



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 3 (2025) pp: 3765-5

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Rancang Bangun Sistem Pengelompokan Berita *Clickbait*/Non-*clickbait* Berita Radar Cirebon dengan K-Medoids

Fathul Ramdhani¹, Rifqi Fahrudin², Mesi Febima³

^{1,2,3} Universitas Catur Insan Cendekia, Indonesia

¹fathulramdhani2@gmail.com, ²rifqi.fahrudin@cic.ac.id, ³mesi.febima@cic.ac.id

Abstrak

Banyak judul berita di internet menggunakan teknik *clickbait* untuk membuat judul yang menarik tetapi seringkali tidak relevan dengan isi berita, yang menantang pengelola media dan pembaca saat ini. Menggunakan algoritma K-Medoids, penelitian ini bertujuan untuk membangun sistem berbasis web yang akan mengelompokkan judul berita dari Radar Cirebon menjadi kategori *clickbait* dan non-*clickbait*. Algoritma ini dipilih karena mampu mengurangi jarak antar data ke medoid atau pusat cluster dan lebih tahan terhadap outlier dibandingkan dengan K-Means. Setelah data dikumpulkan melalui scraping web, proses preprocessing seperti cleaning, tokenizing, stopword, stemming, dan metode TF-IDF juga digunakan untuk mengubah data menjadi data numerik. Indeks Davies Bouldin (DBI) digunakan untuk mengevaluasi kualitas pengelompokan. Untuk mengembangkan sistem, Python, database MySQL, dan antarmuka web yang mudah digunakan digunakan. Radar Cirebon dapat menyajikan berita yang lebih akurat dan dapat diandalkan berkat sistem ini, yang meningkatkan pengalaman dan kepercayaan pembaca.

Kata kunci: Clickbait, K-Medoids, Non-clickbait, Radar Cirebon, Sistem Berbasis Website.

1. Latar Belakang

Seiring dengan kemajuan teknologi, tugas-tugas menjadi lebih mudah untuk diselesaikan dengan cepat dan tepat. Teknologi informasi telah mengubah banyak hal, seperti pendidikan, perkantoran, industri, telekomunikasi, bisnis, dan pariwisata [1]. Dalam sepuluh tahun terakhir, kemajuan teknologi informasi telah mengubah cara masyarakat memandang berita. Dengan penggunaan internet dan media sosial yang meningkat, berita dapat diakses kapan saja dan di mana saja. Karena keadaan ini, ada persaingan sengit di antara media *online* untuk mendapatkan perhatian pembaca. Banyak portal berita menggunakan taktik penulisan judul menarik yang disebut *clickbait* untuk menarik perhatian pembaca dengan kalimat hiperbolis yang terkadang tidak relevan dengan isi berita. Fenomena ini menimbulkan dilema karena meskipun dapat meningkatkan trafik dan kunjungan, kualitas informasi menurun dan kredibilitas media menurun.

Pernyataan umum dan aktual yang ditulis dan disajikan kepada pembaca dalam surat kabar. Untuk membuat berita menarik sehingga layak dimuat dalam majalah atau surat kabar, ada beberapa pertimbangan yang harus diperhatikan. Berita dibuat dan didistribusikan. Untuk memenuhi kebutuhan pelanggan dengan layanan terbaik, wartawan harus menetapkan standar yang disepakati untuk produksi berita. Berita harus diketahui segera. Ini menunjukkan bahwa tulisan harus sederhana dan mudah dipahami. Berita harus singkat dan ringkas tanpa banyak kata [2]. Perkembangan teknologi dan media digital telah membawa perubahan signifikan dalam penyajian informasi. Kehadiran internet mempercepat distribusi berita sehingga dapat diakses dengan cepat, namun bersamaan dengan itu muncul tantangan berupa semakin maraknya praktik *clickbait*. Istilah *clickbait* merujuk pada teknik penulisan judul berita yang provokatif, hiperbolis, atau sengaja dibuat menggoda rasa penasaran agar pembaca tergoda untuk membuka tautan berita. Fenomena *clickbait* bukan hanya sekadar strategi pemasaran konten, tetapi berdampak serius terhadap kualitas informasi yang diterima publik. Banyak penelitian mengungkapkan bahwa *clickbait* mengurangi kredibilitas media, menurunkan kepuasan pembaca, serta memicu penyebaran informasi yang dangkal. Di Indonesia, sejumlah studi menunjukkan bahwa hampir sepertiga berita *online* nasional maupun lokal memanfaatkan teknik ini. Radar Cirebon, salah satu media utama Jawa Barat dengan edisi cetak maupun portal daring, juga menghadapi tantangan serupa.

Meskipun berkomitmen pada jurnalisme berkualitas, portal daringnya (www.radarcirebon.com) masih menampilkan judul berita dengan kecenderungan *clickbait*. Hingga saat ini, klasifikasi berita *clickbait* maupun

non-clickbait hanya dilakukan oleh redaksi secara manual, sehingga memakan waktu dan tenaga serta tidak selalu konsisten. Untuk mengatasi hal ini, penelitian mengusulkan pembangunan sistem klasifikasi berbasis *website* dengan algoritma *clustering*. *Clustering* dipilih karena tidak membutuhkan data label, melainkan cukup dengan menemukan kesamaan antar data untuk dikelompokkan. Algoritma ini memilih medoid sebagai representasi pusat *cluster*, bukan rata-rata, sehingga hasil pengelompokan lebih stabil pada data teks yang kompleks. Penelitian ini tidak hanya membahas penerapan teknis algoritma, tetapi juga menghadirkan sistem berbasis *web* yang praktis digunakan redaksi Radar Cirebon. Hasil keluaran berupa dua kelompok judul, *clickbait* dan *non-clickbait*, serta evaluasi kualitas *clustering* secara kuantitatif menggunakan *Davies Bouldin Index* (DBI). Dengan sistem tersebut, diharapkan Radar Cirebon mampu menyajikan berita digital yang lebih transparan sekaligus memperkuat kepercayaan publik.

Sistem organisasi adalah sistem informasinya. Selain memberikan laporan kepada pihak luar, sistem ini membantu operasi organisasi, serta kegiatan manajemen dan strategis. Selain itu, data dapat disimpan, diambil, diubah, diolah, dan dikomunikasikan oleh sistem informasi atau peralatan sistem lainnya [3]. Saat ini, judul berita sangat penting dalam industri media digital. Kehadiran *clickbait*, sebuah metode pembuatan judul yang menarik untuk mendapatkan klik, telah menjadi komponen penting dari dinamika persaingan antar media *online*. Sebuah penelitian yang dilakukan oleh *Pew Research Center* pada tahun 2023 menemukan bahwa 38% berita *online* di Indonesia menggunakan teknik *clickbait* untuk mendapatkan pengunjung. Praktik ini berkembang karena transformasi digital dan perilaku pembaca yang semakin sering mengakses informasi secara instan melalui perangkat *mobile* mereka.

Radar Cirebon didirikan oleh H. Selama Reformasi, Mahtum Mastoem dan Dahlan Iskan membangun Jawa Pos di bagian barat Jawa untuk memenuhi kebutuhan masyarakat Cirebon akan informasi tentang peristiwa lokal. Radar Cirebon telah mengalami kemajuan besar dalam hal produksi dan sarana dan prasarana. Ini ditunjukkan oleh jumlah halaman yang meningkat, oplah yang meningkat, kemajuan teknologi, dan perubahan dalam pendekatan pemasaran. Selain itu, Radar Cirebon membangun Gedung Graha Pena dan memiliki anak perusahaan bernama Rakyat Cirebon [4]. Sebagai salah satu media terbesar di kota Cirebon, Radar Cirebon menghadapi tantangan untuk mengimbangi jumlah trafik sambil tetap menjaga kredibilitas isi berita. Selama ini, pengelompokan berita belum dilakukan secara sistematis dan otomatis, yang membuatnya sulit bagi editor untuk menentukan apakah judulnya informatif atau hanya *clickbait*. Sistem berbasis *web* yang menggunakan algoritma *clustering* dan *text mining* memungkinkan editor untuk mengelompokkan judul berita secara obyektif, efisien, dan mudah digunakan oleh editor tanpa keahlian teknis khusus.

Permasalahan utama dalam penelitian ini adalah (1) bahwa berita yang diposting di *website* Radar Cirebon belum dikategorikan, sehingga sulit untuk menentukan apakah berita tersebut berupa *clickbait* atau tidak. (2) Radar Cirebon belum memiliki sistem yang dapat mengkategorikan berita secara otomatis berdasarkan kecenderungan *clickbait*.

Berita dikategorikan menurut *clickbait* atau *non-clickbait* untuk meningkatkan presentasi dan membedakan berita sensasional dari berita informatif dan akurat. Untuk meningkatkan kepercayaan pembaca dan meningkatkan pengalaman pengguna dalam mengakses berita yang sesuai dengan preferensi mereka, penelitian ini akan membangun sistem informasi untuk menggabungkan berita di Radar Cirebon berdasarkan pengelompokan *clickbait/non-clickbait*. Metode mengelompokkan data ke dalam berbagai kelompok, atau *cluster*, sehingga tingkat kemiripan data di satu kelompok lebih tinggi sementara tingkat kemiripan data di kelompok lain lebih rendah. Dengan kata lain, *clustering* dapat digunakan dalam teknik pembelajaran mesin yang tidak diawasi karena tidak membutuhkan data pelatihan, atau *training data*, selain mengiringi data [5].

Clickbait juga memiliki efek yang signifikan di tingkat lokal, terutama pada portal berita seperti Radar Cirebon. Karena media lokal harus bersaing dengan *agregator* berita dan media nasional yang memiliki sumber daya yang lebih besar, strategi *clickbait* sering digunakan untuk mempertahankan eksistensi dan meningkatkan angka kunjungan. Namun, pada akhirnya, efek jangka panjang dari *clickbait* justru menurunkan kepercayaan publik terhadap media massa. Studi menunjukkan bahwa judul *clickbait* biasanya mengandung provokasi, hiperbola, dan bahkan misinformasi kecil. Menurut penelitian psikologi media, *clickbait* dapat memengaruhi persepsi pembaca, memengaruhi emosi mereka, dan mungkin membuat mereka salah memahami isi berita. Oleh karena itu, penting bagi media lokal dan nasional untuk menekankan kredibilitas dan akurasi.

Dalam penelitian ini, algoritma *clustering* K-Medoids akan digunakan untuk mengurangi jumlah perbedaan yang ada di antara titik data dalam sebuah *cluster*. Salah satu metode *clustering*, K-Medoids menggunakan obyek yang paling sentral (medoid), rata-rata dari semua obyek dalam *cluster*, sebagai pusat *cluster* untuk nilai-nilai *cluster* dan untuk mengetahui jarak obyek dari titik data lainnya. Metode ini juga dikenal sebagai metode distribusi

[6]. Dengan menggunakan algoritma K-Medoids, tujuan dari penelitian ini adalah untuk mengembangkan sistem informasi yang dapat digunakan untuk mengelompokkan berita.

2. Metode Penelitian

2.1. Bahan Penelitian

Rancang bangun adalah proses membangun sistem, baik untuk mengganti atau memperbaiki sistem yang sudah ada secara keseluruhan atau hanya sebagian [7]. Subagyo, dikutip oleh Syamsul Bahry dan Fakhry Zamzam, mengatakan bahwa teknik penelitian adalah alat yang dapat menyelesaikan masalah apa pun. Priyono menyatakan bahwa metode penelitian adalah suatu cara melakukan sesuatu dengan menggunakan pikiran secara cermat untuk mencapai suatu tujuan [8]. Bahan yang akan digunakan dalam penelitian ini bersumber dari data primer dan data sekunder. Data primer diperoleh secara langsung diperoleh langsung dari *website* resmi Radar Cirebon melalui teknik *web scraping*. Sedangkan data sekunder diperoleh dari data yang dikumpulkan dari sumber yang telah ada, seperti buku, jurnal ilmiah, artikel, laporan penelitian sebelumnya, atau dokumen lain yang relevan dengan topik penelitian.

2.2. Teknik Pengumpulan Data

1. Observasi

Metode observasi adalah mengumpulkan dan mempelajari data dan informasi melalui pengamatan obyek data. Pada metode ini, penulis mengumpulkan serta mempelajari data dan informasi sesuai dengan jurnal yang terkait dengan penelitian.

2. Wawancara

Wawancara adalah metode sistematis untuk mengumpulkan informasi dalam bentuk pernyataan lisan tentang suatu obyek atau peristiwa saat ini, sekarang, dan masa depan. Jenis wawancara sendiri adalah terstruktur, semi terstruktur, dan tidak berstruktur. Pada tahap ini penulis memilih wawancara tidak terstruktur.

3. Studi Literatur

Studi literatur adalah serangkaian kegiatan yang berkaitan dengan membaca dan mencatat bahan penelitian, mengumpulkan data pustaka, dan mengelola bahan penelitian. Pada tahap ini, penulis melakukan dengan cara mengumpulkan informasi melalui jurnal dan *website* berita yang ada di Radar Cirebon.

2.3. Diagram Alur Tahapan Penelitian

Diagram alur tahap penelitian merupakan diagram mengenai alur dari proses pengerjaan penelitian tentang pengelompokan berita yang menggunakan algoritma K-Medoids.



Gambar 1. Diagram Alur Tahapan Penelitian

1. *Dataset*

Dataset berisikan berita dari seribu seratus tujuh puluh (1170) judul berita yang diambil dari *website* resmi Radar Cirebon pada bagian berita dan data diambil dari 01 Mei 2025 sampai dengan 31 Mei 2025.

2. *Preprocessing*

Preprocessing adalah langkah penting dalam menyiapkan data mentah agar siap digunakan dalam analisis atau pemodelan dengan cara *cleaning*, *tokenizing*, *stopword*, *stemming*, dan TF-IDF. Proses ini memastikan kualitas data yang baik sehingga hasil analisis menjadi akurat dan efektif.

3. Pengelompokan berita menggunakan algoritma K-Medoids

Algoritma K-Medoids mengelompokkan data berita berdasarkan judul dengan menggunakan medoid sebagai pusat dari setiap *cluster*. Langkah pertama adalah mempersiapkan data, termasuk mengumpulkan, membersihkan dari duplikasi, nilai kosong, dan *outlier*, serta melakukan normalisasi data. Setelah itu, tentukan berapa banyak *cluster* (K) yang diinginkan menggunakan metode tertentu seperti analisis *elbow* atau berdasarkan pengetahuan *domain*.

4. Pengujian akurasi algoritma K-Medoids

Tahap evaluasi analisa hasil dari proses pengelompokan K-Medoids untuk melihat nilai *Davies Bouldin Index* (DBI) yang didapat dalam pengelompokan ini.

2.3. Metode Penelitian

Menurut Subagyo, yang dikutip oleh Syamsul Bahry dan Fakhry Zamzam, metode penelitian adalah cara melakukan sesuatu dengan menggunakan pikiran secara cermat untuk mencapai suatu tujuan [9]. Obyek penelitian adalah judul berita yang disiarkan di Radar Cirebon (<https://www.radarcirebon.com>) dari 1 Mei 2025 hingga 31 Mei 2025. *Scraping web* digunakan untuk mengumpulkan semua judul berita dan menyimpannya sebagai *dataset* penelitian. Secara keseluruhan, 1.170 judul berita dikumpulkan sebagai data.

2.4. Metode Pengumpulan Informasi

Istilah "metode pengumpulan data" dan "instrumen pengumpulan data" saling terkait dalam penelitian. Metode pengumpulan data digunakan untuk mendapatkan informasi yang diperlukan untuk mencapai tujuan penelitian [10]. Penelitian mengumpulkan dan mengumpulkan data untuk mendapatkan informasi baru untuk mendukung informasi yang telah dikumpulkan di tempat penelitian. Akibatnya, peneliti harus benar-benar memahami metode pengumpulan data [11]. Data dikumpulkan melalui prosedur observasi dengan mempelajari struktur situs *website* Radar Cirebon untuk memahami pola URL, struktur HTML, dan elemen DOM yang memuat judul berita.

2.5. Kebutuhan Sistem

Bahan yang akan digunakan dalam penelitian ini bersumber dari data primer dan data sekunder. Data primer diperoleh secara langsung diperoleh langsung dari *website* resmi Radar Cirebon melalui teknik *web scraping*. Sedangkan data sekunder diperoleh dari data yang dikumpulkan dari sumber yang telah ada, seperti buku, jurnal ilmiah, artikel, laporan penelitian sebelumnya, atau dokumen lain yang relevan dengan topik penelitian. Perangkat lunak termasuk sistem operasi Windows 10, bahasa pemrograman Python, Visual Studio Code, *Database MySQL*, dan XAMPP. Perangkat lunak ini termasuk *preprocessing data*, *clustering nltk* untuk tokenisasi dan *stemming pandas*, *numpy* untuk manipulasi data *matplotlib*, dan *seaborn* untuk menampilkan metrik *scipy* dan *sklearn* untuk penilaian *cluster*, terutama *Davies Bouldin Index*.

2.6. Analisa Perancangan Sistem

Analisa Perancangan Sistem	
1. Analisa sistem berjalan	Sistem pengelolaan berita di Radar Cirebon saat ini kemungkinan masih manual atau belum terintegrasi dengan metode klasifikasi otomatis berdasarkan <i>clickbait</i> atau <i>non-clickbait</i> .
2. Analisa sistem yang diusulkan	Pengelompokan berita dilakukan menggunakan algoritma K-Medoids yang merupakan metode <i>clustering</i> yang efektif untuk mengelompokkan data berdasarkan kemiripan berita.
3. Pemberian informasi berita	Data yang digunakan dalam sistem ini adalah kumpulan berita dari Radar Cirebon yang akan dianalisis dan diklasifikasikan berdasarkan fitur-fitur yang merepresentasikan <i>clickbait</i> atau <i>non-clickbait</i> .

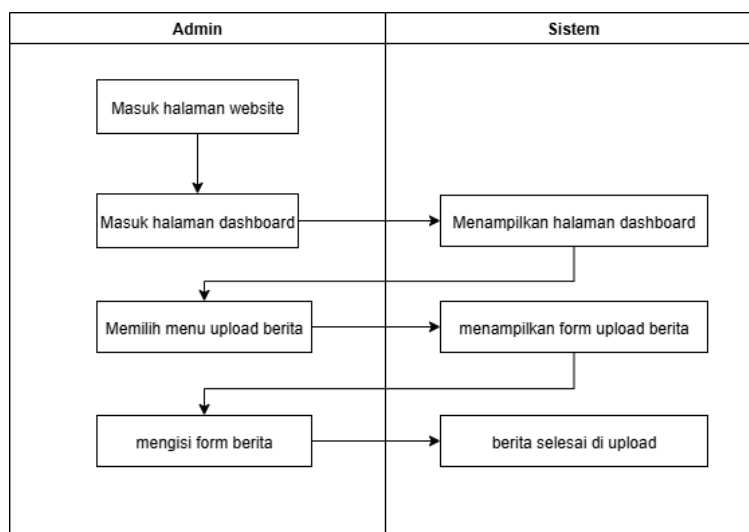
Gambar 2. Analisa Perancangan Sistem

Analisis sistem berjalan pada gambar tersebut menunjukkan pengelolaan berita Radar Cirebon masih dilakukan secara manual dan belum terotomatisasi dalam mengelompokkan berita berdasarkan *clickbait/non-clickbait*. Dengan mengotomatisasi proses klasifikasi berita dan meningkatkan efisiensi dan akurasi pengelolaan berita, sistem yang disarankan terdiri dari *website* yang menggunakan algoritma K-Medoids, menawarkan solusi yang efektif untuk masalah tersebut. Dengan menyediakan informasi data yang sistematis dan terorganisir melalui *website* ini, pengguna akan lebih mudah mengakses berita yang terpercaya dan akan mengurangi penyebaran berita yang dipromosikan. Oleh karena itu, diharapkan bahwa pengembangan sistem ini akan meningkatkan layanan informasi Radar Cirebon dan memberikan manfaat kepada pengguna dan pengelola media berita.

2.7. Analisis Sistem

Analisis sistem ini adalah *website* yang dapat mengambil judul berita dari Radar Cirebon, kemudian memproses dan mengelompokkan judul tersebut menjadi dua kategori yaitu *clickbait/non-clickbait* menggunakan algoritma K-Medoids berbasis TF-IDF.

1. Analisis Sistem Yang Berjalan

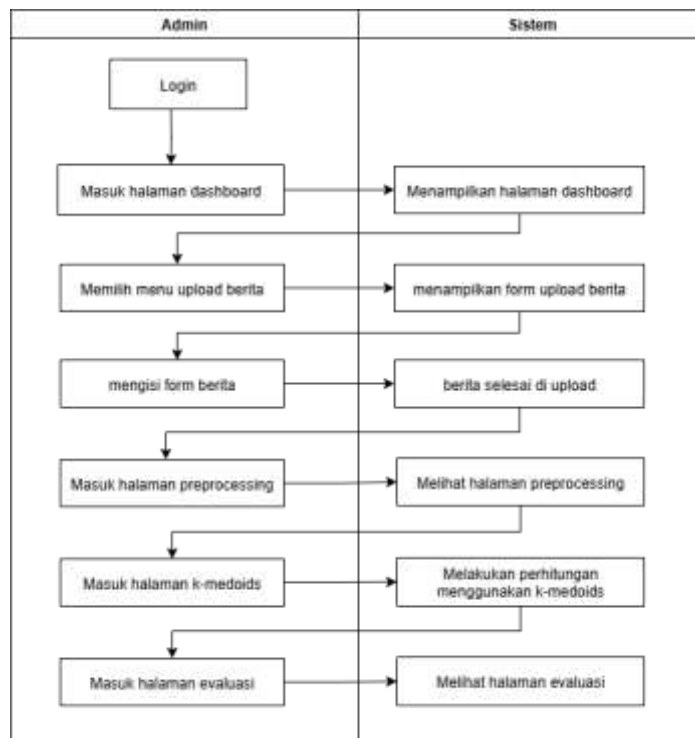


Gambar 3. Diagram Analisis Sistem Yang Berjalan

Analisis sistem yang berjalan, juga dikenal sebagai *flowmap*, menggambarkan urutan langkah-langkah atau proses yang terjadi dalam sebuah sistem. Ini dapat digambarkan dengan diagram atau peta alur, seperti yang ditunjukkan pada gambar diatas.

1. Admin mengakses halaman *website* adalah langkah pertama dalam alur sistem.
2. Setelah masuk ke *website*, admin dapat masuk ke halaman *dashboard*. Admin dapat mengelola konten atau berita.
3. Pada halaman *dashboard*, admin memilih menu untuk mengunggah berita baru.
4. Setelah memilih menu "*upload berita*", admin mengisi *form* berita untuk diunggah.
5. Di sistem, halaman *dashboard* secara otomatis ditampilkan saat admin mengakses *website*.
6. Sistem menampilkan *form upload* berita Setelah admin memilih menu *upload* berita, *form* akan ditampilkan di sistem untuk diisi.
7. Sistem menyelesaikan *upload* berita Setelah admin mengisi *form* berita, sistem memproses dan menyelesaikan *upload* berita, dan berita yang berhasil diunggah.

2. Analisis Sistem Yang Diusulkan



Gambar 4. Diagram Analisis Sistem Yang Diusulkan

Analisis sistem yang diusulkan pada diagram ini melibatkan Admin dan sistem dalam beberapa tahapan yang lebih kompleks, terutama setelah proses upload berita.

1. Admin memulai proses dengan melakukan *login* untuk mengakses sistem.
2. Setelah *login*, Admin diarahkan ke halaman *dashboard* sebagai tempat utama untuk mengelola konten dan melihat status sistem.
3. Admin memilih menu *upload* berita dari *dashboard* untuk memasukkan berita baru ke dalam sistem.
4. Admin mengisi *form* dengan informasi yang diperlukan untuk berita yang akan diunggah.
5. Setelah *login*, sistem menampilkan halaman *dashboard* yang berfungsi sebagai pusat navigasi bagi Admin.
6. Ketika Admin memilih menu *upload* berita, sistem akan menampilkan *form upload* berita yang perlu diisi oleh Admin.
7. Setelah Admin mengisi *form*, sistem akan memproses dan menyelesaikan *upload* berita yang kemudian akan ditampilkan di halaman berita.
8. Admin kemudian masuk ke halaman *preprocessing*, tempat untuk mempersiapkan data atau informasi untuk analisis lebih lanjut.
9. Sistem akan menampilkan halaman *preprocessing* setelah Admin memilih untuk masuk ke halaman tersebut.
10. Admin kemudian mengakses halaman untuk melakukan analisis dengan metode K-Medoids, yang mungkin digunakan untuk pengelompokan data atau informasi.
11. Sistem akan melakukan perhitungan menggunakan algoritma K-Medoids, yang digunakan untuk mengelompokkan data yang relevan.
12. Admin melanjutkan ke halaman evaluasi untuk melihat hasil dari perhitungan dan analisis yang dilakukan.
13. Setelah Admin memilih halaman evaluasi, sistem akan menampilkan hasil evaluasi dari proses yang telah dilakukan.

3. Hasil dan Diskusi

3.1. Gambaran Umum

Surat kabar Radar Cirebon diterbitkan setiap hari di kota Cirebon, Jawa Barat. Dimulai 20 Desember 1999. Selain menerbitkan koran harian, Radar Cirebon juga memiliki stasiun televisi lokal yang disebut Radar Cirebon TV (RCTV) yang disiarkan di wilayah Cirebon dan Indramayu. Selain itu, untuk memenuhi kebutuhan orang-orang di luar wilayah siaran, Radar Cirebon juga memiliki portal berita online dengan domain www.radarcirebon.com. Didirikan pada tahun 2002, Radar Cirebon Com merupakan portal berita *online* lokal terlengkap dan terlengkap di Indonesia [12]. *Clickbait* menjadi sebuah fenomena yang muncul dari adaptasi media terhadap era digital. Strategi ini mempengaruhi kualitas berita melalui penggunaan kata-kata yang berlebihan dan klise serta memadukan ungkapan yang ekspresif [13].

3.2. Proses *Preprocessing Data*

Preprocessing adalah tahap awal pengolahan data teks sebelum pengorganisasian. Pemrosesan teks berarti membersihkan teks dan menyederhanakannya agar dapat diproses lebih lanjut. Proses mengubah bentuk data yang tidak terstruktur menjadi data yang terstruktur sesuai kebutuhan dikenal sebagai *preprocessing*. Agar proses klasifikasi menjadi lebih efektif dan tepat, istilah yang dinilai harus diubah, dikurangi, atau dihilangkan. Ini juga harus menjelaskan isi kalimat.

Tujuannya adalah agar hasil klasifikasi sangat akurat [14]. Dalam penelitian ini, proses *preprocessing* termasuk pembersihan data, juga dikenal sebagai pembersihan data, adalah proses menghapus karakter non-alfabet seperti tanda baca, angka, simbol, dan tautan dari teks judul berita. Memecah kalimat menjadi satuan kata untuk analisis yang lebih mudah dikenal sebagai tokenisasi atau *tokenizing*. Menghapus kata umum atau *stopword* seperti "dan", "yang", "di", "ke", dan sebagainya yang tidak penting untuk mempelajari isi berita. *Stemming*, yaitu mengubah semua kata menjadi bentuk dasar mereka, seperti mengubah kata "Promosi" menjadi "promosi". Metode Frekuensi Kata *Inverse* Dokumen, atau TF IDF, menilai kata berdasarkan frekuensi yang muncul dalam satu dokumen dan dalam korpus secara keseluruhan. Metode ini digunakan untuk mengkonversi kata ke representasi numerik. Hasil dari tahap *preprocessing* menghasilkan vektor numerik yang mewakili setiap judul berita dalam bentuk matriks TF IDF. Matriks ini kemudian berfungsi sebagai input utama dalam algoritma pengelompokan.

3.3. Implementasi Algoritma K Medoids

Untuk membagi *dataset* menjadi kelompok-kelompok, gunakan metode K-medoids. Kelebihan dari pendekatan ini adalah kemampuan untuk mengatasi kelemahan dari pendekatan K-Means yang peka terhadap *outlier*. Algoritma *clustering* seperti K-medoids dan *Partitioning Around Medoids* (PAM) menggunakan obyek sebagai perwakilan (medoid) sebagai pusat *cluster*, sedangkan K-means menggunakan nilai rata-rata (mean) sebagai pusat *cluster*. K-medoids memiliki kelebihan untuk mengatasi kelemahan algoritma K-means, yaitu sensitif terhadap suara dan *outlier*, yang berarti bahwa obyek dengan nilai yang besar dapat menyimpang dari distribusi data. Kelebihan tambahan adalah hasil proses *clustering* ini tidak bergantung pada urutan data masuk [15]. K-Medoids melakukan pengelompokan dengan menggunakan medoids sebagai pusat kelompok untuk setiap kelompok [16]. Hasil percobaan menunjukkan bahwa dua *cluster* dominan. *Cluster* pertama terdiri dari berita yang ditulis dalam gaya yang sensasional dan menimbulkan rasa penasaran, sedangkan *cluster* kedua terdiri dari berita dengan judul yang lugas dan informatif.

Setelah *preprocessing* selesai, algoritma ini melakukan pengelompokan dengan memilih secara acak sejumlah medoid atau titik pusat awal dari data yang tersedia. Selanjutnya, jarak antara masing-masing medoid dihitung, dan data dibagi ke medoid terdekat. Posisi medoid diperbarui berulang kali hingga kondisi konvergen atau posisi medoid tidak berubah. Analisis *Davies Bouldin Index* dan pengamatan visual digunakan untuk menghitung jumlah *cluster* dalam penelitian ini.

3.1. Tampilan Program

1. Halaman *Login*



Gambar 5. Tampilan Halaman *Login*

Pada halaman *login* ini, admin harus memasukkan *username* dan *password* sebelum masuk.

2. Halaman *Dashboard*



Gambar 6. Tampilan Halaman *Dashboard*

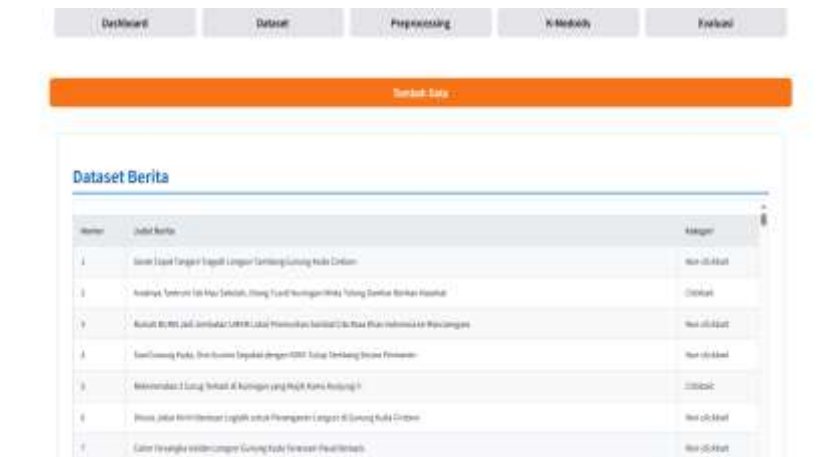
Pada halaman *dashboard*, admin dapat melihat hasil dari perhitungan K-Medoids beserta diagram di sebelahnya.



Gambar 7. Lanjutan Tampilan Halaman *Dashboard*

Pada halaman *dashboard* juga terdapat profil instansi dari Radar Cirebon beserta visi dan misi-nya.

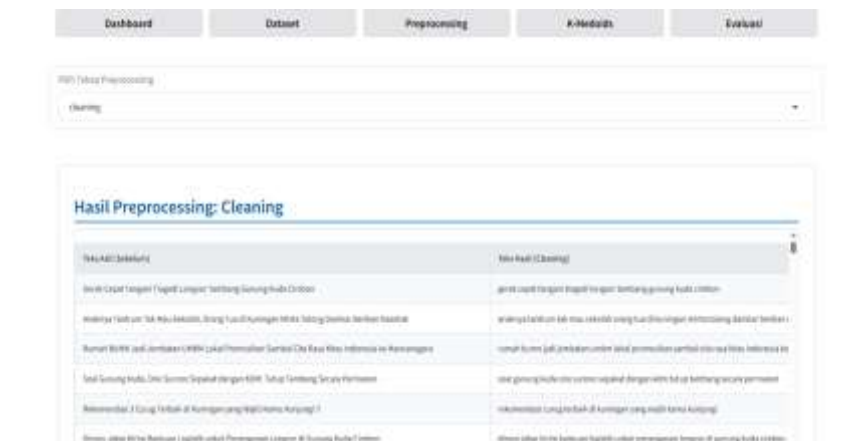
3. Halaman *Dataset*



Gambar 8. Tampilan Halaman *Dataset*

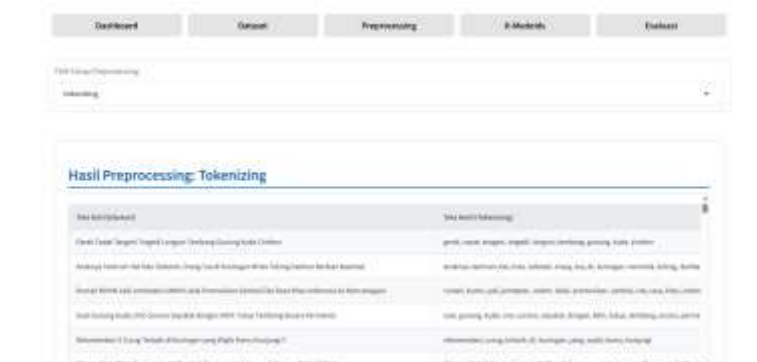
Pada halaman *Dataset*, admin bisa menambah data, *dataset* berita dengan jumlah total 1170 data dengan kategori *clickbait/non-clickbait*.

4. Halaman *Preprocessing*



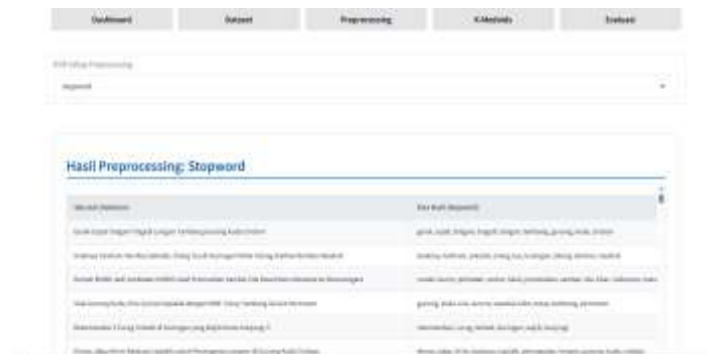
Gambar 9. Tampilan Halaman *Preprocessing Cleaning*

Pada halaman *preprocessing*, terdapat *dropdown menu* diantaranya yaitu *cleaning*, *tokenizing*, *stopword*, *stemming*, dan *TF-IDF*. *Cleaning* adalah menghilangkan url, simbol, angka, tanda baca, *hashtag*, spasi, yang berlebihan maupun karakter yang tidak ada di alphabet.



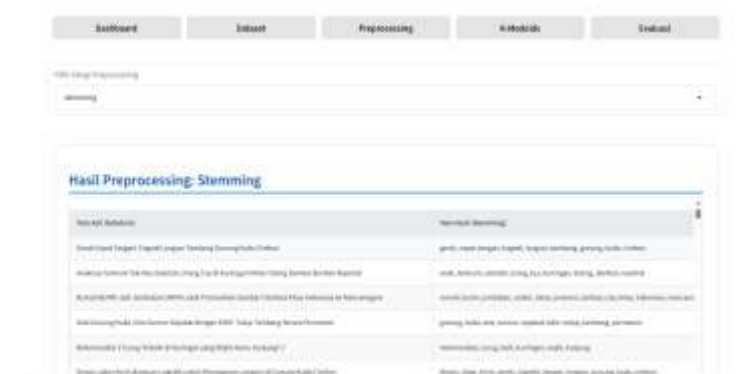
Gambar 10. Tampilan Halaman *Preprocessing Tokenizing*

Selanjutnya ada hasil dari *tokenizing*, *Tokenizing* merupakan proses pemecahan sekumpulan karakter dalam satu teks kedalam satuan kata, pada proses ini juga dapat menghilangkan karakter pembatas menghapus angka dan menghapus tanda baca atau karakter yang bukan merupakan huruf.



Gambar 11. Tampilan Halaman *Preprocessing Stopword*

Selanjutnya ada hasil dari *stopword*, *Stopword* merupakan penghapusan kata tidak penting dari deskripsi dilakukan dengan memeriksa apakah kata-kata yang dihasilkan dari *parsing* deskripsi termasuk dalam daftar kata tidak penting.



Gambar 12. Tampilan Halaman *Preprocessing Stemming*

Selanjutnya ada hasil dari *stemming*, *stemming* merupakan tahapan membentuk teks menjadi kata dasar atau dikenal dengan istilah *term*.



Gambar 13. Tampilan Halaman *Preprocessing TF-IDF*

Selanjutnya ada hasil dari TF-IDF, TF-IDF merupakan ekstraksi data menjadi tipe data numerik, dikarenakan TF-IDF ini mengubah data teks menjadi vektor numerik dengan menghitung frekuensi

kemunculan data dalam dokumen dan mengurangi bobot kata yang sering muncul dalam seluruh dokumen.

5. Halaman K-Medoids

Dashboard	Dataset	Preprocessing	K-Medoids	Evaluasi
Hasil K-Medoids				
Berikut adalah hasil clustering K-Medoids dari file hasil kmedoids.xlsx.				
Cent 1	Cent 2	Kelompok	Cluster	
0.00000	1.113199	0.00000	1.0	
0.10877	0.004760	0.004760	2.0	
0.08789	1.388181	0.08789	1.0	
0.78373	1.388	0.78373	1.0	
0.441011	1.777721	0.441011	1.0	
0.70080	1.940102	0.70080	1.0	
0.70080	1.940102	0.70080	1.0	

Gambar 14. Tampilan Halaman K-Medoids

Pada halaman K-Medoids, menunjukkan hasil dari proses *clustering* yang didapat dari iterasi medoid. Pada tahap kedekatan menunjukkan hasil dari nilai jarak terdekat obyek ke medoid sementara.

6. Halaman Evaluasi

Dashboard

Dataset

Preprocessing

K-Medoids

Evaluasi

Evaluasi Model Clustering

Evaluasi dilakukan pada data uji menggunakan Davies-Bouldin Index.

Skor: 0,898

Davies-Bouldin Index mengukur kualitas clustering. Skor yang lebih rendah menunjukkan pemisahan cluster yang lebih baik dan lebih padat.

Hasil Data Evaluasi

No	Nilai	Cluster	Unweighted 0	Cluster 1	Data	Unweighted 0	Unweighted 1	Unweighted 0	Unweighted 0	Unweighted 10	Unweighted 11	Unweighted 12	Unweighted 13
1	0.00000	0		01	0.00000	0.00789	1.70373	1.401011	0.70080	0.70080	1.123189	1.00338	0.81011
2	0.00789	0		02	0.00789	1.007120	1.400343	1.107120	1.027791	1.000443	1.107797	1.340021	1.107940
3	0.70373	0											
4	0.441011	0											

Gambar 15. Tampilan Halaman Evaluasi

Pada halaman Evaluasi, menunjukkan hasil dari perhitungan dengan menggunakan perhitungan *Davies Bouldin Index* (DBI). Desain yang memungkinkan pengguna untuk mengelola dan memanipulasi data secara efisien.

3.2. Hasil Pengujian Program

Teknik pengujian digunakan pada algoritma K-Medoids untuk mengevaluasi efektivitas dan akurasi hasil *clustering* yang dihasilkannya. Tujuan dari pengujian ini adalah untuk memastikan bahwa algoritma tersebut

memiliki kemampuan untuk mengelompokkan data dengan efektif dan menghasilkan *cluster* yang bermakna.

Proses pengujian DBI dimulai dengan menghitung jarak antar *cluster* untuk menilai pemisahan antar pusat *cluster*. Selanjutnya, dihitung juga kepadatan setiap *cluster*, yaitu jarak rata-rata titik data dalam *cluster* ke medoidnya. DBI dihitung dengan memperhitungkan kedua faktor ini, dan semakin kecil nilai DBI, semakin baik kualitas *clustering* karena menunjukkan bahwa *cluster* lebih padat dan terpisah dengan jelas. Hasil pengujian DBI yang diperoleh adalah 0.898, yang menunjukkan kualitas pengelompokan yang cukup baik. Nilai ini mengindikasikan pemisahan antar *cluster* yang cukup jelas dan konsentrasi data yang baik dalam masing-masing *cluster*.

4. Kesimpulan

Berdasarkan hasil dan analisa serta implementasi sistem maka didapatkan kesimpulan sebagai berikut: 1. Dalam penelitian ini algoritma K-Medoids digunakan untuk membangun sistem pengelompokan berita di Radar Cirebon. Jumlah *cluster* C1 dan C2 didapat dari total *cluster* pada iterasi medoid dari data ke-1 sampai data ke-618. Sistem ini secara otomatis mengelompokkan berita ke dalam kategori *clickbait* dan *non-clickbait*, yang membantu mengelola berita dengan lebih baik dan memudahkan pembaca untuk mencari informasi yang relevan dan sesuai dengan preferensi mereka. 2. Berdasarkan hasil dari evaluasi *Davies Bouldin Index* (DBI) yang mengukur kualitas klasterisasi. Nilai DBI yang lebih rendah menunjukkan *cluster* yang lebih baik. Nilai hasil DBI adalah 0.898, yang menunjukkan kualitas klasterisasi yang cukup baik. Nilai DBI ini bisa dijadikan acuan untuk evaluasi hasil klasterisasi menggunakan algoritma K-Medoids. 3. Radar Cirebon mendapatkan banyak manfaat dari sistem ini, terutama dalam hal meningkatkan kredibilitas berita *clickbait*. Sistem yang dibangun ini membantu Radar Cirebon untuk secara otomatis mengelompokkan berita menjadi kategori *clickbait* dan *non-clickbait*, sehingga dapat meningkatkan kepercayaan pembaca terhadap konten yang disajikan. Dengan menggunakan algoritma K-Medoids, pengelompokan berita dapat dilakukan dengan lebih efisien, yang memungkinkan pengelola media untuk menghindari judul yang menyesatkan dan memastikan bahwa berita yang dipublikasikan relevan dengan isi konten.

Referensi

- [1] A. N. Adi Manggolo, S. Siswanto, and M. Musthofa, "Strategi Peliputan Berita Wartawan Majalah Suluh," *J. Dakwah dan Komun.*, vol. 5, no. 2, p. 101, 2020, doi: 10.29240/jdk.v5i2.2150.
- [2] S. Julianti, "Bab V - 1 Bab V - 2," *Ekon. Islam.*, vol. 11, no. variabel X, pp. 46–47, 2019.
- [3] I. Zulvia, F. A. Hidayatulloh, and E. Rahmawati, "Analisis Kepuasan Pasien Terhadap Pelayanan Kesehatan di Klinik Alkindi Herbal Menggunakan Metode K-Means Clustering," *JIKO (Jurnal Inform. dan Komputer)*, vol. 6, no. 2, p. 261, 2022, doi: 10.26798/jiko.v6i2.612.
- [4] Y. Diana *et al.*, "Analisa Penjualan Menggunakan Algoritma K-Medoids Untuk Mengoptimalkan Penjualan Barang," *JOISIE J. Inf. Syst. Informatics Eng.*, vol. 7, no. 1, pp. 97–103, 2023.
- [5] A. Supriyadi, A. Triayudi, and I. D. Sholihati, "Perbandingan Algoritma K-Means Dengan K-Medoids Pada Pengelompokan Armada Kendaraan Truk Berdasarkan Produktivitas," *JIPi (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 6, no. 2, pp. 229–240, 2021, doi: 10.29100/jipi.v6i2.2008.
- [6] N. Widiawati, B. N. Sari, and T. N. Padilah, "Clustering Data Penduduk Miskin Dampak Covid-19 Menggunakan Algoritma K-Medoids," *J. Appl. Informatics Comput.*, vol. 6, no. 1, pp. 55–63, 2022, doi: 10.30871/jaic.v6i1.3266.
- [7] F. D. Putra, J. Riyanto, and A. F. Zulfikar, "Rancang Bangun Sistem Informasi Manajemen Aset pada Universitas Pamulang Berbasis WEB," *J. Eng. Technol. Appl. Sci.*, vol. 2, no. 1, pp. 32–50, 2020, doi: 10.36079/lamintang.jetas-0201.93.
- [8] M. . Dr. Manotar Tampubolon, S.H., M.A., Metode Penelitian Metode Penelitian, vol. 3, no.17. 2023. [Online]. Available: [http://repository.unpas.ac.id/30547/5/BAB III.pdf](http://repository.unpas.ac.id/30547/5/BAB%20III.pdf)
- [9] A. Azizah, "Strategi Konvergensi Media Cetak pada Surat Kabar Radar Cirebon," Skripsi, pp. 1–156, 2021.
- [10] M. Nafisatur, "Metode Pengumpulan Data Penelitian," Metod. Pengumpulan Data Penelit., vol. 3, no. 5, pp. 5423–5443, 2024.
- [11] M. Nafisatur, "Metode Pengumpulan Data Penelitian," Metod. Pengumpulan Data Penelit., vol. 3, no. 5, pp. 5423–5443, 2024.
- [12] F. Syarifah, "Analisis framing berita persoalan pengelolaan sampah di kota cirebon pada laman media online radarcirebon.com, kabarcirebon.com, dan tribuncirebon.com," Skripsi, pp. 1–68, 2024, [Online]. Available: <https://dspace.uui.ac.id/handle/123456789/50930%0Ahttps://dspace.uui.ac.id/bitstream/handle/123456789/50930/17321129.pdf?sequence=1>
- [13] R. A. Sonya, "No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析 Title," *Otonomi*, vol. 20, pp. 396–406, 2020.
- [14] N. Aula, M. Ula, and L. Rosnita, "Analisis Sentimen Review Customer Terhadap Perusahaan Ekspedisi Jnc, J&T Express Dan Pos Indonesia Menggunakan Metode Support Vector Machine (Svm)," *J. Informatics Comput. Sci.*, vol. 9, no. 1, p. 81, 2023, doi: 10.33143/jics.v9i1.2947.
- [15] S. B. Jawa, "Aji Prasetya Wibawa, Farid Miftahuddin dan Suyono," vol. 10, no. 2, pp. 261–269, 2021.
- [16] R. Gustrianda and D. I. Mulyana, "Penerapan Data Mining Dalam Pemilihan Produk Unggulan dengan Metode Algoritma K-Means Dan K-Medoids," *J. Media Inform. Budidarma*, vol. 6, no.1, p. 27, 2022, doi: 10.30865/mib.v6i1.3294.