



LEMBAR PENGESAHAN
Nomor: 436/LP-LPBK/VIII/2026

Judul : KOMPARASI ALGORITMA *NAIVE BAYES* DAN *SUPPORT VECTOR MACHINE* (SVM) DALAM PENENTUAN KATEGORI ARTIKEL DI WEBSITE MAJELISTABLIGHPWMJATENG.COM

Nama : Agung Dwi Hanggoro

Menerangkan bahwa abstrak dengan judul di atas telah diterjemahkan ke dalam Bahasa Inggris oleh Lembaga Pengembangan Bahasa dan Kerja Sama (LPBK), Universitas Muhammadiyah Pekajangan Pekalongan.

Pekalongan, 19 Agustus 2026

Disahkan oleh,
Kepala Lembaga Pengembangan Bahasa dan Kerja Sama (LPBK)



Aida Rusmariana, S.Kep., Ns., MAN

ABSTRAK

Agung Dwi Hanggoro¹, Aslam Fatkhudin², Fenilinas Adi Artanto³

KOMPARASI ALGORITMA *NAIVE BAYES* DAN *SUPPORT VECTOR MACHINE (SVM)* DALAM PENENTUAN KATEGORI ARTIKEL DI WEBSITE MAJELISTABLIGHPWMJATENG.COM

Perkembangan publikasi digital mendorong meningkatnya jumlah artikel yang perlu dikelola dan dikelompokkan secara sistematis. Website majelistablighpwmjateng.com menyediakan berbagai artikel dakwah dengan kategori yang beragam. Kondisi tersebut menjadikan proses pengategorian secara manual berpotensi memerlukan waktu yang lebih lama serta menghasilkan pelabelan yang kurang konsisten. Penelitian ini bertujuan untuk membangun model klasifikasi kategori artikel berdasarkan judul dan membandingkan kinerja algoritma *Naive Bayes* dengan *Support Vector Machine* dalam mengklasifikasikan artikel dakwah berbahasa Indonesia. Data penelitian berupa judul artikel beserta kategorinya yang diperoleh dari website majelistablighpwmjateng.com pada periode 10 Juni 2023 hingga 9 Juni 2026. Dataset awal berjumlah 1.369 data yang terbagi dalam 19 kategori. Kategori Khutbah Nikah yang hanya memiliki satu data dihapus agar proses pembagian *dataset* dapat dilakukan secara optimal. Setelah penghapusan tersebut, *dataset* akhir terdiri atas 1.368 data dalam 18 kategori. Tahapan penelitian mencakup pengumpulan data, *preprocessing* yang meliputi *cleaning*, *case folding*, *tokenisasi*, dan *stopword removal*, pembobotan fitur dengan metode *Term Frequency-Inverse Document Frequency*, pelatihan model, serta evaluasi menggunakan *metrik accuracy*, *precision*, *recall*, dan *F1-score*. Hasil pengujian menunjukkan bahwa algoritma *SVM* menghasilkan kinerja yang lebih baik daripada *Naive Bayes*. Model *SVM* mencapai *accuracy* sebesar 79,56%, *precision* sebesar 72,11%, *recall* sebesar 79,56%, dan *F1-score* sebesar 74,36%. Adapun model *Naive Bayes* memperoleh *accuracy* sebesar 73,36%, *precision* sebesar 64,73%, *recall* sebesar 73,36%, dan *F1-score* sebesar 66,86%. Perbedaan tersebut mengindikasikan bahwa *SVM* lebih efektif dalam menangani data teks berdimensi tinggi dan pola kategori yang kompleks. Meskipun demikian, ketidakseimbangan jumlah data pada setiap kategori masih memengaruhi kemampuan model dalam mengidentifikasi kelas minoritas. Berdasarkan hasil penelitian, kombinasi TF-IDF dan *SVM* lebih sesuai digunakan untuk mengklasifikasikan kategori judul artikel pada website majelistablighpwmjateng.com dibandingkan kombinasi TF-IDF dan *Naive Bayes*. Penelitian ini diharapkan dapat mendukung proses pengelompokan artikel secara lebih cepat, konsisten, dan efisien, serta menjadi referensi bagi pengembangan sistem klasifikasi teks berbahasa Indonesia dalam domain dakwah.

Kata Kunci: klasifikasi teks, judul artikel, TF-IDF, *Naive Bayes*, *Support Vector Machine*, website dakwah, majelistablighpwmjateng.com.

**Undergraduate Program in Informatics
Faculty of Engineering and Computer Science
University of Muhammadiyah Pekajangan Pekalongan**

ABSTRACT

Agung Dwi Hanggoro¹, Aslam Fatkhudin², Fenilinas Adi Artanto³

THE COMPARISON OF *NAIVE BAYES ALGORITHMS* AND *SUPPORT VECTOR MACHINE (SVM)* IN DETERMINING ARTICLE CATEGORIES ON THE MAJELISTABLIGHPWMJATENG.COM WEBSITE

The development of digital publications has led to an increase in the number of articles that need to be managed and grouped systematically. majelistablighpwmjateng.com website provides a variety of da'wah articles in various categories. This condition makes the manual categorization process potentially take longer and result in less consistent labeling. This study aims to build a model for the classification of article categories based on titles and compare the performance of *the Naive Bayes* algorithm with *the Support Vector Machine* in classifying da'wah articles in Indonesian. The research data is in the form of article titles and their categories obtained from [the majelistablighpwmjateng.com website](https://majelistablighpwmjateng.com) in the period from June 10, 2023 to June 9, 2026. The initial dataset totaled 1,369 data points divided into 19 categories. The category of Marriage Sermon, that only has one data is deleted so that the dataset sharing process could be carried out optimally. After the deletion, the final *dataset* consisted of 1,368 data points in 18 categories. The research stages include data collection, *preprocessing*, which includes *cleaning*, *case folding*, *tokenization*, and *stopword removal*, feature weighting with *the Term Frequency-Inverse Document Frequency* method, model training, and evaluation using *accuracy*, *precision*, and *recall metrics*, then *F1 score*. The test results showed that *the SVM* algorithm produced better performance than *Naive Bayes*. The SVM model achieved an *accuracy* of 79.56%, *precision* of 72.11%, *recall* of 79.56%, and an *F1-score* of 74.36%. The *Naive Bayes* model obtained an *accuracy* of 73.36%, *precision* of 64.73%, *recall* of 73.36%, and an *F1-score* of 66.86%. These differences indicate that SVM is more effective at handling high-dimensional text data and complex category patterns. However, the imbalance in the amount of data in each category still affects the model's ability to identify minority classes. Based on the results of the study, the combination of TF-IDF and SVM is more suitable to be used to classify the category of article titles on the [majelistablighpwmjateng.com website](https://majelistablighpwmjateng.com) than the combination of TF-IDF and *Naive Bayes*. This research is expected to support the process of grouping articles more quickly, consistently, and efficiently, as well as become a reference for the development of a classification system for Indonesian texts in the domain of da'wah.

Keywords: text classification, article title, TF-IDF, *Naive Bayes*, *Support Vector Machine*, da'wah website, majelistablighpwmjateng.com.