JIUBJ

Jurnal Ilmiah Universitas Batanghari Jambi, 21(1), Februari 2021, 259-262
Lembaga Penelitian dan Pengabdian kepada Masyarakat universitas Batanghari Jambi
ISSN 1411-8939 (Online), ISSN 2549-4236 (Print)
DOI 10.33087/jiubj.v21i1.1189

Implementasi Algoritma Naïve Bayes dan K-Nearest Neighbor Dalam Menentukan Tingkat Keberhasilan Immunotherapy Untuk Pengobatan Penyakit Kanker Kulit

F. Lia Dwi Cahyanti[1*], Windu Gata[2], Fajar Sarasati[3]
[1,2]Program Studi Ilmu Komputer, STMIK Nusa Mandiri
[3]Program Studi Sistem Informasi, STMIK Nusa Mandiri
[*]Correspondence email: 14002402@nusamandiri.ac.id, windu@nusamandiri.ac.id, fajar.fss@nusamandiri.ac.id

Abstrak. Penyakit kanker kulit merupakan penyakit yang tumbuh dijaringan kulit dimana kondisi ini ditandai dengan perubahan pada kulit, seperti muncunya benjolan, bercak, atau tahi lalat dengan ukuran tidak normal, salah satu penyebab kanker kulit ialah paparan sinar ultraviolet dari matahari. Salah satu pengobatan penyakit kanker kulit adalah immunotherapy, metode immunotherapy merupakan pengobatan penyakit dengan cara mengaktifkan atau menekan sistem imun pada tubuh. Dalam penelitian ini dilakukan komparasi dengan metode dari data mining untuk klasifikasi yaitu naïve bayes dan K-Nearest Neighbor untuk memprediksi tingkat keberhasilan immunotherapy dalam menyembuhkan penyakit kanker kulit. Dalam proses pengujiannya, peneliti menggunkan aplikasi weka untuk mengolah data dan melakukan pengujian. Hasil pengujian yang telah dilakukan menunjukan model K-Nearest Neighbor didapat nilai akurasi terbaik yaitu 91,1111%. sedangkan Naïve Bayes didapat nilai akurasi yang lebih kecil yaitu 82,2222% dari hasil pengujian dapat disimpulkan bahwa metode K-Nearest Neighbor memiliki akurasi lebih baik dalam menentkan tingkat keberhasilan immunotherapy.

Kata kunci: Immunotherapy; k-nearest neighbour; klasifikasi; naïve bayes

Abstract. *Cancer is a disease that grows in the skin tissue where this condition is characterized by changes in the skin, such as the appearance of lumps, spots, or moles with abnormal sizes, one of the causes of skin cancer is exposure to ultraviolet rays from the sun. One of the treatments for skin cancer is immunotherapy, the immunotherapy method is the treatment of disease by activating or suppressing the immune system in the body. In this study, a comparison with data mining methods for classification was carried out, namely Naïve Bayes and K-Nearest Neighbor to predict the success rate of immunotherapy in curing skin cancer. In the testing process, the researcher uses the Weka application to process data and conduct tests. The results of the tests that have been carried out show that the K-Nearest Neighbor model has the best accuracy value of 91.1111%. while Naïve Bayes obtained a smaller accuracy value, namely 82.2222%. From the test results, it can be concluded that the K-Nearest Neighbor method has better accuracy in determining the success rate of immunotherapy.*

Keywords: *Immunotherapy; classification; k-nearest neighbour; naïve bayes*

PENDAHULUAN

Kesehatan merupakan salah satu harta yang sangat berharga dalam kehidupan manusia, Kesehatan dianggap jauh lebih berharga daripada sebuah harta materi, Karena itu manusia akan selalu berusaha menjaga kesehatannya agar terhindar dari segala penyakit, Salah satu caranya yaitu memperhatikan pola hidup dan segala aspek kondisi kesehatan, Mulai dari aspek makanan, minuman, nutrisi yang dikonsumsi dan perilaku kita sehari-hari. (Hasyim, 2017)

Kanker merupakan salah satu penyakit utama penyebab kematian di dunia. Pada 2012 diperkirakan terdapat 14 juta kasus baru kanker dan 8,2 juta kematian akibat kanker di dunia. (Dewi, 2017). Kualitas hidup penderita kanker dipengaruhi pemahaman individu terhadap penyakitnya sehingga seseorang tahu cara menjaga kesehatan, serta faktor ekonomi dimana hal ini menjadi kekhawatiran khusus terhadap biaya pengobatan. (Anita, 2016). Mahalnya biaya pengobatan kanker tersebut, sangat diperlukan adanya teknologi pengobatan yang tepat untuk penyembuhan kanker dan tingkat keberhasilan yang cukup tinggi

Salah satu penanganan pada penyakit yang disebabkan oleh HPV yaitu Immunotherapy. Metode ini ditemukan pada tahun 1970-an untuk penyembuhan kanker melanoma dan kandung kemih dengan menggunakan agen imun cytokine namun memiliki efek samping yang besar dan seiring mempelajari respon imun, metode ini terus dikembangkan. (Rahmadi et al, 2020)

Algoritma Naive bayes dan K-Nearest Neighbor merupakan algoritma data mining yang digunakan untuk melakukan proses klasifikasi. Penelitian ini kami lakukan dengan tujuan untuk melakukan perbandingan antara algoritma Naive bayes dengan K-Nearest Neighbor dalam memprediksi keberhasilan metode pengobatan imunoterapi pada pengobatan penyakit kanker kulit.

Pada penelitian sebelumnya menunjukkan bahwa *Immunotherapy* memiliki tingkat keberhasilan yang lebih baik dibandingkan dengan metode *Cryotherapy* , dari hasil penelitian yang dilakukan Fahime Khozeimeh, dkk dengan judul *Intralesional immunotherapy compared to cryotherapy in the treatment of warts* ,

menyatakan bahwa hasil yang diperoleh melalui metode immunotherapy dapat memberikan efek positif dengan tingkat respon yang lebih tinggi yaitu 76,7% pasien sembuh total selain itu, terbukti menjadi pengobatan yang memiliki respons terapeutik yang tinggi, jumlah sesi pengobatan yang diperlukan lebih sedikit dan resiko efek samping yang lebih rendah. Sementara hanya 57,7% pasien sembuh dengan metode cryotherapy. (Khozeimeh et al. 2017)

METODE

Metode penelitian yang kami gunakan yaitu komparasi antara dua perbandingan metode klasifikasi dengan algoritma data mining yaitu metode K-Nearest Neighbor (K-NN) dan Naïve Bayes. K-nearest neighbor (k-nn) merupakan salah satu algoritma yang popular pada machine learning, algoritma dengan proses mudah dan sederhana. k-nn merupakan salah satu dari algoritma supervised learning dengan proses belajar. Dalam algoritma k-nn data yang diolah harus memiliki label. Sedangkan Naïve Bayes merupakan metode yang menampilkan label kelas terkait walaupun data training yang digunakan sedikit. Penelitian ini juga menggunakan model klasifikasi. Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. (Rahmadi et al, 2020)

Metode Naïve Bayes

Teori keputusan bayes adalah pendekatan statistik yang fundamental dalam pengenalan pola (pattern recognition), pendekatan ini didasarkan pada kuantifikasi trade-off antara berbagai keputusan klasifikasi dengan menggunakan probabilitas dan ongkos yang ditimbulkan dalam keputusan tersebut. (Rahmadi et al, 2020)

Naïve Bayes Classifier atau sering disebut Bayesian Classification adalah metode pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Naïve Bayes Classifier (NBC) terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar. (Zulfikar and Lukman, 2016)

Menurut Kusirin dalam (Zulfikar and Lukman, 2016) Naïve Bayes Classifier didasarkan kepada teorema Bayes yang memiliki bentuk umum sebagai berikut:

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$$

Keterangan:
X = data dengan class yang belum diketahui
H = hipotesis data X merupakan suatu class spesifik
P(H|X) = probabilitas hipotesis H berdasar kondisi X (posteriori probability)
P(X|H) = probabilitas X berdasar kondisi pada hipotesis H
P(X) = probabilitas dari X

B. K-Nearest Neighbor

Nearest Neighbor adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada. Misalkan diinginkan untuk mencari solusi terhadap seorang pasien baru dengan menggunakan solusi dari pasien terdahulu. Untuk mencari kasus pasien mana yang akan digunakan, maka dihitung kedekatan kasus pasien baru dengan semua kasus pasien lama. Kasus pasien lama dengan kedekatan terbesarkan yang akan diambil solusinya untuk digunakan pada kasus pasien baru. (Zulfikar and Lukman, 2016)

$$\text{similarity} = \frac{\sum_{i=1}^{n} f(T_i, S_i)}{W_i}$$

Keterangan:
T : Kasus baru
S : kasus yang ada dalam penyimpanan
N : jumlah atribut
i : atriut individu antara 1 – n
f : fungsi similarity ATRIBUT I antara kasus T dan kasus S
w : bobot yang diberikan pada atribut ke-i

Weka

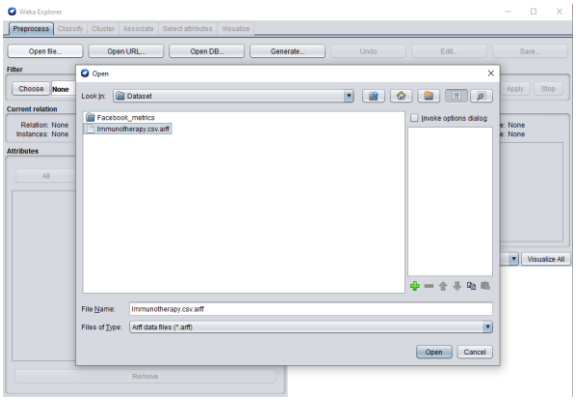
Weka mulai dikembangkan sejak tahun 1994 dan telah menjadi software data mining open source yang paling populer. Weka mempunyai kelebihan seperti mempunyai banyak algoritma data mining dan machine learning, kemudahan dalam penggunaannya, selalu up-to-date dengan algoritma-algoritma yang baru. Software Weka tidak hanya digunakan untuk akademik saja namun cukup banyak dipakai oleh perusahaan untuk meramalkan bisnis dari suatu perusahaan.Semua teknik Weka adalah didasarkan pada asumsi bahwa data tersedia sebagai flat file tunggal atau hubungan, dimana setiap titik data digambarkan oleh sejumlah atribut (biasanya, atribut numerik atau nominal, tetapi beberapa jenis atribut lain juga didukung). (Rahmadi et al, 2020)

HASIL DAN PEMBAHASAN

Naïve Bayes dan K-Nearest Neighbor tools yang di gunakan adalah aplikasi Weka. Disini penelitian menggunakan data set *Immunotherapy* yang di ambil dari repositori UCI (Universitas California, Invene) yang ada pada website dengan link https://archive.ics.uci.edu/ml/datasets/Immunotherapy+

Dataset dengan data sebanyak 90 records. Dibagi menjadi beberapa tahap yaitu :

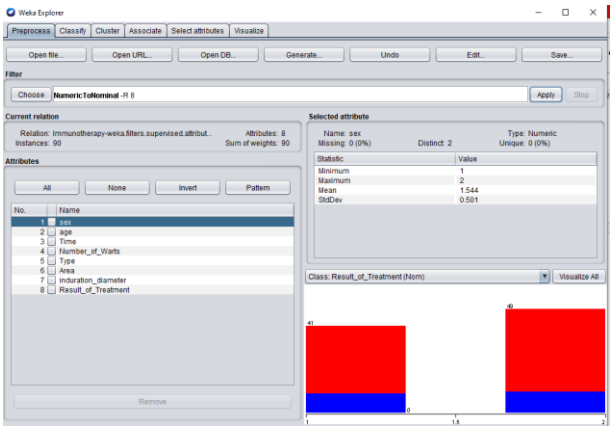
1. Mengubah format data dari dataset, Save As data tersebut dan ganti formatnya yang tadinya (_hasil.arff) untuk digunakan sebagai data testing.
2. Memasukkan dataset



Gambar 1. Mengambil data yang akan diinputkan

Setelah mengaplikasikan dataset, tahap pemodelan pada penelitian ini selanjutnya dilakukan proses klasifikasi menggunakan algoritma yang akan digunakan, yaitu algoritma Naïve Bayes, dan K-Nearest Neighbor.

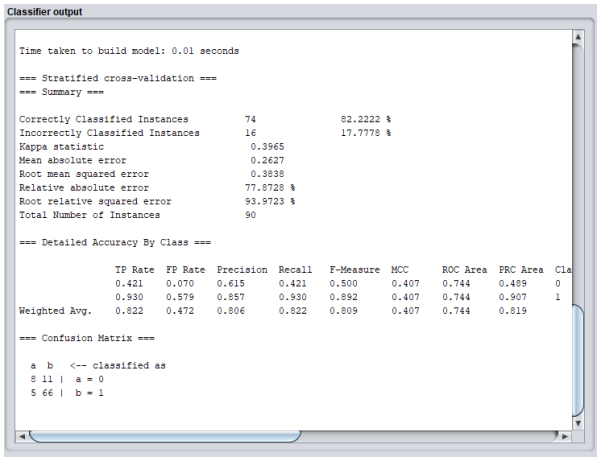
3. Pengaplikasian Dataset *Immunotherapy*



Gambar 2. Pengaplikasian Dataset

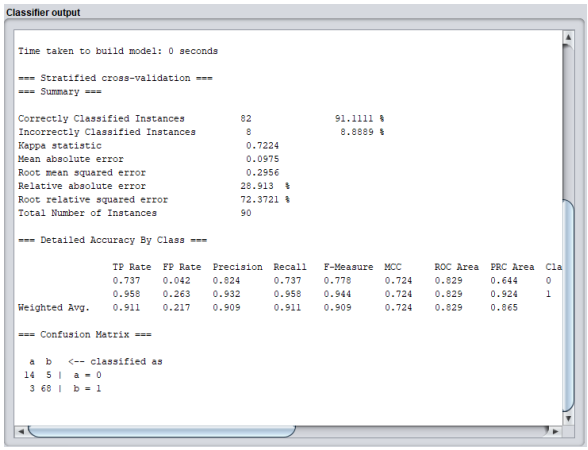
Untuk mengevaluasi permorfa dari masing-masing algoritma yang digunakan, pada penelitian ini dilakukan pembagian persentasi antara data training dan data testing. Class yang digunakan pada penelitian ini adalah result of trearment yang digunakan untuk membagi data menjadi dua set (Label) dengan nilai 1 dan 0. Dimana, 1 adalah berhasil dan 0 berarti gagal.

4. Menentukan metode Naïve Bayes untuk dataset Immunotherapy , Hasil dari metode naïve bayes yang telah dilakukan.



Gambar 3. Data hasil kasifikasi naïve bayes

5. Hasil Klasifikasi menggunakan metode K-Nearest Neighbor



Gambar 4. Data hasil klasifikasi menggunkan metode k-nearest neighbor

Dari data hasil yang diperoleh pengujian menggunakan aplikasi WEKA berisi nilai-nilai tertentu dan untuk membandingkan algoritma klasifikasi sebaik apa data yang akan dipilih dari hasil nilai *f1-score*, akurasi, *kappa statistic*, *mean absolute error* dari setiap algoritma yang diterapkan. Hasil ringkasan pengujian data tercantum dalam tabel 1 dibawah ini.

Tabel 1. Hasil Pengujian Data pada Aplikasi WEKA

Classifier	F1-Score	Akurasi	Kappa Statistic	MAE
Naive Bayes	0,809%	82,2%	0,396	0,262
K-NN	0,909%	91,1%	0,722	0,097

6. Hasil F1-Score

F1-score merupakan alat ukur yang mengombinasikan nilai precision dan recall.(Limantoro, Kristian, and Purwanto 2018). hasil data pengklasifikasian yang telah diuji menggunakan aplikasi WEKA, nikai *f1 score* tertinggi dengan algoritma *K-Nearest Neighbor*

dengan nilai 0,909 , Nilai F1 Score berkisar 0 sampai dengan 1, semakin mendekai nilai 1 maka semakin akurat. (Kautsar and Adi 2016) sedangakan nilai dengan algoritma *Naïve Bayes* 0,809.

7. Hasil Akurasi

   Hasil data pengklasifikasian yang telah diuji menggunakan aplikasi WEKA dengan algoritma *K-Nearest Neighbor* nilai akurasi tertinggi yaitu 91,1%, Sedangkan Naïve Bayes yang memiliki akurasi terendah dengan nilai 82,2%.

8. Hasil *Kappa Statistic*

   Dari pengujian data menggunakan aplikasi WEKA, hasil pengklasifikasian data yang tertinggi menggunakan algoritma *K-Nearest Neighbor*, dengan nilai *Kappa Statistic* 0,7224. Kappa Statistics mendekati nilai 1, sehingga dapat dikategorikan sangat baik untuk digunakan sebagai model. Untuk nilai *Kappa Statistic* terendah dihasilkan algoritma *Naive Bayes* dengan nilai 0,396.

9. Hasil *Mean Absolute Error*

   Dari data yang diuji menggunakan WEKA, maka hasil mengklasifikasian data algoritma *K-Nearest Neighbor memiiki* nilai *mean absolut error* nya rendah 0,0975 dan untuk nilai *mean absolut error* tertinggi algoritma *Naive Bayes* 0,2627. Mean Absolute Error (MAE) adalah salah satu metode yang digunakan untuk mengukur tingkat keakuratan model peramalan, Hasil Mean Absolute Error menunjukkan nilai rata-rata kesalahan multak atau absolut dari nilai sebenarnya dengan nila i peramalan. (Suryanto and Muqtadir 2019)

SIMPULAN

Berdasarkan penelitian, implementasi dan pengujian, maka dapat diambil kesimpulan sebagai berikut :

Dari hasil eksperimen menyimpulkan bahwa klasifikasi yang didapat dari kolaborasi pengujian menggunakan metode machine learning yaitu naïve bayes dan k-nearest neighbor menunjukkan bahwa akurasi terbaik untuk keberhasilan immunotherapy ketika menggunakan metode K-Nearest Neighbor sebesar 91,1111%. Sedangkan menggunakan metode naïve bayes rata-rata akurasi sebesar 82,2222% . Dengan demikian penelitian ini dapat dijadikan acuan dalam menentukan tingkat keberhasilan metode immunotherapy dalam penyembuhan penyakit kanker kulit.

DAFTAR PUSTAKA

Anita. 2016. “PERAWATAN PALIATIF DAN KUALITAS HIDUP PENDERITA KANKER.” *National Institute of Health* VII(3): 508–13.

Dewi, Makassari. 2017. “Sebaran Kanker Di Indonesia, Riset Kesehatan Dasar 2007.” *Indonesian Journal of Cancer* 11(1): 1–8.

Hasyim, lazuardi adi pradana. 2017. “KAJIAN PERANAN VISUAL SEBAGAI PANDUAN REMAJA UNTUK MENJAGA KESEHATAN DALAM MEDIA BUKU Lazuardi.” *JURNAL RUPARUPA* 2(2): 161–72.

Kautsar, Havez Vazirani Al, and Kusworo Adi. 2016. “Implementasi Object Tracking Untuk Mendeteksi Dan Menghitung Jumlah Kendaraan Secara Otomatis Menggunakan Metode Kalman Filter Dan Gaussian Mixture Model.” *Youngster Physics Journal* 5(1): 13–19.

Khozeimeh, Fahime et al. 2017. “Intralesional Immunotherapy Compared to Cryotherapy in the Treatment of Warts.” *International Journal of Dermatology* 56(4): 474–78.

Limantoro, Stephen Ekaputra, Yosi Kristian, and Devi Dwi Purwanto. 2018. “Pemanfaatan Deep Learning Pada Video Dash Cam Untuk Deteksi Pengendara Sepeda Motor.” *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)* 7(2): 3–9.

Rahmadi, Muhammad, Fazriyanor Kaurie, and Tuti Susanti. 2020. “Uji Akurasi Dataset Pasien Pasca Operasi Menggunakan Algoritma Naïve Bayes Menggunakan Weka Tools.” *JURIKOM (Jurnal Riset Komputer)* 7(1): 134–39.

Suryanto, Andik Adi, and Asfan Muqtadir. 2019. “PENERAPAN METODE MEAN ABSOLUTE ERROR (MEA) DALAM ALGORITMA REGRESI LINEAR UNTUK PREDIKSI PRODUKSI PADI.” (1): 78–83.

Zulfikar, Wildan Budiawan, and Nur Lukman. 2016. “Perbandingan Naive Bayes Classifier Dengan Nearest Neighbor Untuk Identifikasi Penyakit Mata.” *Jurnal Online Informatika* 1(2): 82–86.