Hoax News Detection on Urban Issues Using Naive Bayes with Chi-Square Feature Selection

Bimli[1], Wilda Susanti[2]*

[1,2*] Informatics Engineering, Pelita Indonesia Institute, Pekanbaru, Indonesia
*Corresponding author: wilda@lecturer.pelitaindonesia.ac.id

Abstract. The rapid spread of hoax news in digital media has become a serious concern because it can mislead the public, influence opinions, and potentially trigger social unrest, particularly in urban-related issues. The increasing volume of online information also makes manual verification difficult, highlighting the need for automatic hoax detection systems using machine learning approaches. This study aims to develop a hoax news detection model using the Multinomial Naive Bayes algorithm combined with Chi-Square feature selection to improve classification efficiency. The dataset consists of 1,000 Indonesian news articles related to urban issues, consisting of hoax and factual news collected through web scraping from fact-checking websites and online news portals. The research process includes text preprocessing stages such as case folding, tokenizing, stopword removal, and stemming, followed by TF-IDF feature extraction. Chi-Square feature selection was applied to reduce feature dimensionality by selecting the most relevant terms. The dataset was divided into training and testing sets using an 80:20 ratio. Model performance was evaluated using a confusion matrix with accuracy, precision, recall, and F1-score metrics. The experimental results show that the Multinomial Naive Bayes model without feature selection achieved the highest accuracy of 98.5%. Meanwhile, the application of Chi-Square feature selection with 200 selected features achieved an accuracy of 90.5%, while using 100 features resulted in an accuracy of 86.5%. These results indicate that although feature selection slightly reduces classification performance, it significantly reduces feature dimensionality while maintaining acceptable performance. Overall, this study demonstrates that Multinomial Naive Bayes is effective for hoax news classification, while Chi-Square feature selection provides an efficient alternative when dimensionality reduction is required.
Keywords: Hoax News, Naive Bayes, Chi-Square, Text Classification, Feature Selection.

How to cite (in APA style):
Bimli., & Susanti, W. (2026). Hoax News Detection on Urban Issues Using Naive Bayes with Chi-Square Feature Selection. *Proceeding of IJC-MaSTEd*, 1(1), 79-93.

DOI: https://doi.org/10.31571/ijcmasted.v1i1.1038

1. Introduction

The rapid development of digital technology has significantly increased the spread of information through online media. However, this development is also accompanied by the rapid dissemination of hoax news which can mislead the public, influence public opinion, and reduce trust in credible information sources (Agma, 2025). The increasing amount of digital information also makes manual verification difficult, creating the need for automatic hoax detection systems using machine learning approaches (Reddy et al., 2020).

Hoax detection has become an important research topic in the field of text mining and Natural Language Processing (NLP). Various classification algorithms such as Naive Bayes, Support Vector Machine, and Logistic Regression have been widely applied in fake news classification tasks. Among these methods, Naive Bayes is widely used due to its simplicity, efficiency, and relatively stable performance in text classification problems (Rahutomo et al., 2019); (Snegha, 2024).

In addition to classification algorithms, the preprocessing stage also plays an important role in support text classification quality. Several studies show that preprocessing techniques such as tokenization, stopword removal, and stemming significantly influence classification results by improving text quality and reducing noise (Sheviraa et al., 2022); (Saraswati et al., 2025); (Siino et al., 2024).

Furthermore, feature selection techniques are commonly applied to reduce dimensionality and select the most relevant features. One widely used method is Chi-Square feature selection, which evaluates the relationship between terms and class labels. Previous research shows that Chi-Square can help support classification effectiveness by removing irrelevant features and retaining informative terms (Ratiasasadara et al., 2023); (Sami'un et al., 2024).

Although many studies have discussed hoax detection, research focusing specifically on hoax news related to urban issues is still limited. In addition, the implementation of classical machine learning approaches combined with feature selection for domain-specific hoax datasets remains relatively underexplored in the literature. Therefore, further investigation is needed to analyze how feature selection contributes to hoax detection performance in specific thematic contexts.

Based on these considerations, this research proposes a hoax news detection model using Multinomial Naive Bayes combined with Chi-Square feature selection. This study aims to analyze the application of feature selection in hoax news classification related to urban issues and evaluate the performance of the proposed approach.

Table 1. Summary of Related Research in Hoax News Detection

Reference	**Methodology/Model**	**Key Contribution**	**Limitations/Context**
Rahutomo et al. (2019)	Naive Bayes	Demonstrated effectiveness of Naive Bayes for Indonesian hoax classification	No feature selection applied to improve model performance
Reddy et al. (2020)	Machine learning fake news detection	Provided framework for automated fake news classification	Dataset not focused on specific thematic issues
Snegha (2024)	Naive Bayes classifier	Showed Naive Bayes efficiency in text classification	No analysis of feature optimization techniques
Sheviraa et al. (2022)	Text preprocessing pipeline	Demonstrated importance of preprocessing order in classification	Did not evaluate classification algorithm performance
Saraswati et al. (2025)	Text normalization techniques	Showed stemming improves feature quality	Not applied to hoax detection context
Siino et al. (2024)	Preprocessing comparison	Highlighted importance of text cleaning process	Context limited to general text classification
Ratiasasadara et al. (2023)	Naive Bayes + Chi-Square	Improved classification accuracy using feature selection	Applied only to sentiment analysis data

Based on Table 1. previous studies demonstrate that Naive Bayes is widely used for text classification due to its simplicity and effectiveness. Additionally, preprocessing and feature selection methods

such as Chi-Square have been proven to improve classification performance. However, most studies focus on sentiment analysis or general text classification rather than hoax news detection. Furthermore, limited studies evaluate the impact of feature selection on Naive Bayes performance in domain-specific hoax datasets, particularly urban issue news. Therefore, further research is required to analyze the effectiveness of Chi-Square feature selection in improving Naive Bayes performance for hoax news detection.

2. Method

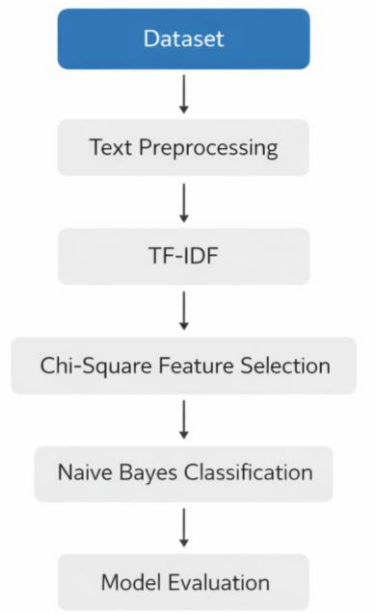

Figure 1. Research workflow

The research methodology used in this study consists of several stages, starting from dataset collection, text preprocessing, feature extraction using TF-IDF, feature selection using Chi-Square, classification using Multinomial Naive Bayes, and model evaluation. The overall research workflow is illustrated in Figure 1.

2.1. Dataset

This study uses a dataset consisting of 1,000 Indonesian news articles related to urban issues, consisting of 500 hoax news and 500 factual news. The balanced class distribution helps avoid classification bias and improves model performance.

Hoax data were collected from TurnBackHoax.id, while factual news were obtained from Urbannews.id. Each data instance consists of two attributes: narrative and label. The narrative contains the news text, while the label indicates the class (hoax or fact).

Table 2. Dataset Structure

Attribute	Description
Narrative	News article text
Label	Hoax or Fact classification

Table 2. shows the structure of the dataset used in this study. The narrative attribute serves as the input feature for text processing, while the label attribute is used as the target class in the classification process.

The dataset was divided into training and testing data using an 80:20 ratio.

2.2. Text Preprocessing

Text preprocessing aims to clean and standardize text data before classification. The preprocessing stages include:

a. Case Folding
Case folding converts all characters into lowercase to ensure consistency.
Example:
"Berita HOAX Tentang Banjir" becomes: "berita hoax tentang banjir".

b. Tokenizing
Tokenizing is the process of splitting sentences into individual words.
Example: "berita hoax tentang banjir" becomes: [berita, hoax, tentang, banjir].

c. Stopword Removal
Stopword removal removes common words that do not contribute significant meaning such as: “dan, yang, di, ke, dari”.

d. Stemming
Stemming converts words into their root forms.
Example: "pembangunan" → "bangun", "kemacetan" → "macet".

This process reduces dimensionality and helps prepare text data for classification.

2.3. TF-IDF Feature Extraction

After preprocessing, text data was transformed into numerical representation using TF-IDF (Term Frequency – Inverse Document Frequency).

a. TF measures how often a word appears in a document:
TF formula:

$$TF(t,d) = \frac{f(t,d)}{\sum f(t,d)} \quad (1)$$

b. IDF measures the importance of a word across documents:
IDF formula:

$$IDF(t) = \log\left(\frac{N}{df(t)}\right) \quad (2)$$

c. TF-IDF weight is calculated as:

$$TF - IDF\ (t,d) = TF(t,d) \times IDF(t) \quad (3)$$

Where:
t = term
d = document
N = total documents
df = document frequency

TF-IDF helps assign higher weight to important words and lower weight to common words.

2.4. Feature Selection using Chi-Square

Chi-Square feature selection was used to select the most relevant features. This method measures the dependency between terms and class labels.
The Chi-Square formula is:

$$X^2(t,c) = \frac{N \times (AD - BC)^2}{(A+C) \times (B+D) \times (A+B) \times (C+D)} \quad (4)$$

Where:
A = term appears in hoax class
B = term appears in factual class
C = term absent in hoax class
D = term absent in factual class
Features with high Chi-Square scores indicate strong relationships with class labels.

The Chi-Square method was used to select the most relevant features based on their statistical relationship with the class labels. In this study, the top 200 features with the highest Chi-Square scores were selected and used as input for the classification model.

Feature selection was performed to reduce dimensionality and improve model performance by removing irrelevant terms.

2.5. Multinomial Naive Bayes

Multinomial Naive Bayes was used for classification because it performs well on text classification problems.

Naive Bayes is based on Bayes theorem:

$$P(C_i|X) = \frac{P(X|C_i) \cdot P(C_i)}{P(X)} \quad (5)$$

P(C|X) = posterior probability
P(X|C) = likelihood
P(C) = prior probability
P(X) = evidence
Multinomial Naive Bayes calculates probability based on word frequency distributions.

The classification decision is determined by selecting the class with the highest posterior probability.

2.6. Model Evaluation

Model performance was evaluated using a Confusion Matrix.

The evaluation metrics include:
Accuracy
Accuracy measures overall prediction correctness:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

Precision
Precision measures prediction correctness in positive class:

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

Recall
Recall measures model ability to detect positive class:

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

F1-Score
F1 Score balances precision and recall:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{9}$$

Where:

TP = True Positive
TN = True Negative
FP = False Positive
FN = False Negative

These metrics provide comprehensive evaluation of classification performance.

2.7. Experimental Setup
The experiment was conducted using the Python programming language in the Google Colaboratory cloud environment. The libraries used in this study include Scikit-learn for classification, Pandas for data processing, NumPy for numerical computation, Sastrawi for Indonesian text stemming, and Matplotlib for data visualization.

The experiments were executed on the Google Colab platform which provides cloud-based computational resources for machine learning experiments.

3. Results and Discussion
This section presents the experimental results obtained from each stage of the proposed hoax news detection pipeline, including dataset distribution, preprocessing results, feature extraction, feature selection, and classification performance evaluation.

3.1. Dataset
The dataset used in this study consists of 1,000 news articles related to urban issues, with two balanced classes: 500 hoax news and 500 factual news. The balanced dataset helps the classification model learn patterns without bias toward a particular class.

After applying the 80:20 data splitting strategy, 800 data were used for training and 200 data were used for testing.

Table 3. Dataset Distribution After Splitting

Class	Training	Testing	Total
Hoax	400	100	500
Fact	400	100	500
Total	800	200	1000

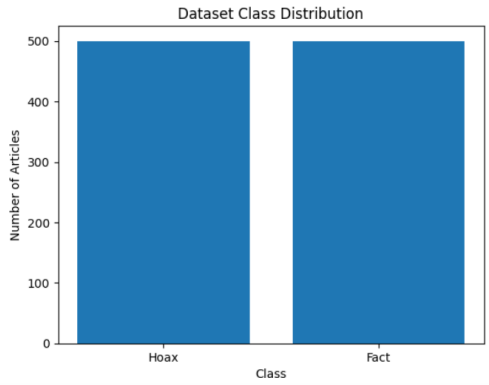

Figure 2. Dataset Class Distribution

Table 3 presents the numerical distribution of the dataset after the data splitting process, while Figure 2 illustrates the class balance visually. The consistent distribution between hoax and factual news across training and testing data indicates that the dataset is well balanced. This condition is essential in classification tasks because class imbalance may negatively affect model generalization and produce misleading performance metrics. Therefore, the balanced dataset supports fair model training and more reliable performance evaluation.

3.2. Text Preprocessing

Text preprocessing was performed to clean and normalize the news text before the feature extraction process. This stage includes case folding, tokenizing, stopword removal, and stemming.

Table 4. Text Preprocessing

Original Text	Preprocessing Result
bencana hari ini~lenyap??baru saja papua berduka gelombang hebat sapu bersih 1.085 rumah pagi ini?	bencana, lenyap, papua, berduka, gelombang, hebat, sapu, bersih, rumah, pagi
baru saja blitar bergetar dahsyat 19 februari 2024 gempa magnitude 10 6 guncang jatim	blitar, bergetar, dahsyat, februari, gempa, magnitude, guncang, jatim
bencana hari ini~inalilla??banten berduka gelombang hebat tiba² melibas habis 2 kabupaten & hancur?	bencana, inalilla, banten, berduka, gelombang, hebat, tiba², melibas, habis, kabupaten, hancur
waspada ada penculik anak-anak yang berumur 1-12 tahun. bapak-bapak ibu-ibu harus menjaga anak kita dengan hati-hati. penculik sedang ada dalam kampung-kampung dan dia menyamar sebagai: tolong disebarkan terima kasih	waspada, penculik, anak, anak, berumur, menjaga, anak, hati, hati, penculik, kampung, kampung, menyamar, tolong, disebarkan, terima, kasih

besaran tarif parkir khusus di bandara soekarno hatta jakarta dikabarkan melonjak drastis. beredar kabar per 1 september 2025 tarif parkir khusus naik menjadi rp 150.000.	besaran, tarif, parkir, khusus, bandara, soekarno, hatta, jakarta, dikabarkan, melonjak, drastis, beredar, kabar, september, tarif, parkir, khusus, rp

Table 4. shows the transformation of raw text into cleaned text. The preprocessing process successfully converts text into lowercase form, removes punctuation, eliminates stopwords, and reduces words to their root form. This process helps reduce noise and improves the quality of features used in classification.

3.3. TF-IDF Feature Extraction

After preprocessing, feature extraction was performed using TF-IDF to convert text data into numerical feature vectors. This process produced thousands of unique terms representing the vocabulary of the dataset.

The TF-IDF weighting process highlights important terms based on their frequency distribution across documents, enabling the model to better capture textual characteristics. Several terms related to urban issues such as pertamina, indonesia, gempa, kota, and banjir emerge as dominant features, indicating that these topics frequently appear and contribute to distinguishing document patterns. Examples of mean TF-IDF feature scores and the distribution of important terms are presented in Table 5 and Figure 3. These results demonstrate how TF-IDF emphasizes discriminative words that support the classification process before further feature selection using the Chi-Square method.

Table 5. Top TF-IDF Features Based on Mean TF-IDF Scores

Term	Mean TF-IDF Score
pertamina	0.022321
indonesia	0.020991
gempa	0.020680
kota	0.020578
banjir	0.019774
pt	0.015597
wilayah	0.015405
jakarta	0.014550
rumah	0.014505
bencana	0.013768
energi	0.013385
guncang	0.011783
tsunami	0.011648
warga	0.011492
akibat	0.011079

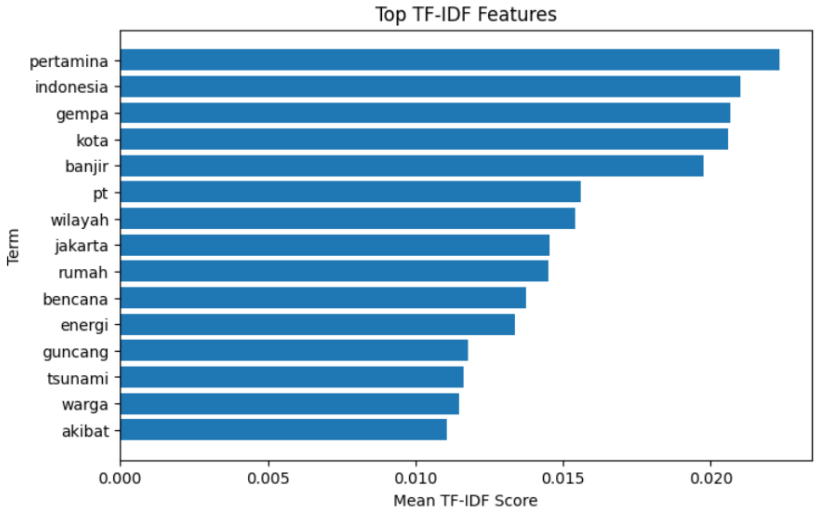

Figure 3. Top TF-IDF Features Extracted from the Dataset

3.4. Chi-Square Feature Selection

Feature selection was performed using the Chi-Square method to identify the most relevant features obtained from the TF-IDF extraction. The initial TF-IDF process produced 4664 features, which could increase computational cost and include irrelevant terms that do not significantly contribute to classification. Therefore, feature selection is required to reduce dimensionality while maintaining important information for the classification process.

Based on the Chi-Square scoring results, feature selection was performed by selecting the top-ranked features based on their relevance to the class labels. In this study, two feature subsets consisting of the top 100 and top 200 features were evaluated to analyze the impact of feature size on classification performance. This selection helps the model focus on the most discriminative terms between hoax and factual news while also enabling evaluation of how different feature sizes affect classification performance. Several important terms such as pertamina, gempa, and banjir appear as dominant features, indicating strong statistical relationships with the class labels. Examples of the selected features along with their Chi-Square scores can be observed in Table 6, while the distribution of the top-ranked features is illustrated in Figure 4. Higher Chi-Square scores indicate stronger feature relevance in distinguishing hoax and factual news. This result demonstrates that the Chi-Square method effectively removes less relevant features while retaining important terms that support the classification process.

Table 6. Top 15 Features Selected by Chi-Square

Feature	Chi-Square Score
pertamina	22.320712
gempa	20.679754
banjir	19.042531
pt	14.279077
energi	13.384753
guncang	11.782836
sapu	11.055531
tsunami	10.484212
detik	10.135781
persero	9.806211
gelombang	9.624432

Feature	Chi-Square Score
magnitude	9.588706
hancur	9.515458
agustus	9.408448
menteri	9.337199

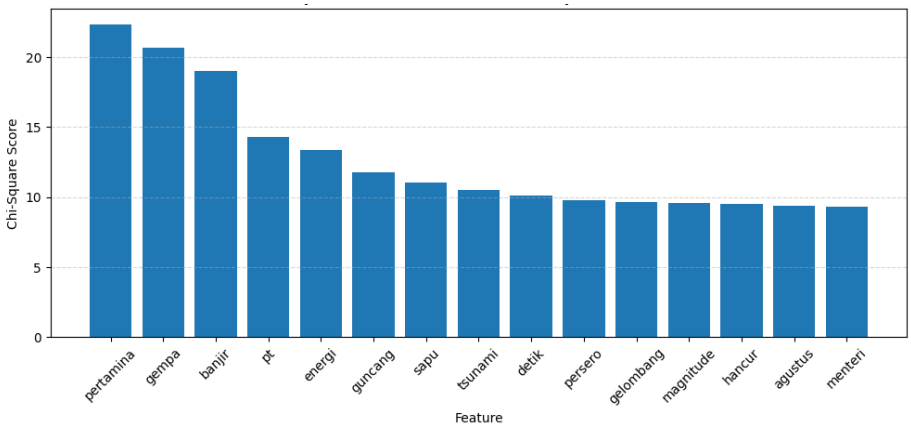

Figure 4. Top 15 Features Based on Chi-Square Scores

3.5. Multinomial Naive Bayes Classification

The classification process was performed using the Multinomial Naive Bayes algorithm to classify news articles into hoax and factual categories. The model was selected due to its effectiveness and efficiency in text classification problems, particularly when working with TF-IDF feature representations.

In this study, three experimental scenarios were evaluated to analyze the impact of feature selection on classification performance. The first scenario used the complete TF-IDF feature set consisting of 4664 features without feature selection. The second and third scenarios applied Chi-Square feature selection using the top 100 and top 200 selected features, respectively. This experimental design allows analysis of how feature reduction affects model performance.

The dataset was divided into training and testing data using an 80:20 ratio, resulting in 800 training data and 200 testing documents used for evaluation.

Table 7 shows the confusion matrix obtained from the Multinomial Naive Bayes classification without feature selection. This baseline model produced the highest classification performance because all TF-IDF features were retained.

Table 7. Confusion Matrix of Naive Bayes without Feature Selection

Actual Class	Predicted Hoax	Predicted Fact
Hoaks	94	2
Fakta	1	103

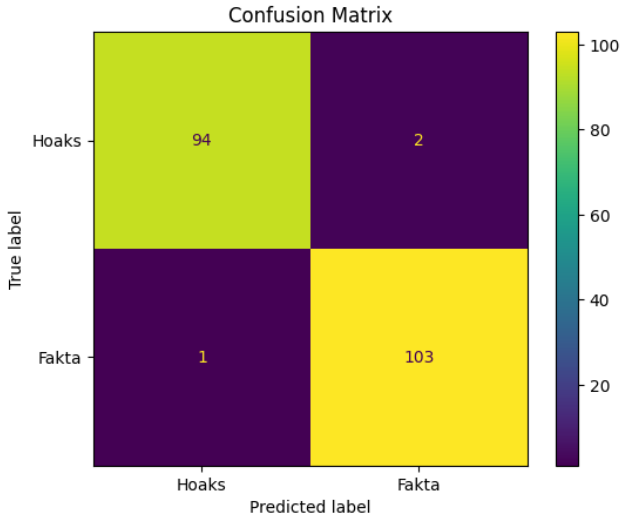

Figure 5. Confusion Matrix of Naive Bayes without Feature Selection

Based on Table 7, the model correctly classified 197 out of 200 testing documents, resulting in an accuracy of 98.5%.

Table 8 shows the confusion matrix obtained from the Naive Bayes classification using Chi-Square feature selection with the top 100 features.

Table 8. Confusion Matrix of Naive Bayes with Chi-Square (Top 100 Features)

Actual Class	Predicted Hoax	Predicted Fact
Hoax	97	3
Fact	24	76

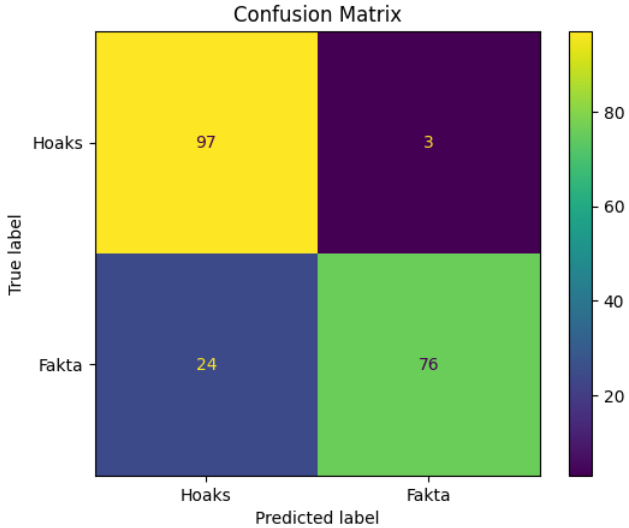

Figure 6. Confusion Matrix of Naive Bayes with Chi-Square 100 Feature Selection

Based on Table 8, the model correctly classified 173 out of 200 testing documents, resulting in an accuracy of 86.5%.

Table 9 shows the confusion matrix obtained from the Naive Bayes classification using Chi-Square feature selection with the top 200 features.

Table 9. Confusion Matrix of Naive Bayes with Chi-Square (Top 200 Features)

Actual Class	Predicted Hoax	Predicted Fact
Hoax	97	3
Fact	16	84

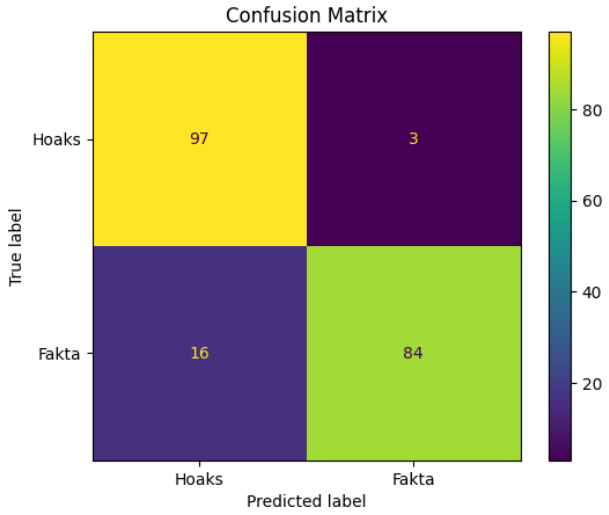

Figure 7. Confusion Matrix of Naive Bayes with Chi-Square (Top 200 Features)

Based on Table 9, the model correctly classified 181 out of 200 testing documents, resulting in an accuracy of 90.5%.

3.6. Performance Evaluation

The performance of the classification models was evaluated using accuracy, precision, recall, and F1-score. These metrics provide a comprehensive evaluation of classification performance across different model configurations.

Table 10. Performance Comparison of All Models

Model	Features	Accuracy	Macro Precision	Macro Recall	Macro F1-score
NB	4664	0.985	0.99	0.98	0.98
NB + Chi-Square	200	0.905	0.91	0.91	0.90
NB + Chi-Square	100	0.865	0.88	0.86	0.86

Based on Table 10, the Multinomial Naive Bayes model without feature selection achieved the highest accuracy of 98.5%. This indicates that retaining the complete TF-IDF feature set provides the richest textual representation for classification.

Meanwhile, the Chi-Square feature selection significantly reduced the number of features while still maintaining relatively good classification performance. The model using 200 selected features achieved an accuracy of 90.5%, while the model using 100 features achieved an accuracy of 86.5%.

These results indicate that reducing the number of features may decrease classification performance due to the removal of important discriminative terms. However, feature selection remains useful for reducing feature dimensionality.

A more detailed comparison and discussion of these results is presented in the comparison analysis section.

3.7. Comparison Analysis

This section compares the performance of the Multinomial Naive Bayes model without feature selection and the models using Chi-Square feature selection with 200 and 100 selected features.

Based on the experimental results, the Multinomial Naive Bayes model without feature selection achieved the best performance with an accuracy of 98.5% and a macro F1-score of 0.98. This result indicates that using the complete TF-IDF feature set enables the model to capture more comprehensive textual information, which leads to better classification performance.

As illustrated in Figure 10, the Multinomial Naive Bayes model without feature selection clearly outperforms the models that apply Chi-Square feature selection. The figure also shows a consistent decrease in accuracy as the number of selected features is reduced.

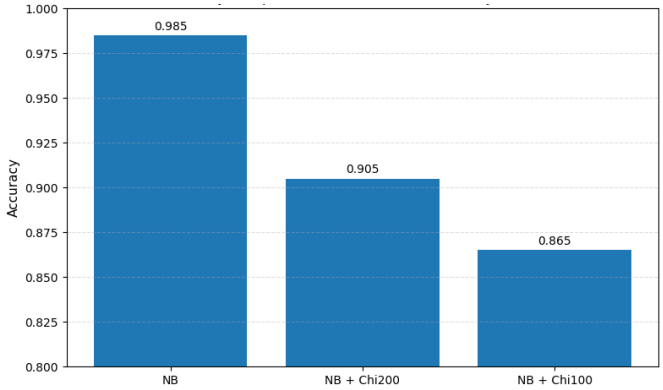

Figure 10. Accuracy comparison of Multinomial Naive Bayes models

Meanwhile, the Chi-Square feature selection method successfully reduced the number of features from 4664 to 200 and 100 features. Despite this substantial reduction, the models were still able to maintain acceptable performance levels, achieving accuracies of 90.5% and 86.5%, respectively. These results indicate that Chi-Square is capable of selecting the most relevant features while eliminating less important terms.

As shown in Figure 10, the model using 200 selected features performs better than the model using only 100 features. This indicates that 200 features still preserve sufficient textual information for effective classification, whereas reducing the features to 100 begins to remove important contextual information required by the classifier.

However, the results also reveal a clear trade-off between dimensionality reduction and classification performance. The downward trend shown in Figure 10 confirms that as the number of selected features decreases, both accuracy and F1-score also decrease. This suggests that excessive feature reduction may remove important discriminative terms required for accurate classification.

Overall, these findings indicate that the Multinomial Naive Bayes model without feature selection provides the best classification performance. Meanwhile, Chi-Square feature selection offers an effective alternative for reducing feature dimensionality while maintaining relatively good

performance, particularly when computational efficiency and dimensionality reduction are important considerations.

3.8. Discussion

Based on the experimental results, the Multinomial Naive Bayes model demonstrates strong performance in classifying hoax news related to urban issues. The baseline model without feature selection achieved the highest accuracy, indicating that retaining the complete TF-IDF feature set allows the classifier to utilize richer textual information for better decision making.

Meanwhile, the use of Chi-Square feature selection shows that dimensionality reduction can still maintain relatively good classification performance while significantly reducing the number of features. This indicates that many irrelevant or less informative terms can be removed without severely degrading model performance.

The experimental results also show that the model achieved high precision in the hoax class, indicating that most predictions labeled as hoax are correct. However, the recall of the hoax class in the reduced feature models indicates that some hoax news articles are still misclassified as factual news. This may occur because certain hoax articles share similar linguistic patterns with factual news, which increases classification difficulty.

The comparison results further indicate that reducing the number of selected features affects classification performance. Models with 200 selected features perform better than those with 100 features, suggesting that excessive feature reduction may remove important discriminative information required for accurate classification.

Despite showing strong performance, this study still has several limitations. First, the dataset size is relatively limited, which may affect the generalization ability of the model. Second, this study only evaluates a single classification algorithm, which limits the comparative analysis of model performance. Future research may consider comparing multiple machine learning algorithms, increasing dataset size, or applying additional feature engineering techniques to obtain more comprehensive results.

Overall, the results demonstrate that the Multinomial Naive Bayes model is effective for hoax news classification, while Chi-Square feature selection provides an alternative approach when feature dimensionality reduction is required.

4. Conclusion

This study proposed a hoax news classification model on urban issues using the Multinomial Naive Bayes algorithm with TF-IDF feature extraction and Chi-Square feature selection. The research process included data collection, text preprocessing, TF-IDF feature extraction, feature selection, classification, and performance evaluation.

Based on the experimental results, the Multinomial Naive Bayes model without feature selection achieved the best performance with an accuracy of 98.5%. Meanwhile, the models using Chi-Square feature selection achieved accuracies of 90.5% and 86.5% for 200 and 100 selected features, respectively. These results indicate that while feature selection reduces dimensionality, it may also slightly decrease classification performance.

This study demonstrates that Chi-Square feature selection can effectively reduce feature dimensionality while maintaining competitive classification performance and significantly reducing

the number of features. This makes Chi-Square useful when computational efficiency or dimensionality reduction becomes an important consideration.

Future research may consider using larger datasets, comparing multiple classification algorithms, or applying additional optimization techniques to further improve classification performance.

Reference

Agma, A. R. (2025). Hoaks dan Disinformasi di Media Sosial: Strategi Literasi Digital untuk Meningkatkan Kesadaran Publik. *Jurnal Komunikasi Dan Media*, *1*(1), 30–36. https://jurnal.samudrailmu.com/index.php/jkomed/article/view/28

Rahutomo, F., Yanuar, I., Pratiwi, R., & Ramadhani, D. M. (2019). Naïve Bayes's Experiment On Hoax News Detection In Indonesian Language. *Jurnal Penelitian Komunikasi Dan Opini Publik*, *23*(1), 1–15. https://doi.org/10.33299/jpkop.23.1.1805

Ratiasasadara, P. W., Sudarno, & Tarno. (2023). *Analisis sentimen penerapan ppkm pada twitter menggunakan naïve bayes classifier dengan seleksi fitur chi-square 1,2,3*. *11*, 580–590. https://doi.org/10.14710/j.gauss.11.4.580-590

Reddy, H., Raj, N., Gala, M., & Basava, A. (2020). *Text-mining-based Fake News Detection*. https://doi.org/10.1007/s11633-019-1216-5

Sami'un, D. C., Sugiharto, A., & Jie, F. (2024). Chi Square Feature Selection For Improving Sentiment Analysis Of News Data Privacy Treats. *Journal of Theoretical and Applied Information Technology*, *102*(18), 6601–6610. http://jatit.org/volumes/Vol102No18/3Vol102No18.pdf

Saraswati, N. W. S., Yanti, C. P., Muku, I. D. M. K., & Dewi, D. A. P. R. (2025). Evaluation Analysis of the Necessity of Stemming and Lemmatization in Text Classification. *Jurnal Manajemen, Teknik Informatika, Dan Rekayasa Komputer*, *24*(2), 321–332. https://doi.org/10.30812/matrik.v24i2.4833

Sheviraa, S., Suarjaya, I. M. A. D., & Buana, P. W. (2022). Pengaruh Kombinasi dan Urutan Pre-Processing pada Tweets Bahasa Indonesia. *Jurnal Ilmiah Teknologi Dan Komputer*, *3*(2). https://doi.org/10.24843/jtrti.2022.v03.i02.p06

Siino, M., Tinnirello, I., & Cascia, M. La. (2024). Is text preprocessing still worth the time ? A comparative survey on the influence of popular preprocessing methods on Transformers and traditional classifiers. *Information Systems*, *121*, 102342. https://doi.org/10.1016/j.is.2023.102342

Snegha, M. (2024). News Article Classification Using Navie Bayes Algorthim. *International Journal of Progressive Research in Engineering Managemengt and Science (IJPREMS)*, *4*(3), 425–427.