

BAB II

TINJAUAN PUSTAKA

Bagian ini menguraikan penelitian-penelitian terdahulu yang dijadikan sebagai bahan referensi penulis dalam penelitian ini. Penulis melakukan refleksi terhadap penelitian-penelitian terdahulu yang dibahas dalam makalah, mencoba melakukan perbandingan untuk menemukan perbedaan dan persamaan dalam penelitian ini. Dengan demikian dapat diketahui keunggulan penelitian penulis dibandingkan penelitian-penelitian sebelumnya.

Penelitian ini dilakukan oleh Fathir pada tahun 2023 dengan judul "Analisis Sentimen Artikel Pemilu Berdasarkan Metodologi Taksonomi " untuk mengevaluasi efektivitas metode taksonomi dalam memberikan analisis sentimen pada artikel pemilu. Penelitian ini menggunakan metode SMOTE untuk mendukung kinerja algoritme klasifikasi dalam menangani data yang tidak seimbang. algoritme Naïve Bayes menghasilkan akurasi sebesar 84,44%, sedangkan Support Vector Machine memperoleh akurasi sebesar 90,70%. Pada pengujian kedua, yaitu tanpa menggunakan algoritme SMOTE, algoritme yang memberikan akurasi tertinggi adalah Support Vector Machine dengan nilai akurasi sebesar 92,05%, recall sebesar 56,26%, recovery 43,72%, dan nilai F1 sebesar 46,97%. Sedangkan algoritme Naïve Bayes mencapai akurasi 89,98% dan algoritme Random Forest mencapai akurasi 90,78% [4].

Penelitian ini dilakukan oleh Ayu Kadek Nadya Oktaviana, Ngurah Agus Sanjaya, Ida Bagus Made Mahendra, I Gede Santi Astawa, I Gede Arta Wibawa, dan I Komang Ari Mogi pada tahun 2023 dengan judul "Pemodelan Topik Artikel Berita Menggunakan Structural Topic Modeling dan Latent Dirichlet Allocation". Hasil penelitian menunjukkan bahwa metode LDA menghasilkan 8 topik berdasarkan perhitungan nilai kombinasi probabilitas tertinggi. Metode STM menghasilkan 11 topik

berdasarkan konsistensi semantik dan nilai eksklusivitas tertinggi. Di antara hasil dua subjek untuk masing-masing metode, LDA menghasilkan 2 dari 8 subjek yang tidak mudah diinterpretasikan. Sementara itu, STM mencoba memfasilitasi penafsiran topik dengan mengkaji eksklusivitas topik. Dalam metode STM terdapat dua pasang topik yang saling berkaitan, yaitu Topik 11 dan Topik 2 yang mempunyai kesamaan dalam permasalahan ekonomi, serta Topik 8 dan Topik 4 yang memiliki kesamaan mengenai pemerintahan [16].

Penelitian ini dilakukan oleh Alifiyah Rohmatul Hidayati, Arief Senja Fitriani, dan Mochamad Alfian Rosid pada tahun 2022 dengan judul "Analisa Sentimen Pemilu 2019 Pada Judul Berita Online Menggunakan Metode Logistic Regression". Hasil penelitian ini mencakup semua tahapan mulai dari preprocessing hingga processing dan evaluasi. Hasil yang diperoleh pada penelitian ini mencapai akurasi sebesar 86%. Kombinasi metode logistic regression dan randomized search cross-validation menunjukkan skor akurasi sebesar 86% [1].

Penelitian ini dilakukan oleh Sely Karmila dan Visda Intan Ardianti pada tahun 2022 dengan judul "Metode Latent Dirichlet Allocation Untuk Menentukan Topik Teks Suatu Berita". Hasil penelitian menunjukkan bahwa metode Latent Dirichlet Allocation (LDA) digunakan untuk memodelkan topik hangat dari detik.com dengan berbagai kategori. Penelitian ini menciptakan 3 model topik di mana hasil pemodelan ditampilkan secara grafis untuk mengidentifikasi tren topik yang sedang menjadi berita. Hasil pengujian relevansi judul dengan subjek menghasilkan akurasi sebesar 67% [15].

Penelitian ini dilakukan oleh Ihsan Ahsanu Amala, Donni Richasdy, dan Mahendra Dwi Febri Purbolaksono pada tahun 2022 dengan judul "Telkom University News Topic Modeling Using Latent Semantic Analysis (LSA) Method on Online News Portal". Hasil penelitian menunjukkan bahwa metode *Latent Semantic Analysis* (LSA) digunakan untuk menganalisis data berupa judul berita yang berhubungan

dengan Telkom University. Proses ini dilakukan dengan penerapan penghapusan stopwords dan stemming. Jumlah topik yang optimal adalah enam dengan skor koherensi sebesar 0,524. Hal ini dikarenakan stemming menghilangkan imbuhan kata dan penghapusan stopwords menghilangkan kata yang sering muncul namun tidak bermakna. Kesimpulannya, terdapat enam topik yang diberitakan di portal berita online tentang Telkom University, di mana topik satu dan topik dua sama-sama bertemakan pemeringkatan [19].



Judul	Peneleti	Tahun	Metode	Hasil
Analisis Sentimen Artikel Pemilu Berdasarkan Metodologi Taksonomi	Fathir	2023	Metode SMOTE	Algoritme yang memberikan akurasi tertinggi adalah Support Vector Machine dengan nilai akurasi sebesar 92,05%, recall sebesar 56,26%, recovery 43,72%, dan nilai F1 sebesar 46,97%. Sedangkan algoritma Naïve Bayes mencapai akurasi 89,98% dan algoritma Random Forest mencapai akurasi 90,78%.
Analisa Sentimen Pemilu 2019 Pada Judul Berita Online Menggunakan Metode Logistic Regressionn	Alifiyah Rohmatul Hidayati, Arief Senja Fitriani, dan Mochamad Alfan Rosid	2022	Metode Logistic Regression	Hasil yang diperoleh pada penelitian ini mencapai akurasi sebesar 86%. Kombinasi metode logistic regression dan randomized search cross-validation menunjukkan skor akurasi sebesar 86%
Telkom University News Topic Modeling Using Latent Semantic Analysis	Ihsan Ahsanu Amala, Donni Richasdy, dan	2022	Metode Latent Semantic Analysis (LSA)	Jumlah topik yang optimal adalah enam dengan skor koherensi sebesar 0,524. Kesimpulannya, terdapat enam

(LSA) Method on Online News Portal	Mahendra Dwi Febri Purbolaksono			topik yang diberitakan di portal berita online tentang Telkom University, di mana topik satu dan topik dua sama-sama bertemakan pemeringkatan.
Metode Latent Dirichlet Allocation Untuk Menentukan Topik Teks Suatu Berita	Sely Karmila dan Visda Intan Ardianti	2023	Metode Latent Dirichlet Allocation (LDA)	Penelitian ini menciptakan 3 model topik di mana hasil pemodelan ditampilkan secara grafis untuk mengidentifikasi tren topik yang sedang menjadi berita. Hasil pengujian relevansi judul dengan subjek menghasilkan akurasi sebesar 67%
Pemodelan Topik Artikel Berita Menggunakan Structural Topic Modeling dan Latent Dirichlet Allocation	Ayu Kadek Nadya Oktaviana, Ngurah Agus Sanjaya, Ida Bagus Made Mahendra, I	2023	Metode STM dan metode LDA	Hasil penelitian menunjukkan bahwa metode LDA menghasilkan 8 topik berdasarkan perhitungan nilai kombinasi probabilitas tertinggi. Metode STM menghasilkan 11 topik berdasarkan konsistensi semantik dan nilai eksklusivitas tertinggi.

	Gede Santi Astawa, I Gede Arta Wibawa, dan I Komang Ari Mogi			
Analisis Sentimen dan Topik Berita Terkait Pemilu 2024 Di Indonesia	Penulis	2024	Metode TF-IDF, Bag of Words (BoW) dan Kombinasi keduanya dan Metode LDA dan LSA untuk pemodelan topik	Metode yang digunakan kombinasi TF- IDF dan bag of words (BoW) yang paling tinggi pada algoritme SVM dengan nilai akurasi 90.89% dan F1 skor 90.86%. Hasil jumlah topik tertinggi untuk LDA adalah satu dengan coherence score 0.6089501917687159 topik. LSA adalah satu topik dengan coherence score 0.5654880904606634

Tabel 2. 1 Tabel Perbandingan Tinjauan Pustaka