

# Deteksi Email Spam dan Non Spam Bahasa Indonesia Menggunakan Metode Naïve Bayes

*by* John Doe

---

**Submission date:** 09-Jun-2021 02:46PM (UTC+0700)

**Submission ID:** 1603328646

**File name:** Paper\_kolokium\_17523225\_fix.docx (51.69K)

**Word count:** 1917

**Character count:** 12274

# Deteksi Email Spam dan Non Spam Bahasa Indonesia Menggunakan Metode Naïve Bayes

**Abstrak** — Penggunaan email yang mudah saat ini banyak sekali dimanfaatkan banyak orang sehingga menimbulkan dampak positif maupun negatif. Email negatif biasa kita sebut dengan email spam yang berisi berupa iklan, penipuan, virus dan malware yang berpotensi untuk merugikan orang lain. Sehingga masalah tersebut memerlukan penanganan untuk mengatasinya. Penelitian ini bertujuan untuk membuat sebuah model klasifikasi email spam dan non spam berbahasa Indonesia dengan menggunakan algoritma *Naïve Bayes*. Berdasarkan hasil penelitian yang telah dilakukan bahwa algoritma *Naïve Bayes* dengan menggunakan fitur N-gram telah berhasil melakukan klasifikasi sangat baik dengan nilai akurasi 89% hingga 94%, Nilai *precision* 89% hingga 99%, *recall* 97%, dan *F-Score* 97%

**Kata Kunci**—Email, Spam, Klasifikasi, Naïve Bayes

## I. PENDAHULUAN

Email merupakan sarana komunikasi dalam jaringan internal maupun internet untuk pertukaran informasi. Email masih digunakan hingga saat ini karena kemudahan dalam hal penggunaannya. Saat ini selain digunakan untuk komunikasi email juga digunakan untuk kebutuhan otentikasi aplikasi dan sinkronisasi media sosial seperti instagram, facebook dan twitter. Penggunaan email yang tinggi bisa berdampak positif dan berdampak negatif karena tidak semua orang menggunakan email dengan baik dan bahkan ada banyak sekali penyalahgunaan email sehingga berpotensi merugikan pengguna email lainnya. Email yang disalahgunakan ini disebut sebagai spam atau email sampah yang mana memiliki konten tentang iklan, Scam, dan virus[1].

Email spam yang beredar di kalangan pengguna sebenarnya memiliki pola tertentu hanya saja banyak sekali pengguna awam tidak banyak mengetahui. biasanya kasus yang banyak terjadi adalah email spam berjenis iklan yang memenuhi kotak masuk email korban padahal email tersebut tidak diinginkan. Spam dapat menyebabkan ketidakefisienan bandwidth karena merupakan kapasitas dari sebuah jaringan agar dapat dilewati oleh paket data. Bagi banyak orang hal ini sangat mengganggu sehingga dibutuhkan penanganan mengatasi email spam ini.

Permasalahan ini dapat diminimalisir dengan membuat sebuah model anti spam yang bertujuan untuk mengklasifikasikan email dan memberikan informasi terhadap pengguna email apabila terdapat pesan yang diprediksi sebagai pesan spam. Salah satu metode menciptakan anti spam adalah dengan metode Naive Bayes untuk mengklasifikasikan email spam dan non spam. Penelitian menggunakan metode naïve bayes sebenarnya telah banyak dilakukan untuk mengklasifikasikan email spam berbahasa inggris namun penelitian untuk email yang berbahasa Indonesia sangat jarang dilakukan. Oleh karena itu, penelitian ini akan mengukur metode tersebut seberapa baik performa metode Naive Bayes untuk menangani permasalahan email spam yang berbahasa Indonesia.

## II. LANDASAN TEORI

### A. Spam Email

Spam atau *stupid pointless annoying messages* merupakan serangan pesan yang dikirimkan ke sejumlah pengguna layanan pesan yang tidak meminta pesan tersebut. Spam juga dapat didefinisikan sebagai pengiriman pesan secara terus-menerus. berikut merupakan tipe-tipe email spam [2]:

Iklan : digunakan untuk mempromosikan suatu barang atau layanan yang dimiliki suatu perusahaan maupun individu perorangan.

Phising: menyamar sebagai perusahaan besar/lembaga terpercaya untuk memikat para korban untuk mengunjungi situs web palsu yang tertera dalam pesan dan mengambil data pribadi korban. Malware: memperdaya korban dengan mengirimkan sebuah file yang berisikan sebuah virus malware.

Scam: upaya penyamaran yang dilakukan untuk mendapatkan simpati korban sehingga bisa mendapatkan sesuatu hal yang berharga seperti data maupun uang.

Perbedaan Spam dan Non Spam dapat dilihat dari struktur email sebagai berikut:

Subject: merupakan judul topik yang mewakili isi email biasanya dalam email spam terdapat kata-kata “Ada Diskon” yang sering dijumpai pada korban yang terkena serangan email spam.

Body: merupakan inti dari pesan email yang diberikan dan isi email spam sangat mudah dikenali dengan melihat kata-kata yang dikirimkan oleh pengirim

### B. Crowdsourcing

Crowdsourcing adalah suatu kegiatan yang dilakukan untuk mendapatkan sebuah ide dan data untuk menyelesaikan masalah yang kompleks dengan tidak memandang latar belakang pendidikan, kewarganegaraan, agama, amatir maupun profesional setiap individu diperbolehkan untuk ikut berpartisipasi dengan pengetahuan dan pengalaman sehingga dalam permasalahan yang ada dapat ditangani secara cepat, tepat, dan hemat biaya.

Pengguna nantinya akan mendapatkan kepuasan dari hasil yang telah didapat baik itu ekonomi, pengkayaan sosial, maupun pengembangan keterampilan individu [3].

### C. Text Mining

Text mining adalah proses yang dilakukan untuk menggali data dengan format teks. Text Mining mempunyai tujuan untuk mengambil kata dan memperoleh sebuah informasi sehingga dari hasil yang didapat bisa dilakukan sebuah analisa yang memiliki nilai untuk kepentingan tertentu. Terdapat beberapa tahapan proses dalam implementasi text mining yaitu text preprocessing, teks tranformasi, seleksi fitur, dan pattern discovery[4].

#### D. Naive Bayes

Merupakan sebuah algoritma yang dikemukakan oleh ilmuwan Inggris Thomas Bayes untuk pengklasifikasian yang dihitung dari gabungan probabilitas dengan melakukan penjumlahan antar frekuensi dan kombinasi nilai dari data yang ada. Kelebihan dalam menggunakan algoritma naive Bayes adalah metode ini hanya membutuhkan training yang sedikit[4].

Berikut merupakan persamaan naive Bayes:

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

Keterangan:

X : Data dengan kelas yang belum diketahui.

H : Hipotesis data merupakan suatu kelas spesifik

P(H|X) : Probabilitas hipotesis H berdasarkan 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

Dalam Bayes memiliki aturan jika  $P(H1|x) < P(H2|x)$ , maka x diklasifikasikan sebagai h2. Pernyataan  $P(H1|x)$  mengindikasikan probabilitas hipotesis h1 berdasarkan kondisi x terjadi, begitu pula dengan h2. Sehingga dapat klasifikasi dari x sesuai dengan probabilitas terbesar diantara probabilitas x terhadap semua kelas.

#### E. Performance Evaluation Measure

PEM memiliki tujuan untuk mengevaluasi model yang telah dibuat dan merepresentasikan sebuah kondisi aktual dari data yang dihasilkan oleh model yang digunakan. Banyak perhitungan untuk mendapat hasil dari nilai PEM yaitu[4]:

##### - Precision

Mengukur tingkat kepastian terhadap data yang diklasifikasi secara tepat oleh model yang telah dibangun[5].

Rumus Precision:

$$pre = \frac{TP}{TP + FP}$$

##### - Accuration

merupakan perbandingan antara informasi yang dijawab oleh sistem dengan benar oleh keseluruhan informasi

Rumus Accuration:

$$acc = \frac{TN}{FP + FN + TP + TN}$$

##### - Recall

Recall mengukur sensitifitas model terhadap data yang berlabel positif dengan benar

Rumus Recall:

$$rec = \frac{TP}{FN + TP}$$

Keterangan:

TP: prediksi positif dan dinilai benar sebagai positif

TN: prediksi negatif dan dinilai benar sebagai negatif

FP: prediksi negatif dan dinilai salah sebagai positif

FN: prediksi positif dan dinilai salah sebagai negatif

##### - F-Score

F-Score adalah perbandingan rata-rata recall dan presisi.

Rumus F-Score:

$$Fscore = \frac{2 \times Recall \times Precision}{Recall + precision}$$

#### F. Penelitian Terkait

Terdapat beberapa penelitian tentang klasifikasi email spam telah dilakukan sebelum penelitian ini. Hengki, et al.[6] memakai algoritma Naive Bayes dan SVM dengan berbasis PSO untuk mengklasifikasikan email berbahasa Inggris dan keduanya memiliki akurasi yang sangat baik. Chandra et al.[7] melakukan perbandingan antara metode Pos Tagger dan Naive Bayes dalam mengklasifikasikan dari hasil penelitian yang dilakukan Naive Bayes memiliki skor yang lebih tinggi dibandingkan Pos Tagger.

Menurut Juang[8] Dari hasil penelitian tersebut bahwa algoritma naive Bayes dapat mengklasifikasikan email ke dalam dua kelas yaitu spam dan non spam. Hasil dari proses tersebut sangat dipengaruhi oleh proses training model.

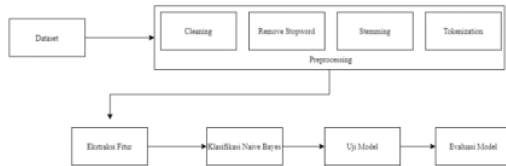
Hayuningtyas[2] melakukan penelitian menggunakan algoritma Naive Bayes dan menggunakan confusion matrix sebagai evaluasi model. Hasil pengujian menggunakan Naive Bayes sangat baik dalam mendeteksi spam dan memiliki tingkat akurasi 75,9%.

Alwani, et al[9] melakukan penelitian untuk membangun email spam filtering berbahasa Arab dalam penelitiannya mereka menggunakan fitur N-gram dan menggunakan algoritma Bayesian. Keakuratan spam filtering berbahasa Arab yang mereka bangun sangat baik mencapai 80%.

Abdulhamid, et al [10] melakukan perbandingan antar algoritma untuk email spam detection dalam penelitiannya mereka membandingkan algoritma Rotation Forest, Random Tree, NNge, Radial Basis Function (RBF) Network, Lazy Bayesian Rule, Logit Boost, Hidden Naive Bayes, Bayesian Logistic Regression, J48, REP Tree, Naive Bayes, Voted Perceptron, Multilayer Perceptron, dan Logistic Model Tree. Hasil yang didapat algoritma Rotation Forest mendapatkan akurasi paling tinggi dan algoritma yang memiliki akurasi paling rendah untuk spam detection adalah algoritma REP Tree.

### 6 III. METODOLOGI PENELITIAN

Tahapan proses dalam penelitian ini dapat dilihat pada Gambar I.



GAMBAR I. ALUR PENELITIAN

#### A. Dataset

Pengumpulan data dilakukan dengan metode *crowdsourcing* yaitu dengan meminta sejumlah email berbahasa Indonesia kepada setiap individu yang terindikasi terkena serangan email spam. Jumlah data yang didapat sebanyak 317 *Email Spam* dan 300 *Email Non Spam*.

#### B. Pre-processing

Terdapat langkah-langkah yang dilakukan untuk menjadikan data yang telah didapat sebelumnya menjadi data terstruktur, yaitu:

1. *Cleaning*: Membersihkan data yang didapat dari tanda baca atau punctuation, hashtag maupun mention, dan mengubah semua kata menjadi lowercase.

TABEL I. CLEANING

| Sebelum  | Sesudah  |
|--|--|
| saya sudah menginstal = virus Trojan pada Sistem Operasi | saya sudah menginstal virus Trojan pada sistem operasi |
| Absensi bisa langsung menemui saya di ruangan,           | absensi bisa langsung menemui saya di ruangan          |

2. *Remove Stopwords*: Penghapusan kata-kata yang kurang memiliki makna yang berarti seperti kata: dan, saya, atau.

TABEL II. REMOVE STOPWORDS

| Sebelum  | Sesudah                                |
|--|--|
| saya sudah menginstal virus Trojan pada sistem operasi | menginstal virus trojan sistem operasi |
| absensi bisa langsung menemui saya di ruangan          | absensi langsung menemui ruangan       |

3. *Stemming*: Mengubah kata-kata yang memiliki afiks menjadi sebuah kata dasar aslinya.

TABEL III. STEMMING

| Sebelum | Sesudah |
|---------|---------|
|---------|---------|

|  |                                     |
|--|-------------------------------------|
| menginstal virus trojan sistem operasi | install virus trojan sistem operasi |
| absensi langsung menemui ruangan       | absen langsung nemu ruang           |

4. *Tokenization*: Memecah dokumen menjadi bagian-bagian yang lebih kecil sehingga memudahkan untuk analisa.

TABEL IV. TOKENIZATION

| Sebelum                             | Sesudah   |
|-------------------------------------|---|
| install virus trojan sistem operasi | ['install', 'virus', 'trojan', 'sistem', 'operasi'] |
| absen langsung nemu ruang           | ['absen', 'langsung', 'nemu', 'ruang']              |

#### C. Ekstraksi Fitur

Tahapan yang dilakukan setelah tahap preprocessing bertujuan untuk mempermudah proses klasifikasi. Tahapan ini menggunakan fitur N-gram yaitu memecah kalimat sesuai dengan nilai n yang telah ditentukan. Misal apabila n=2 sama saja dapat disebut *bigram* maka kalimat akan dibagi menjadi dua kata pada setiap data. Metode selanjutnya adalah dengan menggunakan *term frequency* yaitu digunakan untuk melakukan perhitungan pembobotan dengan menghitung kemunculan istilah tertentu pada data yang diolah.[5]

#### D. Klasifikasi Naïve Bayes

Metode yang dipakai dalam mengklasifikasikan data email untuk mendapatkan prediksi spam atau non spam. Untuk proses klasifikasi akan digunakan data hasil preprocessing dan ekstraksi fitur yang dilakukan pada tahapan sebelumnya. Setelah melakukan proses training lalu akan diuji menggunakan data uji sehingga mendapatkan ketepatan hasil klasifikasi yang dilakukan oleh model yang telah dibuat.

#### E. Uji Model

Proses ini dilakukan setelah training data selesai dengan tujuan untuk mengetahui kinerja model yang telah kita buat.

#### F. Evaluasi Model

Evaluasi dan Validasi Hasil dilakukan dengan cara menghitung nilai *recall*, *precision*, dan *f-score*. *precision* menghitung tingkatan kepastian yang diklasifikasikan dengan benar, *recall* rasio prediksi dari data pada setiap label yang diklasifikasikan dengan benar terhadap data yang salah diklasifikasikan ke label lainnya. *F-score* merupakan *harmonic mean* antara nilai *precision* dan nilai *recall*.

## 1 IV. HASIL DAN PEMBAHASAN

### A. Perhitungan Akurasi pada model klasifikasi

Proses training dan testing dilakukan dengan fitur N-gram dan menggunakan algoritma Naïve Bayes. Berikut merupakan hasil perhitungan akurasi dapat dilihat pada Tabel V.

TABEL V. HASIL PERHITUNGAN MODEL KLASIFIKASI

| N-gram | Akurasi |
|--------|---------|
| 1      | 0,89    |
| 2      | 0,93    |
| 3      | 0,94    |

Berdasarkan Tabel V, Hasil yang didapat menunjukkan bahwa model dengan menggunakan algoritma Naïve Bayes dan menggunakan fitur N-gram memiliki akurasi yang cukup baik. Dari percobaan yang dilakukan nilai n=3 memiliki akurasi yang paling tinggi yaitu 94% dibandingkan dengan n=1 dan n=2 yang masing-masing hanya memiliki akurasi 89% dan 93%.

#### 1 B. Perhitungan Precision, Recall, dan F-Score

Hasil evaluasi menggunakan precision, recall, f-score dapat dilihat dalam Tabel VI.

TABEL VI PERHITUNGAN PRECISION, RECALL, dan F-SCORE

| N-gram | Precision | Recall | F-Score |
|--------|-----------|--------|---------|
| 1      | 0,89      | 0,97   | 0,97    |
| 2      | 0,99      | 0,97   | 0,97    |
| 3      | 0,99      | 0,97   | 0,97    |

Berdasarkan hasil yang ditampilkan pada Tabel VI, untuk nilai F-Score pada table tersebut ketiganya memiliki nilai F-Score yang sama yaitu 97% dan pada nilai precision fitur N-gram dengan nilai n=2 dan n=3 memiliki nilai yang sama yaitu 99%. Pada nilai recall untuk n=1, n=2, n=3 memiliki nilai yang sama yaitu 97%

#### V. KESIMPULAN

1 Dari hasil penelitian ini data email spam bahasa Indonesia memiliki karakteristik yang sangat tidak terstruktur dan memiliki banyak noise seperti tag html, tanda baca, dan kata-kata tidak bermakna sehingga langkah-langkah dalam melakukan preprocessing akan sangat berpengaruh terhadap kualitas model yang dihasilkan. Algoritma Naïve Bayes dapat digunakan untuk mendeteksi email spam berbahasa Indonesia dengan akurasi yang sangat baik. Berdasarkan eksperimen yang telah dilakukan perpaduan fitur N-gram dengan nilai n=3 dan algoritma Naïve Bayes memiliki akurasi yang paling tinggi yaitu 94%

#### REFERENSI

# Deteksi Email Spam dan Non Spam Bahasa Indonesia Menggunakan Metode Naïve Bayes

## ORIGINALITY REPORT

18%

SIMILARITY INDEX

17%

INTERNET SOURCES

9%

PUBLICATIONS

3%

STUDENT PAPERS

## PRIMARY SOURCES

|   |  |    |
|---|--|----|
| 1 | <a href="http://media.neliti.com">media.neliti.com</a><br>Internet Source  | 6% |
| 2 | <a href="http://dspace.uii.ac.id">dspace.uii.ac.id</a><br>Internet Source  | 5% |
| 3 | <a href="http://ejurnal.plm.ac.id">ejurnal.plm.ac.id</a><br>Internet Source  | 1% |
| 4 | Elly Indrayuni. "Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes", Jurnal Khatulistiwa Informatika, 2019<br>Publication   | 1% |
| 5 | Shafi'i Muhammad Abdulhamid, Maryam Shuaib, Oluwafemi Osho, Idris Ismaila, John K. Alhassan. "Comparative Analysis of Classification Algorithms for Email Spam Detection", International Journal of Computer Network and Information Security, 2018<br>Publication | 1% |
| 6 | <a href="http://doku.pub">doku.pub</a><br>Internet Source  | 1% |

---

|    |  |      |
|----|--|------|
| 7  | <a href="https://export.arxiv.org">export.arxiv.org</a><br>Internet Source   | 1 %  |
| 8  | <a href="http://www.neliti.com">www.neliti.com</a><br>Internet Source  | <1 % |
| 9  | <a href="http://tentangmediabarur.wordpress.com">tentangmediabarur.wordpress.com</a><br>Internet Source  | <1 % |
| 10 | Maryam Hasan. "PREDIKSI TINGKAT KELANCARAN PEMBAYARAN KREDIT BANK MENGGUNAKAN ALGORITMA NAÏVE BAYES BERBASIS FORWARD SELECTION", <i>ILKOM Jurnal Ilmiah</i> , 2017<br>Publication                                  | <1 % |
| 11 | Merio Hengki, Mochamad Wahyudi. "Klasifikasi Algoritma Naïve Bayes dan SVM Berbasis PSO Dalam Memprediksi Spam Email Pada Hotline-Sapto", <i>Paradigma - Jurnal Komputer dan Informatika</i> , 2020<br>Publication | <1 % |
| 12 | <a href="http://core.ac.uk">core.ac.uk</a><br>Internet Source  | <1 % |
| 13 | <a href="http://digilibadmin.unismuh.ac.id">digilibadmin.unismuh.ac.id</a><br>Internet Source  | <1 % |
| 14 | <a href="http://download.garuda.ristekdikti.go.id">download.garuda.ristekdikti.go.id</a><br>Internet Source  | <1 % |
| 15 | <a href="http://makalahcopy4paste.blogspot.com">makalahcopy4paste.blogspot.com</a><br>Internet Source  | <1 % |

---

---

16

[www.scribd.com](http://www.scribd.com)

Internet Source

<1 %

---

17

[ojs.uajy.ac.id](http://ojs.uajy.ac.id)

Internet Source

<1 %

---

---

Exclude quotes      Off

Exclude matches      Off

Exclude bibliography      On