

# PENGEMBANGAN SISTEM REKOMENDASI KONTEN WEBSITE BERITA MENGGUNAKAN ALGORITMA NAÏVE BAYES

Rizqy Khoirul Waritsin<sup>1</sup>

<sup>1</sup>Informatika, Universitas Pembangunan Nasional “VETERAN” Jawa Timur  
[20081010081@student.upnjatim.ac.id](mailto:20081010081@student.upnjatim.ac.id)

## **ABSTRAK**

Dalam era digital yang dipenuhi dengan informasi melimpah, tantangan utama yang dihadapi oleh pengguna internet adalah menemukan berita yang sesuai dengan minat dan preferensi mereka. Sistem rekomendasi berita menjadi solusi krusial untuk membantu pengguna menyaring informasi yang relevan di antara banyaknya konten yang tersedia. Penelitian ini bertujuan untuk mengembangkan Sistem Rekomendasi Konten Website Berita dengan menggunakan Algoritma Naïve Bayes. Fokus penelitian melibatkan pengukuran tingkat personalisasi pada sistem rekomendasi berita yang ada, implementasi algoritma Naïve Bayes untuk memahami preferensi pengguna, dan analisis kemampuan algoritma dalam mengatasi asumsi kemandirian fitur. Pertanyaan penelitian mencakup langkah-langkah untuk meningkatkan akurasi rekomendasi dengan Naïve Bayes dan dampak penggunaannya terhadap pengalaman dan retensi pengguna di website berita. Teori-teori terkait, seperti Sistem Rekomendasi Berita, Algoritma Naïve Bayes, dan konsep personalisasi, menjadi dasar dalam merancang dan mengimplementasikan solusi. Metode penelitian mencakup studi literatur, analisis kebutuhan, pengumpulan dataset, implementasi algoritma, analisis tingkat personalisasi, dan pengujian aplikasi. Hasil penelitian diharapkan dapat memberikan kontribusi dalam meningkatkan tingkat personalisasi rekomendasi berita, memberikan pemahaman yang lebih baik tentang preferensi pengguna, dan mengoptimalkan pengalaman pengguna di website berita.

**Kata Kunci** : Sistem Rekomendasi Berita, Algoritma Naïve Bayes, Personalisasi

## **ABSTRACT**

*In the digital era filled with abundant information, the primary challenge faced by internet users is to discover news that aligns with their interests and preferences. News recommendation systems have become a crucial solution to assist users in filtering relevant information from the vast pool of available content. This research aims to develop a Content News Recommendation System for Websites using the Naïve Bayes Algorithm. The research focuses on measuring the level of personalization in existing news recommendation systems, implementing the Naïve Bayes algorithm to understand user preferences, and analyzing the algorithm's ability to address feature independence assumptions. Research questions include steps to enhance recommendation accuracy with Naïve Bayes and the impact of its implementation on user experience and retention on news websites. Relevant theories such as News Recommendation Systems, Naïve Bayes Algorithm, and the concept of personalization form the basis for designing and implementing the solution. The research methodology involves literature review, needs analysis, dataset collection, algorithm implementation, personalization level analysis,*

*and application testing. The research results are expected to contribute to improving the personalization level of news recommendations, providing a better understanding of user preferences, and optimizing the user experience on news websites.*

**Keywords** : News Recommendation System, Naïve Bayes Algorithm, Personalization

## **1. PENDAHULUAN**

Dalam era digital yang penuh dengan informasi yang berlimpah, pengguna internet sering kali menghadapi tantangan dalam menemukan berita yang relevan dengan minat dan preferensi mereka. Website berita dan platform media sosial telah menjadi sumber utama berita bagi banyak orang, tetapi ketersediaan berita yang begitu banyak membuat sulit untuk memilah informasi yang paling relevan dan bermanfaat. Sistem rekomendasi berita telah menjadi solusi yang semakin penting untuk membantu pengguna menavigasi hiruk-pikuk informasi dan menemukan berita yang paling sesuai dengan minat mereka. Namun, keberhasilan sistem rekomendasi ini seringkali tergantung pada sejauh mana personalisasi yang dapat mereka tawarkan kepada pengguna. Dalam banyak kasus, personalisasi ini belum mencapai tingkat yang memuaskan. Algoritma Naïve Bayes adalah salah satu alat yang digunakan dalam machine learning untuk membuat prediksi berdasarkan data. Dalam konteks rekomendasi berita, penggunaan Naïve Bayes dapat memberikan pendekatan yang menarik untuk meningkatkan tingkat personalisasi rekomendasi berita.

## **2. METODE PENELITIAN**

### **2.1. Studi Literatur**

Studi literatur melibatkan pencarian, pemahaman, dan sintesis literatur yang relevan dengan topik penelitian. Dari sudut pandang penelitian, ini mencakup memahami konsep sistem rekomendasi berita, algoritma Naïve Bayes, dan personalisasi dalam konteks rekomendasi berita. Studi literatur membantu membangun landasan teoritis dan memahami perkembangan terkini dalam bidang ini.

### **2.2. Analisis Kebutuhan**

Analisis kebutuhan adalah langkah penting dalam merancang solusi yang efektif. Analisis kebutuhan mencakup pemahaman mendalam tentang pengguna website berita, tujuan dari sistem rekomendasi, dan preferensi pengguna. Ini melibatkan wawancara dengan pengguna potensial, pemilik website berita, dan pemahaman yang lebih baik tentang apa yang diharapkan dari sistem rekomendasi.

### **2.3. Pengumpulan Dataset**

Pada tahap pengumpulan dataset, nantinya akan dilakukan pengumpulan data tentang beberapa kategori berita favorit tiap pengguna dan rekam jejak berita yang telah dibaca oleh pengguna. Data tersebut akan dikumpulkan dan diklasifikasikan ke beberapa kategori gabungan antara rekam jejak dan kategori berita favorit pengguna. Tahap ini sangat diperlukan untuk melatih dan menguji model algoritma Naïve Bayes. Pengumpulan dataset adalah fondasi penting untuk pengembangan model.

### **2.4. Implementasi Algoritma Naïve Bayes di Sistem Website**

Implementasi algoritma Naïve Bayes adalah langkah teknis dalam pengembangan solusi. Dalam tahapan ini sistem website berita akan diintegrasikan dengan model yang dibuat berdasarkan dataset yang telah dikumpulkan. Tahapan Ini melibatkan perubahan teknis dalam infrastruktur website untuk menghasilkan rekomendasi berita yang lebih personal sesuai dengan preferensi pengguna.

### **2.5. Analisis Tingkat Personalisasi**

Analisis tingkat personalisasi adalah evaluasi terhadap sistem rekomendasi berita yang sudah ada di website. Tahapan ini mencakup penggunaan data yang telah dikumpulkan dan metrik yang sesuai untuk mengukur sejauh mana sistem rekomendasi saat ini memenuhi preferensi pengguna. Hasil analisis ini akan menjadi pembandingan dengan sistem yang akan diimplementasikan.

### **2.6. Testing Aplikasi**

Testing aplikasi adalah tahap penting dalam memvalidasi solusi yang diimplementasikan. Dari perspektif penelitian, ini mencakup pengujian aplikasi yang telah diterapkan dengan pengguna sebenarnya. Tujuannya adalah untuk mengukur sejauh mana sistem menghasilkan rekomendasi yang lebih personal, mengukur akurasi rekomendasi, serta memahami dampak penggunaan sistem baru terhadap pengalaman pengguna dan retensi mereka di website berita.

## **3. CONCLUSIONS**

### **3.1. Pengumpulan Data**

Pengumpulan data dilakukan secara online dengan cara mencari data yang sudah dikumpulkan di website Kaggle. Data ini berisi tentang berita, kategori berita, *headline*, deskripsi pendek, dan tanggal berita diupload. Data berjumlah sekitar 200000 data di dalam file dengan extensi *json*. Berikut tampilan sekilas tentang dataset yang digunakan

```

8675 {"category": "CRIME", "headline": "There Were 2 Mass Shootings In Texas Last Week, But Only 1 On TV", "authors": "Melissa Jeltsen", "link": "https://
8676 {"category": "ENTERTAINMENT", "headline": "Will Smith Joins Diplo And Nicky Jam For The 2018 World Cup's Official Song", "authors": "Andy McDonald", "link": "https://www.huffingto
8677 {"category": "ENTERTAINMENT", "headline": "Hugh Grant Marries For The First Time At Age 57", "authors": "Ron Dicker", "link": "https://www.huffingto
8678 {"category": "ENTERTAINMENT", "headline": "Jim Carrey Blasts 'Castrato' Adam Schiff And Democrats In New Artwork", "authors": "Ron Dicker", "link": "https://www.huffingto
8679 {"category": "ENTERTAINMENT", "headline": "Julianne Margulies Uses Donald Trump Poop Bags To Pick Up After Her Dog", "authors": "Ron Dicker", "link": "https://www.huffingto
8680 {"category": "ENTERTAINMENT", "headline": "Morgan Freeman 'Devastated' That Sexual Harassment Claims Could Undermine Legacy", "authors": "Ron Dicker", "link": "https://www.huffingto
8681 {"category": "ENTERTAINMENT", "headline": "Donald Trump Is Lovin' New McDonald's Jingle In 'Tonight Show' Bit", "authors": "Ron Dicker", "link": "https://www.huffingto
8682 {"category": "ENTERTAINMENT", "headline": "What To Watch On Amazon Prime That's New This Week", "authors": "Todd Van Luling", "link": "https://www.huffingto
8683 {"category": "ENTERTAINMENT", "headline": "Mike Myers Reveals He'd 'Like To' Do A Fourth Austin Powers Film", "authors": "Andy McDonald", "link": "https://www.huffingto
8684 {"category": "ENTERTAINMENT", "headline": "What To Watch On Hulu That's New This Week", "authors": "Todd Van Luling", "link": "https://www.huffingto
8685 {"category": "ENTERTAINMENT", "headline": "Justin Timberlake Visits Texas School Shooting Victims", "authors": "Sebastian Murdock", "link": "https://www.huffingto
8686 {"category": "WORLD NEWS", "headline": "South Korean President Meets North Korea's Kim Jong Un To Talk Trump Summit", "authors": "", "link": "https://www.huffingto
8687 {"category": "IMPACT", "headline": "With Its Way Of Life At Risk, This Remote Oyster-Growing Region Called In Robots", "authors": "Karen Pinchin", "link": "https://www.huffingto
8688 {"category": "POLITICS", "headline": "Trump's Crackdown On Immigrant Parents Puts More Kids In An Already Strained System", "authors": "Elise Foley", "link": "https://www.huffingto
8689 {"category": "POLITICS", "headline": "Trump's Son Should Be Concerned: FBI Obtained Wiretaps Of Putin Ally Who Met With Trump Jr.", "authors": "Mary Papenfuss", "link": "https://www.huffingto
8690 {"category": "POLITICS", "headline": "Edward Snowden: There's No One Trump Loves More Than Vladimir Putin", "authors": "Mary Papenfuss", "link": "https://www.huffingto
8691 {"category": "POLITICS", "headline": "Booyah: Obama Photographer Hilariously Trolls Trump's 'Spy' Claim", "authors": "Mary Papenfuss", "link": "https://www.huffingto
8692 {"category": "POLITICS", "headline": "Ireland Votes To Repeal Abortion Amendment In Landslide Referendum", "authors": "Laura Bassett", "link": "https://www.huffingto
8693 {"category": "POLITICS", "headline": "Ryan Zinke Looks To Reel Back Some Critics With 'Grand Pivot' To Conservation", "authors": "Chris D'Angelo", "link": "https://www.huffingto
8694 {"category": "POLITICS", "headline": "Trump's Scottish Golf Resort Pays Women Significantly Less Than Men: Report", "authors": "Mary Papenfuss", "link": "https://www.huffingto
8695 {"category": "WEIRD NEWS", "headline": "Weird Father's Day Gifts Your Dad Doesn't Know He Wants (But He Does)", "authors": "David Moye", "link": "https://www.huffingto
8696 {"category": "ENTERTAINMENT", "headline": "Twitter #PutStarWarsInOtherFilms And It Was Universally Entertaining", "authors": "Andy McDonald", "link": "https://www.huffingto
8697 {"category": "WEIRD NEWS", "headline": "Mystery 'Wolf-Like' Animal Reportedly Shot In Montana, Baffles Wildlife Officials", "authors": "Wilary Hanso", "link": "https://www.huffingto
8698 {"category": "WORLD NEWS", "headline": "North Korea Still Open To Talks After Trump Cancels Summit", "authors": "Josh Smith and Christine Kim, Reute", "link": "https://www.huffingto
8699 {"category": "WORLD NEWS", "headline": "2 Men Detonate Bomb Inside Indian Restaurant Near Toronto, Authorities Say", "authors": "", "link": "https://www.huffingto
8700 {"category": "WORLD NEWS", "headline": "Thousands Travel Home To Ireland To Vote On Abortion Access", "authors": "Antonia Blumberg", "link": "https://www.huffingto
8701 {"category": "WORLD NEWS", "headline": "Irish Voters Set To Liberalize Abortion Laws In Landslide, Exit Poll Signals", "authors": "", "link": "https://www.huffingto
8702 {"category": "BLACK VOICES", "headline": "Warriors Coach Steve Kerr Calls NFL Ban On Protests 'Fake Patriotism'", "authors": "Sebastian Murdock", "link": "https://www.huffingto
8703 {"category": "BLACK VOICES", "headline": "In Historic Victory, Barbados Elects First Female Prime Minister", "authors": "Kenya Downs", "link": "https://www.huffingto

```

Gambar 3.1 Dataset

## 3.2. Hasil Penelitian

Hasil dari penelitian terdapat pada tampilan html yang menampilkan berita sesuai dengan kategori yang dipilih oleh pengguna. Berikut tampilan program yang telah diteliti.

### News Recommendation

Choose a news category:

Gambar 3.2 Tampilan input kategori

## Recommended News in Category: CRIME

- **Biden Says U.S. Forces Would Defend Taiwan If China Invaded**  
Category: *POLITICS*  
President issues vow as tensions with China rise.
- **'Beautiful And Sad At The Same Time': Ukrainian Cultural Festival Takes On A Deeper Meaning This Year**  
Category: *POLITICS*  
An annual celebration took on a different feel as Russia's invasion dragged into Day 206.
- **Biden Says Queen's Death Left 'Giant Hole' For Royal Family**  
Category: *POLITICS*  
U.S. President Joe Biden, in London for the funeral of Queen Elizabeth II, says his heart went out to the royal family, adding the queen's death left a "giant hole."
- **Bill To Help Afghans Who Escaped Taliban Faces Long Odds In The Senate**  
Category: *POLITICS*  
Republican outrage over the shoddy U.S. withdrawal from Afghanistan hasn't spurred support for resettling refugees.
- **Mark Meadows Complies With Justice Dept. Subpoena: Report**  
Category: *POLITICS*  
The former White House chief of staff has turned over records as part of a federal investigation into the Jan. 6, 2021 assault on the Capitol.
- **Democrats Nominate Seth Magaziner In Key Rhode Island House Race**  
Category: *POLITICS*  
The state's general treasurer is slated to face former Cranston Mayor Allan Fung (R) in the general election.
- **Joe Biden Urges National Unity In Speech On Renewed 'Cancer Moonshot'**  
Category: *POLITICS*  
"Cancer does not discriminate red and blue," the president said Monday. "Beating cancer is something we can do together."
- **Sen. Tim Scott Downplays Electability Concerns Over Struggling Senate GOP Candidates**  
Category: *POLITICS*  
"Who we have on the field is who we're gonna play," Scott told "Fox News Sunday."
- **Electability Is On The Ballot In Key Rhode Island House Primary**  
Category: *POLITICS*  
Democrats in the state's more moderate House district are deciding who will take on a popular Republican former mayor.
- **In A Nod To JFK, Joe Biden Pushing 'Moonshot' To Fight Cancer**

Gambar 3.3 Tampilan berita yang terpilih

## 4. KESIMPULAN

Penelitian ini bertujuan untuk mengembangkan Sistem Rekomendasi Konten Website Berita dengan menggunakan Algoritma Naïve Bayes. Fokus penelitian melibatkan pengukuran tingkat personalisasi pada sistem rekomendasi berita yang ada, implementasi algoritma Naïve Bayes untuk memahami preferensi pengguna, dan analisis kemampuan algoritma dalam mengatasi asumsi kemandirian fitur. Metode penelitian mencakup studi literatur, analisis kebutuhan, pengumpulan dataset, implementasi algoritma, analisis tingkat personalisasi, dan pengujian aplikasi. Studi literatur melibatkan pemahaman konsep sistem rekomendasi berita, algoritma Naïve Bayes, dan personalisasi. Analisis kebutuhan memahami tujuan sistem rekomendasi, preferensi pengguna, dan harapan pemilik website berita. Pengumpulan dataset dilakukan secara online dengan data yang sudah dikumpulkan di website Kaggle. Hasil penelitian belum sempurna karena masih belum terdapat personalisasi tiap user. Kedepannya, diharapkan penelitian ini dapat digunakan sebagai referensi untuk meningkatkan tingkat personalisasi rekomendasi berita, memberikan pemahaman yang lebih baik tentang preferensi pengguna, dan mengoptimalkan pengalaman pengguna di website berita. Kesimpulan dari penjelasan di atas menunjukkan bahwa penelitian ini melibatkan langkah-langkah yang komprehensif untuk mengatasi tantangan dalam menyajikan berita yang sesuai dengan preferensi pengguna dalam era digital yang penuh informasi.

## REFERENCES

- Alvina Felicia Watratan, Arwini Puspita. B, & Dikwan Moeis. (2020). Implementasi Algoritma naive Bayes Untuk Memprediksi Tingkat Penyebaran covid-19 di Indonesia. *Journal of Applied Computer Science and Technology*, 1(1), 7–14. <https://doi.org/10.52158/jacost.v1i1.9>

- Anita, M., Susanto, B., & Larwuy, L. (2023). Perbandingan metode random forest dan naïve Bayes Dalam email spam filtering. *KUBIK: Jurnal Publikasi Ilmiah Matematika*, 7(2), 88–96. <https://doi.org/10.15575/kubik.v7i2.18933>
- Arif Yudianto, M. R., Agustin, T., Morgan James, R., Imannisa Rahma, F., Rahim, A., & Utami, E. (2021). Rainfall forecasting to recommend crops varieties using moving average and naïve bayes methods. *International Journal of Modern Education and Computer Science*, 13(3), 23–33. <https://doi.org/10.5815/ijmeecs.2021.03.03>
- Kurniadi, D., Abdurachman, E., Warnars, H. L., & Suparta, W. (2019). A proposed framework in an intelligent recommender system for the college student. *Journal of Physics: Conference Series*, 1402(6), 066100. <https://doi.org/10.1088/1742-6596/1402/6/066100>
- Parlina, I., Yusuf Arnol, M., Febriati, N. A., Dewi, R., Wanto, A., Lubis, M. R., & Susiani. (2019). Naive Bayes algorithm analysis to determine the percentage level of visitors the most dominant zoo visit by age category. *Journal of Physics: Conference Series*, 1255(1), 012031. <https://doi.org/10.1088/1742-6596/1255/1/012031>
- Salmi, N., & Rustam, Z. (2019). Naïve bayes classifier models for predicting the colon cancer. *IOP Conference Series: Materials Science and Engineering*, 546(5), 052068. <https://doi.org/10.1088/1757-899x/546/5/052068>
- Saritas, M. M., & Yasar, A. (2019). Performance analysis of ann and naïve Bayes classification algorithm for Data Classification. *International Journal of Intelligent Systems and Applications in Engineering*, 7(2), 88–91. <https://doi.org/10.18201/ijisae.2019252786>
- Utari, D. R. (2023). Sentiment analysis of the COVID-19 booster vaccine with the naïve Bayes algorithm. *Procedia of Engineering and Life Science*, 4. <https://doi.org/10.21070/pels.v4i0.1379>
- Wibisono, A. D., Dadi Rizkiono, S., & Wantoro, A. (2020). Filtering spam email Menggunakan metode naïve Bayes. *TELEFORTECH : Journal of Telematics and Information Technology*, 1(1). <https://doi.org/10.33365/tft.v1i1.685>