10-fold Cross Validation* adalah validasi yang dilakukan dengan cara membagi suatu set data menjadi sepuluh segmen yang berukuran sama besar dengan cara melakukan pengacakan data. Validasi dan pengujian dilakukan untuk mengetahui tingkat akurasi dari hasil prediksi klasifikasi.

HASIL DAN PEMBAHASAN

Dalam penelitian ini metode yang digunakan adalah metode algoritma *Random Forest* dan *Decision Tree* kemudian *tool* yang digunakan untuk pengolahan data yaitu *Python*.

Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data hasil pengobatan penyakit kutil menggunakan immunotherapy yang ditunjukkan pada gambar 2 dibawah ini.

```
In [22]: dataset = pd.read_excel("Immunotherapy.xlsx")
dataset.head(91)
```

```
Out[22]:
```

	sex	age	Time	Number_of_Warts	Type	Area	induration_diameter	Result_of_Treatment
0	1	22	2.25	14	3	51	50	1
1	1	15	3.00	2	3	900	70	1
2	1	16	10.50	2	1	100	25	1
3	1	27	4.50	9	3	80	30	1
4	1	20	8.00	6	1	45	8	1
...
85	1	40	5.50	8	3	69	5	1
86	1	38	7.50	8	2	56	45	1
87	1	46	11.50	4	1	91	25	0
88	1	32	12.00	9	1	43	50	0
89	2	23	6.75	6	1	19	2	1

90 rows x 8 columns

Gambar 2. Data Immunotherapy

Validasi dan Pengujian

Pengujian dilakukan dengan validasi silang. Salah satu jenis validasi silang adalah *cross validation*. Berikut hasil dari *cross validation* yang ditunjukkan pada gambar 3 dibawah ini.

```
In [11]: from sklearn.tree import DecisionTreeClassifier
clf = DecisionTreeClassifier()
clf.fit(X,y)
scores = cross_val_score(clf, X, y, cv=10,scoring='accuracy')
print("Accuracy:" + str(scores.mean()))
```

Accuracy:0.8444444444444444

```
In [12]: clf = RandomForestClassifier()
clf.fit(X,y)
scores = cross_val_score(clf, X, y, cv=10,scoring='accuracy')
print("Accuracy:" + str(scores.mean()))
```

Accuracy:0.8555555555555555

Gambar 3. Hasil Accuracy Cross Validation

Berdasarkan hasil pengujian dengan menggunakan *Cross-Validation* maka di dapatkan nilai akurasi 84,4 % dengan menggunakan metode *Decision Tree* kemudian nilai akurasi 85,5% dengan menggunakan metode *Random Forest*.

SIMPULAN DAN SARAN

Berdasarkan hasil pengujian metode *Random Forest* dan *Decision Tree* terhadap prediksi keberhasilan pengobatan imunoterapi untuk penyakit kutil menggunakan bahasa Python. Maka dapat diambil kesimpulan bahwa hasil pengujian menggunakan metode *Decision Tree* masih terdapat prediksi yang tidak tepat dengan tingkat akurasi 84,4% kemudian metode *Random Forest* prediksinya tepat dengan tingkat akurasi 85,5%. Hasil pengujian tersebut menunjukkan bahwa metode *Random Forest* merupakan metode yang lebih unggul dibandingkan dengan metode *Decision Tree*.

DAFTAR PUSTAKA

[1] Arifin, T., & Syalwah, S. (2020). Prediksi Keberhasilan Immunotherapy Pada Penyakit Kutil Dengan Menggunakan Algoritma Naïve Bayes. *Jurnal Responsif*, 2(1), 38–43.

[2] Cahyanti, F. L. D., Gata, W., & Sarasati, F. (2021). Implementasi Algoritma Naïve Bayes dan K-Nearest Neighbor Dalam Menentukan Tingkat Keberhasilan Immunotherapy Untuk Pengobatan Penyakit Kanker Kulit. *Jurnal Ilmiah Universitas Batanghari Jambi*,

- 21(1), 259. <https://doi.org/10.33087/jiubj.v21i1.1189>
- [3] Devella, S., Yohannes, Y., & Rahmawati, F. N. (2020). Implementasi Random Forest Untuk Klasifikasi Motif Songket Palembang Berdasarkan SIFT. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 7(2), 310–320. <https://doi.org/10.35957/jatisi.v7i2.289>
- [4] Fazriansyah, A., Azis, M. A., & Yudhistira, Y. (2020). Analysis of Neural Network Classification Algorithm To Know the Success Level of Immunotherapy. *Jurnal Techno Nusa Mandiri*, 17(1), 57–62. <https://doi.org/10.33480/techno.v17i1.1089>
- [5] Fitriyani, F., & Arifin, T. (2020). Implementasi Greedy Forward Selection untuk Prediksi Metode Penyakit Kutil Menggunakan Decision Tree. *JST (Jurnal Sains Dan Teknologi)*, 9(1), 76–85. <https://doi.org/10.23887/jst-undiksha.v9i1.24896>
- [6] Khozeimeh, F., Jabbari Azad, F., Mahboubi Oskouei, Y., Jafari, M., Tehranian, S., Alizadehsani, R., & Layegh, P. (2017). Intralesional immunotherapy compared to cryotherapy in the treatment of warts. *International Journal of Dermatology*, 56(4), 474–478. <https://doi.org/10.1111/ijd.13535>
- [7] Putra, J. L., & Raharjo, M. (2019). Penerapan Neural Network Dalam Menentukan Tingkat Keberhasilan Immunotherapy. *IJCIT (Indonesian Journal on Computer and Information Technology)*, 4(2), 132–136. <https://doi.org/10.31294/ijcit.v4i2.6242>
- [8] Sa, H., Indriani, F., Nugrahadi, D. T., Yani, J. A., Banjarbaru, K., & Selatan, K. (2019). *METODE NAIVE BAYES GAUSSIAN DALAM PEMILIHAN PENGOBATAN KUTIL (IMMUNOTHERAPY DAN. 2, 24–36.*
- [9] Sutoyo, I. (2018). Implementasi Algoritma Decision Tree Untuk Klasifikasi Data Peserta Didik. *Jurnal Pilar Nusa Mandiri*, 14(2), 217. <https://doi.org/10.33480/pilar.v14i2.926>