KLASIFIKASI JENIS BERITA MENGGUNAKAN ALGORITMA *C4.5*

ARTIKEL JURNAL

RUSLINA MOHI
NIM. A01417044

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN ILMU KOMPUTER
UNVERSITAS MUHAMMADIYAH GORONTALO
GORONTALO
2023

KLASIFIKASI JENIS BERITA MENGGUNAKAN ALGORITMA *C4.5*

Ruslina Mohi [1], Rubiyanto Maku[2], Irawan Ibrahim[3]

[2]Program Studi Sistem Informasi Universitas Muhammadiyah Gorontalo. Gorontalo 9600. Indonesia

Email Penulis: ruslinamohi@gmail.com

ABSTRAK

Radio merupakan media auditif yang murah dan dapat dinikmati siapa saja. Walaupun hanya dapat dinikmati secara auditif, radio memiliki kelebihan, yakni kemampuan menciptakan imajinasi audiens. Radio menimbulkan stimuli dengan beragam suara, dan berupaya memvisualisasikan suara penyiar melalui telinga pendengarnya. Penelitian ini bertujuan untuk mengetahui hasil klasifikasi berita berdasarkan Algoritma C45. Data yang digunakan pada penelitian ini adalah data berita berjumlah 160 record dari tahun 2019-2021. Berdasarkan hasil penelitian menggunakan penerapan data mining algoritma C4.5 mampu memberikan solusi dalam menentukan jenis berita. Tingkat akurasi yang dihasilkan oleh metode tersebut yaitu sebesar 96.25%.

Kata Kunci : Algoritma C45, Jenis Berita, Klasifikasi.

ABSTRACT

Radio is an auditive media that is cheap and can be enjoyed by anyone. Although it can only be enjoyed auditively, radio has its advantages, namely the ability to create the audience's imagination. Radio creates stimuli with a variety of sounds, and attempts to visualize the announcer's voice through the listener's ear. This research aims to determine the results of news classification based on the C45 Algorithm. The data used in this research is news data totaling 160 records from 2019-2021. Based on the results of research using the application of data mining algorithm C4.5 is able to provide solutions in determining the type of news. The level of accuracy produced by this method is equal to 96.25%.

Keywords: C45 Algorithm, Types of News, Classification.

1. PENDAHULUAN

Berita radio adalah berita yang disiarkan atau di publikasikan lewat media radio.Berita Radio (*Radio News*)merupakan materi utama program siaran radio berita (*News Radio*),yaitu radio dengan format *All News*.Hampir semua radio,dengan format apapun memiliki program siaran(news programmers) yaitu program yang bermaterikan informasi aktual. Hard News berita lempang atau langsung di sebut juga *Straight News* dan *Spot News*.Sebagai media penyiaran, radio acapkali tidak memiliki struktur pemberitaan dalam program-programnya. Hyde mengatakan bahwa radio tidak memiliki produser berita maupun reporter dan penasihat siaran berita. Mereka hanya muncul pada saat-saat tertentu saja (2004: 220). Pada umumnya, laporan berita dari sebuah stasiun radio di suplai dari kantor berita nasional, seperti KBR 68 H Jakarta. Sementara di stasiun lainnya, berita radio diambil dari internet, surat kabar, lalu dibacakan tanpa diedit. Di stasiun radio yang demikian, kemampuan jurnalistik penyiar biasanya diabaikan; ketrampilan membaca dan mempresentasika berita menjadi yang utama. Sementara stasiun radio yang mengambil berita dari stasiun radio lain atau dari kantor berita sangat menggantungkan diri mereka pada kerja reporter dan penyaji berita. Sementara, untuk sedikit stasiun radio yang memproduksi program beritanya sendiri, biasanya sangat bergantung pada reporter lapangan dan stringers, yakni reporter part ime yang dibayar untuk menulis berita tertentu. Lalu, bagaimanakah bentuk berita radio tersebut?

Radio merupakan media auditif yang murah dan dapat dinikmati siapa saja. Walaupun hanya dapat dinikmati secara auditif, radio memiliki kelebihan, yakni kemampuan menciptakan imajinasi audiens. Radio menimbulkan stimuli dengan beragam suara, dan berupaya memvisualisasikan suara penyiar melalui telinga pendengarnya. Dengan mempertimbangkan karakteristik ini, berita radio pun memiliki kekhususannya sendiri. Riswandi (2009: 21) mencatat beberapa definisi berita radio.

Sedangkan karakteristik berita radio yang berisi keterangan secara global dapat mendekati kebenaran dari aspek emosionalitas melalui usaha penyederhanaan fakta-fakta yang terinci. Tujuannya bukan hanya agar berita lebih mudah dipahami, namun juga agar penyampaian beritanya mampu mengeksplorasi unsur rasa yang terdapat dalam berita maupun perasaan keyakinan pendengarnya. Keterangan secara global didisain tidak semata agar pemberitaan tidak membuat kening pendengar berkerut karena banyaknya aktivitas berpikir, tetapi juga memberi kesempatan kepada

mendengar mencerna atau ‘merasakan’ pemberitaannya untuk mendapatkan keyakinan akan kebenaran beritanya.

Menurut Paul D Maessenner, berita adalah sebuah informasi baru tentang suatu peristiwa yang penting dan menarik perhatian serta minat pendengar. Unsur kebaruan peristiwa merupakan poin penting dalam definisi ini. Sedangkan Mitchel Charnley memaparkan berita radio sebagai laporan tentang suatu peristiwa, opini, kecenderungan situasi kondisi, interpretasi yang penting, menarik, masih baru, dan harus secepatnya disampaikan kepada khalayak. Menurut Curttis Beckman, berita adalah laporan atas opini atau peristiwa yang penting bagi sejumlah khalayak. Berita yang besar adalah liputan opini atau peristiwa yang dibutuhkan bagi banyak orang (21-22). Dari beberapa pendapat tersebut, Riswandi menyimpulkan bahwa berita radio adalah suatu sajian laporan berupa fakta dan opini yang mempunyai nilai berita, penting, dan menarik bagi sebanyak mungkin orang , dan disiarkan melalui media radio secara berkala.

1.1 Klasifikasi

> Klasifikasi adalah suatu kegiatan mengelempokkan. Dimana klasifikasi sangat dibutuhkan dalam peprustakaan, karena klasifikasi bertujuan untuk mengelompokkan satu koleksi yang sejenis, yang pengelompokkannya berdasarkan judul, pengarang, dan lain sebagainya (Cristea, 2016).

Klasifikasi adalah suatu teknik yang dilakukan untuk memprediksi class atau properti dari setiap instace data. Teknik ini dilakukan pada data baru dengan memanipulasi data yang ada dan telah diklasifikasikan menggunakan hasilnya untuk memberikan sejumlah aturan. a. Konsep Entropy Entropy (S) merupakan jumlah bit yang diperkirakan dibutuhkan untuk dapat mengekstrak suatu kelas (+ atau -) dari sejumlah data acak pada ruang sampel S. Entropy dapat dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas. Semakin kecil nilai entropy maka akan semakin baik entropy yang digunakan dalam mengekstrak suatu kelas. Entropy digunakan untuk mengukur ketidakaslian(Nugroho, 2020)

1.2 Jenis Berita

Berita radio adalah berita yang disiarkan atau di publikasikan lewat media radio.Berita Radio (Radio News)merupakan materi utama program siaran radio berita (News Radio),yaitu radio dengan format All News.Hampir semua radio,dengan format apapun memiliki program siaran(news programmers) yaitu program yang bermaterikan informasi aktual. Hard News berita lempang atau langsung di sebut juga *Straight News* dan *Spot News*.Sebagai media penyiaran, radio acapkali tidak memiliki struktur pemberitaan dalam program-programnya. Hyde mengatakan bahwa radio tidak memiliki produser berita maupun reporter dan penasihat siaran berita. Mereka

hanya muncul pada saat-saat tertentu saja (2004: 220). Pada umumnya, laporan berita dari sebuah stasiun radio di suplai dari kantor berita nasional, seperti KBR 68 H Jakarta. Sementara di stasiun lainnya, berita radio diambil dari internet, surat kabar, lalu dibacakan tanpa diedit. Di stasiun radio yang demikian, kemampuan jurnalistik penyiar biasanya diabaikan; ketrampilan membaca dan mempresentasika berita menjadi yang utama. Sementara stasiun radio yang mengambil berita dari stasiun radio lain atau dari kantor berita sangat menggantungkan diri mereka pada kerja reporter dan penyaji berita. Sementara, untuk sedikit stasiun radio yang memproduksi program beritanya sendiri, biasanya sangat bergantung pada reporter lapangan dan stringers, yakni reporter part ime yang dibayar untuk menulis berita tertentu. Lalu, bagaimanakah bentuk berita radio tersebut(Bangun, A Koagouw, & Kalangi, 2019)

.

1.3 Algoritma *C4.5*

Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau pengelompokan pada dataset. Dasar dari algoritma C4.5 adalah pembentukan pohon keputusan (*Decision Tree*). Cabang-cabang dari pohon keputusan merupakan pertanyaan klasifikasi sedangkan untuk daun-daunnya merupakan kelas-kelas atau kelompoknya. Karena tujuan dari algortima C4.5 untuk melakukan klasifikasi, sehingga hasil dari pengolahan dataset berupa pengelompokan data ke dalam kelas-kelas tertentu.

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Pilih variabel sebagai akar
2. Buat cabang untuk masing-masing nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Hal pertama yang dilakukan untuk membentuk pohon keputusan adalah menentukan atribut/variabel mana yang menjadi akar dari pohon keputusan tersebut. Cara menentukan variabel yang menjadi akar adalah dengan menggunakan *entropy*, *gain*, *split info,* dan *gain ratio.*

2. METODE

3.1 Objek Penelitian

Objek penelitian adalah sebuah variabel atau sasaran dalam penelitian yang dilakukan oleh peneliti. Yaitu klasifikasi jenis berita menggunakan algoritma C 45 dengan lokasipenelitian di RRI Gorontalo

.

3.2 Pengumpulan Data

Dalam proses pengumpulan atau pengambilan data, dalam penelitian ini menggunakan dua cara meliputi:

1. Observasi

Penulis melakukan pengumpulan atau pengambilan data dengan cara melakukan pengamatan langsung terhadap objek yang akan diteliti dengan instansi terkait untuk mengumpulkan data dan informasi yang berkaitan dengan permasalahan yang ada.

2. Wawancara

Penulis melakukan pengambilan data dan informasi dengan cara mewawancarai atau berdiskusi dengan Pimpinan RRI Gorontalo, Bidang pemberitaan dan penyiar

3.3 Tahapan Penelitian

Adapun tahapan dalam penelitian yang akan dilakukan sebagai berikut :

1. Pemilihan Topik Penelitian

   Dalam tahapan ini peneliti mencari dan menentukan ide atau topik penelitian berdasarkan masalah yang didapat pada lokasi yang telah ditentukan.

2. Perumusan Masalah Penelitian

   Pada tahap ini peneliti melakukan perumusan masalah dari penelitian yakni,, bagaimana cara melakukan klasifikasi berita dan bagaimana menerapkan Algoritma C4.5 untuk klasifikasi berita.

3. Analisis Data

   Data yang diperoleh dilakukan analisis data, untuk menentukan target klasifikasi dengan menggunakan Algoritma C4.5.

4. Membersihkan data

   Pada tahapan ini Data yang kosong atau tidak lengkap di isi dan dibersihkan untuk menghindari kesalahan dan tingkat akurasi yang jauh dari harapan.

4.1 Hasil Penelitian

4.1.1 Pengumpulan Data

Hasil Penelitian memuat data penelitian yang telah dilakukan di Kantor LPP RRI Gorontalo dimana melakukan pengumpulan seluruh data berita. Data yang diperoleh melalui wawancara dengan pegawai LPP RRI Gorontalo Gorontalo bagian layanan dan pengembangan usaha, survei yang dilakukan langsung di lapangan, studi kasus dan studi pustaka yang di lakukan secara bertahap untuk mendapatkan data yang sesuai, kemudian diolah agar mendapatkan kesimpulan yang sesuai dengan penelitian yang dilakukan. Data yang diperoleh secara langsung dari Kantor LPP RRI. Data berita berjumlah 157 record dari tahun 2019-2021.

Atribut dalam penelitian ini adalah

1. Kategori mitra

   A1 = Lintas Sektoral / Pemerintah

   A2 = BUMN, Lembaga Negara dab sejenisnya

   A3 = Perusahaan Swasta

   A4 = Agensi dan Pihak Ketiga lain

   A5 = Pribadi / Perorangan

2. Kategori Siaran

   Dibaca langsung : berita dibaca langsung oleh penyiar tanpa diproduksi seperti berta keluarga dan berita kehiilangan

   Produksi Spot : berita dari instansi lain yang melalui tahap seleksi seperti iklan

   Produksi : berita yang di produksi oleh RRI seperti iklan dan dialog

3. Sifat

Biasa : berita yang diproduksi dan bisa di terbitkan sesuai dengan permintaan klien

Penting : berita yang segera diberitakan, seperti radiogram dan pengumunan

Sangat penting : berita yang segera diberitakan, seperti berita kedukaan dan dialog

4.1.2 Pre-Processing

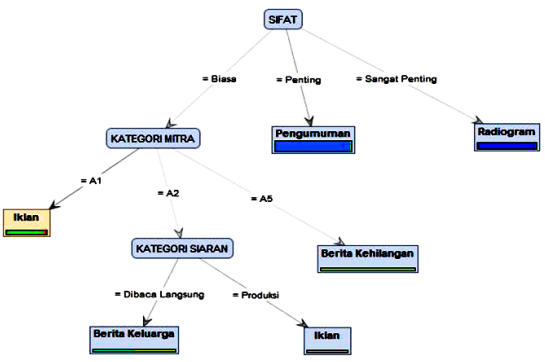
Pre - Processing merupakan tahap awal sebelum pengolahan data. Data diseleksi dan dibentuk menjadi data yang lebih berguna. Tahapan *pre -processing* sangat berguna untuk mendapatkan atribut dan data yang akan digunakan dalam penelitian ini. Atribut yang digunakan pada penelitian ini adalah kategori mitra, kategori siaran dan sifat berita periode 2019 sampai 2021.

Pada tahap ini dilakukan pemodelan data, metode yang dipakai pada

penelitian ini adalah decision tree (pohon keputusan) dengan menggunakan algoritma c45. Data yang telah di kumpul, diseleksi dan di transformasi akan di kelola menggunakan metode decision tree. Metode ini adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan.

Tahapan ini ada 3 atribut yang digunakan dalam mengklasifikasi data berita. Atribut tersebut yaitu kategori mitra, kategori siaran, sifat dan untuk jenis berita merupakan target dari proses klasifikasi. Atribut data tersebut dianalisis dengan menggunakan software RapidMiner. Data berita LPP RRI Gorontalo memiliki 157 record data. Data berita dengan format .xls yang akan diolah melalui software RapidMiner. Hasil preprocessingnya dapat dilihat pada tabel berikut

Untuk hasil pohon keputusan dapat dilihat bahwa *node* tertinggi adalah atribut sifat dengan demikian kondisi dari *node* tertinggi bisa di jadikan sebagai faktor yang paling menentukan dalam klasifikasi jenis berita. Berdasarkan dari rule/knowledge yang dihasilkan terdapat beberapa Rules cukup sesuai dengan hasil perhiungan manual algoritma c45 yakni nilai entropi atribut sifat lebih tinggi dari nilai entropi atribut kategori mitra dan atribut kategori siaran sedangkan nilai gain atribut sifat lebih rendah dari nilai entropi atribut kategori mitra dan atribut kategori siaran. Pohon keputusan yang terbentuk pada gambar berikut.



Gambar 1. Pohon Keputusan

3. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan dengan menggunakan Algoritma C4.5, dimana menghasilkan suatu pohon keputusan dengan aturan yang telah dibuat berdasarkan pohon keputusan yang sama maka kesimpulan yang didapat adalah:

2. Penerapan algoritma *C4.5* dengan metode pohon keputusan dapat memberikan informasi rule prediksi untuk menggambarkan proses yang terkait dengan klasisfikasi jenis berita
3. Dengan adanya penerapan data mining algoritma *C4.5* mampu memberikan solusi dalam menentukan jenis berita. Tingkat akurasi yang dihasilkan oleh metode tersebut yaitu sebesar 98.09%

.

DAFTAR PUSTAKA

Asroni, A., Masajeng Respati, B., & Riyadi, S. (2018). Penerapan Algoritma C4.5 untuk Klasifikasi Jenis Pekerjaan Alumni di Universitas Muhammadiyah Yogyakarta. *Semesta Teknika*, *21*(2), 158–165. https://doi.org/10.18196/st.212222

Bangun, E. P., A Koagouw, F. V. I., & Kalangi, J. S. (2019). Analisis isi unsur kelengkapan berita pada media online manadopostonline.com. *Acta Diurna Komunikasi*.

Cristea, A. (2016). Metode C45 Untuk Mengklarifikasi Pelanggan Perusahaan Telekomunikasi Seluler. *Revista Brasileira de Ergonomia*.

Latifah, K. (2018). Analisis dan penerapan algorithma c45 dalam data mining untuk menunjang strategi promosi prodi informatika upgris. *Jurnal teknik informatika*. https://doi.org/10.15408/jti.v11i2.6706

Nasrullah, A. H. (2018). Penerapan Metode C4.5 untuk Klasifikasi Mahasiswa Berpotensi Drop Out. *ILKOM Jurnal Ilmiah*, *10*(2), 244–250. https://doi.org/10.33096/ilkom.v10i2.300.244-250

Nugroho, F. A. (2020). Implementasi Data Mining Pada Pasien Setelah Koma dengan Menggunakan Metode Algoritma C 4 . 5 dan Naivy Bayes untuk Mengetahui Intelegensi (IQ) Pasien. *Seminar Nasional Dinamika Informatika 2020*.