

PROJECT UJIAN AKHIR SEMESTER
KECERDASAN BUATAN
“SPAM EMAIL DETECTION”



Dosen Pengampu:

Yuyun Umaidah, S.Kom., M.Kom.

Disusun Oleh:

Rizky Fahrureza 2210631170149

Joyce Rosita Firdaus 2210631170075

PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SINGAPERBANGSA KARAWANG

2024

ABSTRAK

Email menjadi salah satu entitas yang penting dalam hal komunikasi, baik personal, korporat, pemerintah dan komunitas lainnya. Meluasnya penggunaan email sebagai media komunikasi mempunyai dampak negatif. Salah satu dampak negatif yang dihasilkan yaitu banyak email yang tidak diinginkan masuk pada *inbox*. Kondisi seperti ini biasa dikenal dengan *spam*. Kehadiran *spam* pada *email* menyebabkan penelitian terhadap pembangunan piranti lunak *spam* filter untuk mengklasifikasikan *email* *spam* dan *non spam* secara otomatis meningkat. *Naive Bayes* merupakan fungsi yang banyak digunakan oleh pengembang *spam* filter sebagai fungsi untuk mengklasifikasikan *email* karena sederhana dan mudah untuk diimplementasikan. Hasil penelitian menunjukkan bahwa *Naive Bayes* dapat memfilter *email spam* dengan klasifikasi *text*.

Kata Kunci: *spam, email, naive bayes*.

ABSTRACT

Email is an important entity in terms of communication, whether personal, corporate, government and other communities. The widespread use of email as a communication medium has a negative impact. One of the negative impacts is that many unwanted emails enter the inbox. This condition is commonly known as spam. The presence of spam in emails causes research into the development of spam filter software to classify spam and non-spam emails automatically. Naive Bayes is a function that is widely used by spam filter developers as a function to classify emails because it is simple and easy to implement. The results show that Naive Bayes can filter spam emails with text classification.

Keyword: spam, email, naive bayes.

DAFTAR ISI

ABSTRAK.....	i
DAFTAR ISI.....	ii
BAB I PENDAHULUAN.....	1
BAB II LANDASAN TEORI.....	3
2.1 Spam Email.....	3
2.2 Klasifikasi.....	3
2.3 Naive Bayes.....	3
BAB III HASIL DAN PEMBAHASAN.....	5
3.1 Dataset.....	5
3.2 Implementasi Spam Email Detection.....	6
BAB IV PENUTUP.....	10
4.1 Kesimpulan.....	10
4.2 Saran.....	10

BAB I

PENDAHULUAN

Email merupakan suatu entitas penting yang digunakan untuk berkomunikasi digital melalui internet (Andriansyah dan Suhendra, 2005). Pesan Elektronik menjadi primadona untuk berkomunikasi saat ini. Hanya terhubung dengan koneksi internet, berkirim pesan elektronik dapat dengan mudah dilakukan (Andros dkk, 2015). Semakin banyak orang yang terhubung ke internet menjadikan email sebagai salah satu alat komunikasi paling cepat dan ekonomis (Ananda, 2011). Email dengan segala kelebihannya membawa konsep komunikasi dan pengelolaannya menjadi lebih mudah sehingga pengguna email semakin meningkat dari hari ke hari (Mujtaba, Shuib, Raj, Majeed, & Al-Garadi, 2017). Penelitian menunjukkan bahwa terdapat lebih dari 3 miliar akun email di dunia dengan frekuensi pengiriman email sekitar 205 – 294 miliar setiap hari (Alurkar, et al., 2017) (Wijayanto & Takdir, 2014).

Dengan meningkatnya pengguna email memikat beberapa pihak untuk membombardir email dengan pesan yang tidak diminta (Andros dkk, 2015) yang berisi promosi produk atau jasa, pornografi, virus dan hal-hal yang tidak penting (Widiasari dan Bayu, 2013). Email inilah yang disebut dengan spam email penerimanya (Widiasari dan Bayu, 2013). Spam email adalah email yang tidak penting atau tidak berarti, yang dikirim oleh seseorang ke banyak penerima atau pengguna email, pada umumnya berisi promosi, informasi sampah, bahkan bisa juga berisi informasi penipuan (Chandra, Indrawan, & Sukajaya, 2016) (Vyas, Prajapati, & Gadhwal, 2015)

Mengirimkan email spam adalah sebuah pelanggaran terhadap Acceptable Use Policy (AUP) atau peraturan penggunaan yang bisa diterima pada hampir semua Internet Service Provider (ISP), dan dapat menyebabkan penghapusan pada account pengirim (Ananda, 2011). Banyak pengguna email yang merasa terganggu dengan adanya spam (Supri, 2010). Dampak buruk yang paling utama dari spam email adalah waktu yang terbuang dengan percuma untuk menghapus spam (Widiasari dan Bayu, 2013).

Untuk mengatasi hal ini diperlukan suatu filter antispam dengan algoritma tertentu yang dapat memisahkan antara spam-mail dengan non spam-mail (Supri, 2010). Algoritma Naive Bayes adalah yang paling sering digunakan karena kesederhanaan konsepnya. Naive bayes bekerja berdasarkan kemunculan fitur atau ciri terhadap suatu kelas yang kemudian dihitung probabilitasnya, sehingga sesuai dengan karakteristik fitur atau ciri email yang dijadikan acuan untuk membedakan mana yang termasuk spam email dan mana yang tidak (Chandra, Indrawan, & Sukajaya, 2016) (Rusland, Wahid, Kasim, & Hafit, 2017).

Penelitian ini akan menggunakan algoritma Naive Bayes untuk melakukan deteksi spam email, sehingga dari dataset email akan dikelompokkan menjadi 2

yaitu spam (spam email) dan ham (bukan spam email), yang selanjutnya akan dilakukan analisis terhadap hasil klasifikasi yang telah didapatkan, sedangkan dataset yang dipakai adalah mengambil dari UCI Machine Learning Repository yaitu dataset SMS Spam Collection.

Dataset SMS Spam Collection:

<https://archive.ics.uci.edu/dataset/228/sms+spam+collection>

BAB II

LANDASAN TEORI

2.1 Spam Email

SPAM merupakan akronim dari Stupid Pointless Annoying Message (Ananda, 2011). Spam email yaitu email yang tidak diinginkan atau diminta oleh penerimanya (Widiasari dan Bayu, 2013). Spam muncul pertama kali pada bulan Mei tahun 1978. Spam tersebut bersifat iklan yang dikirimkan oleh Digital Equipment Corporation (DEC) (Widiasari dan Bayu, 2013). Spam juga dapat berupa pengiriman pesan secara berulang-ulang ke berbagai newsgroup atau server milis dengan pokok bahasan yang tidak berkaitan (Ananda, 2011).

Tipe-tipe email spam (Supri, 2010):

1. Iklan
2. Spam dapat digunakan untuk mempromosikan suatu produk ataupun layanan.
3. Mengirimkan Malware
4. Spam merupakan salah satu cara utama untuk mendistribusikan virus dan malware.
5. Phishing
6. Bersembunyi dibalik nama besar perusahaan, lembaga keuangan, lembaga pemerintah, lembaga amal, para phisher mencoba memikat korban untuk mengunjungi website palsu.
7. Scam
8. Berita elektronik dalam internet yang bersifat menipu sehingga pengirimnya dapat mendapatkan manfaat atau keuntungan.
9. Pesan yang tak berarti
10. Sebuah potongan pesan sampah seperti ini dapat memenuhi inbox mail kita.

2.2 Klasifikasi

Klasifikasi adalah proses dengan model yang menggambarkan dan membedakan kelas data atau konsep. Model merupakan analisis objek yang label kelasnya belum diketahui (Widiasari dan Bayu, 2013). Klasifikasi pada penelitian kami menggunakan metode *Naive Bayes* yang dapat mengklasifikasikan email sebagai spam atau not spam.

2.3 Naive Bayes

Naive Bayes merupakan suatu metode klasifikasi yang menggunakan perhitungan probabilitas (Ling dkk, 2014). Teori Naive Bayes diadopsi dari nama penemunya yaitu Thomas Bayes sekitar tahun 1950 (Andriansyah dan Suhendra, 2005). Naive Bayes menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi

nilai dari dataset yang diberikan (Sukardi dkk, 2014). Persamaan Metode Naive Bayes (Sukardi dkk, 2014):

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

Keterangan:

X : Data dengan class yang belum diketahui

H : Hipotesis data merupakan suatu class spesifik

$P(H|X)$: Probabilitas hipotesis H berdasar kondisi X (posterior probabilitas)

$P(H)$: Probabilitas hipotesis H (prior probabilitas)

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$: Probabilitas X

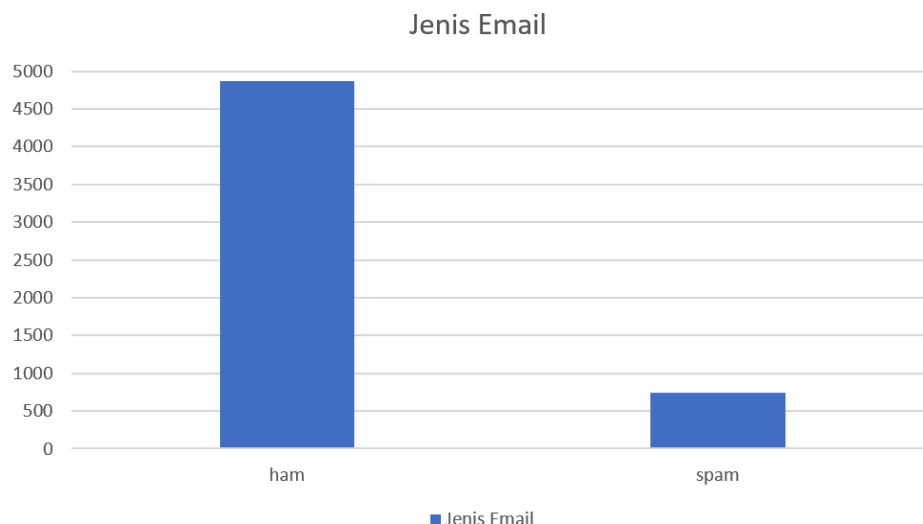
BAB III

HASIL DAN PEMBAHASAN

3.1 Dataset

Dataset yang digunakan diambil dari UCI Machine learning Repository yaitu dataset SMS Spam Collection, dataset ini mengandung koleksi antara 425 pesan spam SMS yang diambil secara manual dari situs Web Grumbletext. Ini adalah forum di Inggris di mana pengguna ponsel membuat klaim publik tentang pesan spam SMS, sebagian besar dari mereka tanpa melaporkan pesan spam yang diterima. Identifikasi teks pesan spam adalah tugas yang sangat sulit dan memakan waktu, dan melibatkan pemindaian ratusan halaman web dengan hati-hati. Daftar 450 pesan SMS ham yang dikumpulkan dari Tesis PhD Caroline Tag. Subset dari 3.375 pesan SMS ham dari NUS SMS Corpus (NSC), yang merupakan sebuah korpus yang terdiri dari sekitar 10.000 pesan yang sah yang dikumpulkan untuk penelitian di Departemen Ilmu Komputer di National University of Singapore. Pesan-pesan tersebut sebagian besar berasal dari warga Singapura dan sebagian besar dari mahasiswa yang kuliah di Universitas. Pesan-pesan ini dikumpulkan dari para sukarelawan yang telah diberitahu bahwa kontribusi mereka akan tersedia untuk umum. Jumlah 1.002 pesan SMS ham dan 322 pesan spam yang diekstrak dari SMS Spam Corpus v.0.1 Big yang dibuat oleh Jos Mar a G mez Hidalgo.

Total dataset yang kami miliki adalah 5574 pesan email dengan rincian 4827 data dikategorikan sebagai email ham (not spam email) dan 747 data dikategorikan sebagai email spam (spam email).




```
File Edit View
ham Go until jurong point, crazy.. Available only in bugis n great world la e buffet... Cine there got amore wat...
ham Ok lar... Joking wif u oai...
spam Free entry in 2 a wely comp to win FA Cup final tixts 21st May 2005. Text FA to 87121 to receive entry question(std txt rate)&c's apply 08452810075over18's
ham U dun say so early hor... U c already then say...
ham Nah I don't think he goes to usf, he lives around here though
spam Ereasing my there darling it's been 3 week's now and no word back! I'd like some fun you up for it still? Tb ok! xox std chgs to send, £1.50 to rcy
ham Even my brother is not like to speak with me. They treat me like aids patent.
ham As per your request 'Belle Belle (Osu Minamunginte Nungu Yottam)' has been set as your callertune for all Callers. Press *9 to copy your friends Callertune
spam WINNER!! As a valued network customer you have been selected to excelsior £900 prize reward! To claim call 08007071461. Claim code KL341. Valid 12 hours only.
ham Had your mobile 11 months or more? U R entitled to update to the latest colour mobiles with camera for Free! Call The Mobile Update Co FREE on 08002986030
ham I'm gonna be home soon and i don't want to talk about this stuff anymore tonight, k? I've cried enough today.
spam SIX chances to win CASH! From 100 to 20,000 pounds txt: C5H11 and send to 87575. Cost 150p/day, 6days, 16+ Standcs apply Reply HL 4 info
ham URGENT!! You have won a 1 week FREE membership in our £100,000 Prize Jackpot! Txt the word: CLAIM to No: 81010 T&C www.duk.net LCCLTD POBOX 4403LDNW1A7R9J18
ham I've been searching for the right words to thank you for this breather. I promise i wont take your help for granted and will fulfil my promise. You have been wonderful and a blessing at all times.
ham I HAVE A DATE ON SUNDAY WITH WILL!!
spam XXXMobileMovieClub: To use your credit, click the WAP link in the next txt message or click here>> http://wap. xxxmobilemovieclub.com?n=QJXG1GH4J3GCB1
ham Oh k...I'm watching here:)
ham Eh u remember how 2 spell his name... Yes i did. He v naughty make until i v wet.
ham Fine if thatz the way u feel. Thatz the way its gota b
spam England v Macedonia - dont miss the goals/team news. Txt uc national team to 87077 eg ENGLAND to 87077 TrxWALS, SCOTLAND 4txt/£1.20 POBOX636504M45WQ 16+
ham Is that seriously how you spell his name?
ham I'm going to try for 2 months ha ha only joking
ham So u pay first lar... Then when is da stock comin...
ham Aff i finish my lunch then i go str down lor. grd 3 sat lor. U finish uc lunch already?
ham Effffff. Alright no way i can meet up with you sooner?
ham Just forced myself to eat a slice. I'm really not hungry tho. This sucks. Mark is getting worried. He knows I'm sick when I turn down pizza. Lol
ham Lol your always so convincing.
ham Did you catch the bus ? Are you frying an egg ? Did you make a tea? Are you eating your mom's left over dinner ? Do you feel my Love ?
ham I'm back kamp; we're packing the car now. I'll let you know if there's room
ham Abhh.. work. I vaguely remember that! what does it feel like? lol
ham Wait that's still not all that clear, were you not sure about me being sarcastic or that that's why x doesn't want to live with us
ham Yeah he got in at 2 and was v apologetic. n had fallen out and she was actin like spoilt child and he got caught up in that. Till 2! But we won't go there! Not doing too badly
cheers. You?
ham K tell me anything about you.
ham For fear of fainting with the of all that housework you just did? Quick have a cuppa
spam Thanks for your subscription to Ringtone UK your mobile will be charged £5/month please confirm by replying YES or NO. If you reply NO you will not be charged
ham Yup... OK i go home look at the timings then i msg u again... Xuhd going to learn o 2nd may too but her lesson is at 8am
ham Oops, I'll let you know when my roommate's done
ham I see the letter B on my car
ham Amvktine 1m u 4ur 4fe
Ln 3177, Col 1 7 of 477203 characters 100% Unix (LF) UTF-8
```

3.2 Implementasi Spam Email Detection

Pada project ini, kami menggunakan Google Colab dan Python sebagai implementasi untuk mendeteksi email spam menggunakan algoritma *Naive Bayes*. Berikut Kode Program yang kami buat:

1. Download Dataset

Dataset pada kode Google Colab tersebut digunakan untuk deteksi spam email menggunakan algoritma Naive Bayes. Dataset tersebut berisi contoh-contoh email yang dikategorikan sebagai spam atau non-spam, yang digunakan untuk melatih model Naive Bayes dan mengklasifikasikan email yang baru masuk sebagai spam atau non-spam berdasarkan fitur-fitur yang diekstrak dari teks email.

```
[1] !wget https://archive.ics.uci.edu/ml/machine-learning-databases/00228/smsspamcollection.zip
!unzip smsspamcollection.zip

--2024-06-12 11:48:05-- https://archive.ics.uci.edu/ml/machine-learning-databases/00228/smsspamcollection.zip
Resolving archive.ics.uci.edu (archive.ics.uci.edu)... 128.195.10.252
Connecting to archive.ics.uci.edu (archive.ics.uci.edu)|128.195.10.252|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: unspecified
Saving to: 'smsspamcollection.zip'

smsspamcollection.z  [ <=> ] 198.65K --.-KB/s in 0.1s

2024-06-12 11:48:05 (1.83 MB/s) - 'smsspamcollection.zip' saved [203415]

Archive: smsspamcollection.zip
  inflating: SMSSpamCollection
  inflating: readme
```

2. Import Library

Kode ini memuat beberapa library yang diperlukan untuk mengimplementasikan algoritma Naive Bayes. Library yang digunakan adalah:

- numpy: untuk melakukan operasi numerik dan manipulasi array.
- pandas: untuk bekerja dengan data dalam bentuk tabel dan melakukan operasi data.
- sklearn: untuk mengimplementasikan algoritma Naive Bayes dan melakukan evaluasi model.

```

✓ [2] import pandas as pd
      import numpy as np
      from sklearn.feature_extraction.text import CountVectorizer
      from sklearn.model_selection import train_test_split
      from sklearn.naive_bayes import MultinomialNB
      from sklearn.metrics import accuracy_score, classification_report

```

3. Load Dataset

Dataset diproses agar file dataset dapat dibaca oleh program dan dapat digunakan untuk melatih program naive bayes dalam mengklasifikasikan jenis email.

```

✓ [3] df = pd.read_csv('SMS Spam Collection', sep='\t', names=['label', 'message'])
      df.head()

```

	label	message
0	ham	Go until jurong point, crazy.. Available only ...
1	ham	Ok lar... Joking wif u oni...
2	spam	Free entry in 2 a wkly comp to win FA Cup fina...
3	ham	U dun say so early hor... U c already then say...
4	ham	Nah I don't think he goes to usf, he lives aro...

Next steps: [View recommended plots](#)

4. Data Pre-processing

Data diproses dengan mengkonversi teks ke huruf kecil, menghapus tanda baca, dan menghapus kata-kata yang tidak relevan. Data pre-processing melibatkan converting label menjadi binary dan pemisahan data menjadi data training untuk melatih program dan data testing untuk menguji kinerja program.

```

✓ [4] # Encode labels: ham = 0, spam = 1
      df['label'] = df['label'].map({'ham': 0, 'spam': 1})

      # Split dataset into training and testing sets
      X_train, X_test, y_train, y_test = train_test_split(df['message'], df['label'], test_size=0.2, random_state=42)

      # Vectorize text data
      vectorizer = CountVectorizer()
      X_train_vec = vectorizer.fit_transform(X_train)
      X_test_vec = vectorizer.transform(X_test)

```

5. Training Model Naive Bayes

Kode ini memanfaatkan fungsi MultinomialNB() pada library sklearn untuk melatih program sesuai algoritma naive bayes.

```

✓ [5] # Train Naive Bayes model
      model = MultinomialNB()
      model.fit(X_train_vec, y_train)

```

▼ MultinomialNB
MultinomialNB()

6. Evaluasi Menggunakan Data Testing

Kode ini digunakan untuk mengklasifikasikan dataset program. Kode ini menggunakan fungsi `classification_report()` untuk menghitung laporan klasifikasi. Laporan klasifikasi menunjukkan performa program dalam mengklasifikasikan dataset. Kinerja program dievaluasi menggunakan beberapa metrik evaluasi:

- Akurasi: Akurasi adalah rasio jumlah email yang diklasifikasikan dengan benar terhadap jumlah total email.
- Presisi: Presisi adalah rasio jumlah email yang diklasifikasikan sebagai spam yang sebenarnya spam terhadap jumlah total email yang diklasifikasikan sebagai spam.
- Recall: Recall adalah rasio jumlah email yang diklasifikasikan sebagai spam yang sebenarnya spam terhadap jumlah total email yang sebenarnya spam.
- F1-score: F1-score adalah rasio dari presisi dan recall.

Kinerja model dievaluasi menggunakan beberapa metrik evaluasi. Hasil evaluasi digunakan untuk mengevaluasi kinerja model dan memperbaiki model jika diperlukan.

```
[6] # Predict on test data
    y_pred = model.predict(X_test_vec)

    # Evaluate model
    accuracy = accuracy_score(y_test, y_pred)
    print(f'Accuracy: {accuracy:.2f}')

    # Classification report
    print(classification_report(y_test, y_pred))
```

➡ Accuracy: 0.99

	precision	recall	f1-score	support
0	0.99	1.00	1.00	966
1	1.00	0.94	0.97	149
accuracy			0.99	1115
macro avg	1.00	0.97	0.98	1115
weighted avg	0.99	0.99	0.99	1115

7. Prediksi dan Deteksi Email

Algoritma Naive Bayes pada program ini dipakai untuk memprediksi kelas email yang baru. Email yang baru diinputkan ke kode program, dan program menghitung posterior probability untuk setiap kelas. Email diklasifikasikan sebagai spam atau not spam berdasarkan posterior probability yang terbesar. Berikut ini implementasi ketika melakukan percobaan prediksi email baru untuk mendeteksi apakah email tersebut spam atau not spam.

✓
0s

▶

```
# Function to predict if a message is spam or not
def predict_spam(message):
    message_vec = vectorizer.transform([message])
    prediction = model.predict(message_vec)
    return 'Spam' if prediction[0] == 1 else 'Not Spam'

# Test the function
test_message = "Congratulations! You've won a free ticket to Bahamas. Reply WIN to claim."
print(predict_spam(test_message))
```

↔

 Spam

BAB IV

PENUTUP

4.1 Kesimpulan

Deteksi email spam menggunakan algoritma Naive Bayes telah terbukti sederhana dan efektif untuk klasifikasi teks. Dengan meningkatnya penggunaan email sebagai sarana komunikasi, ancaman spam email juga meningkat. Spam email dapat mengganggu dan mengakibatkan hilangnya produktivitas karena waktu yang terbuang untuk menghapusnya.

Penelitian ini menggunakan dataset dari UCI Machine Learning Repository yang berisi koleksi SMS spam dan ham (non-spam). Model Naive Bayes dilatih menggunakan dataset ini dan diimplementasikan menggunakan Google Colab dengan bahasa pemrograman Python. Data yang digunakan terdiri dari 5574 pesan email, dimana 4827 dikategorikan sebagai ham dan 747 sebagai spam.

Proses implementasi mencakup:

1. Pre-processing Data: Mengkonversi teks ke huruf kecil, menghapus tanda baca, dan kata-kata yang tidak relevan.
2. Training Model: Menggunakan algoritma Naive Bayes untuk melatih model dengan dataset yang telah diproses.
3. Evaluasi Model: Menggunakan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score untuk menilai performa model dalam mengklasifikasikan email.
4. Prediksi Email: Menginputkan email baru untuk diprediksi/dideteksi apakah email tersebut termasuk email spam atau not spam.

Hasil evaluasi menunjukkan bahwa program Spam Email Detection menggunakan metode Naive Bayes mampu memfilter email spam dengan cukup baik berdasarkan fitur-fitur teks. Penggunaan Naive Bayes dalam deteksi spam email dapat menjadi solusi efektif dalam mengatasi masalah spam email.

4.2 Saran

Diharapkan pembaca dapat mengambil manfaat dari project yang telah kami buat dan pembaca dapat membantu menyempurnakan laporan ini dikarenakan penulis menyadari bahwa laporan kami jauh dari kata sempurna. Kami sangat menghargai kritik serta saran dari para pembaca sebagai motivasi agar para penulis lebih berkembang.