



**PENERAPAN ALGORITMA *NAIVE BAYES* UNTUK KLASIFIKASI
PRODUK *SELF DECLARE* (STUDI KASUS : BADAN PENYELENGGARA
JAMINAN PRODUK HALAL)**

SKRIPSI

GUSTIAN ABRARY SHIDQI

NIM. 1910512097

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA**

2023



**PENERAPAN ALGORITMA *NAIVE BAYES* UNTUK KLASIFIKASI
PRODUK *SELF DECLARE* (STUDI KASUS : BADAN PENYELENGGARA
JAMINAN PRODUK HALAL)**

SKRIPSI

**Diajukan sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana
Komputer**

GUSTIAN ABRARY SHIDQI

NIM. 1910512097

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA

2023

PERNYATAAN ORISINALITAS

Skripsi ini merupakan hasil karya sendiri serta semua sumber referensi yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Gustian Abrary Shidqi

NIM : 1910512097

Tanggal : 28 November 2023

Bilamana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 28 November 2023

Yang Menyatakan,



(Gustian Abrary Shidqi)

**PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK
KEPENTINGAN AKADEMIS**

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Gustian Abrary Shidqi
NIM : 1910512097
Fakultas : Ilmu Komputer
Program Studi : S1 - Sistem Informasi

Memberikan persetujuan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non Eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah yang berjudul:

**PENERAPAN ALGORITMA *NAIVE BAYES* UNTUK KLASIFIKASI
PRODUK *SELF DECLARE* (STUDI KASUS : BADAN PENYELENGGARA
JAMINAN PRODUK HALAL)**

Beserta perangkat lain (apabila diperlukan). Hak Bebas Royalti Non Eksklusif ini memberikan hak Universitas Pembangunan Nasional Veteran Jakarta dalam menyimpan, mengalih media, mengelola dalam bentuk pangkalan data, merawat, dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sesungguhnya.

Dibuat di : Jakarta
Pada tanggal : 28 November 2023

Yang Menyatakan,



(Gustian Abrary Shidqi)

LEMBAR PENGESAHAN

Skripsi ini diajukan oleh :

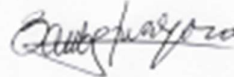
Nama : Gustian Abary Shidqi
NIM : 1910512097
Program Studi : S-1 Sistem Informasi
Judul Tugas Akhir : Penerapan Algoritma Naive Bayes Untuk Klasifikasi
Produk Self Declare (Studi Kasus : Badan Penyelenggara
Jaminan Produk Halal)

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Rio Wirawan, S.Kom., M.M.S.I

Pembimbing I



Bambang Triwahyono, S.Kom., M.Si

Pembimbing II



Ika Nurfaul Isnainiyah, S.Kom., M.Sc.

Penguji I



Niady Intavika, S.Si., M.T

Penguji II



Fyzy Kurniasnik, S.Kom., M.M.

Pit Dekan



Anita Muliawati, S.Kom., M.TI

Kepala Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 15 Desember 2023



ABSTRAK

Badan Penyelenggara Jaminan Produk Halal (BPJPH) adalah unit kerja di bawah Kementerian Agama yang dibentuk sesuai amanat Undang-undang Nomor 33 Tahun 2014 tentang Jaminan Produk Halal. Dalam rangka menyukseskan target 10 juta produk bersertifikat halal, BPJPH, Kementerian/Lembaga, Pemerintah Daerah, dan mitra BPJPH lainnya hadir untuk memberikan penguatan pelaku usaha mikro dan kecil melalui program Sertifikasi Halal Gratis (SEHATI). Sebanyak 1 juta kuota pendaftar UMK dipersiapkan untuk mendapatkan sertifikat halal secara gratis dengan mekanisme pernyataan pelaku usaha (*Self Declare*). *Self Declare* adalah suatu pernyataan status halal produk Usaha Mikro dan Kecil (UMK) oleh pelaku usaha itu sendiri. Pada tahap verifikasi produk yang dilakukan BPJPH, proses klasifikasi produk *Self Declare* masih dilakukan oleh staf BPJPH. Dengan cara ini masih membutuhkan waktu yang lama dan rentan terhadap kesalahan seperti kesalahan pemeriksaan oleh staf BPJPH. Tujuan dari penelitian ini yaitu untuk melakukan klasifikasi produk *Self Declare* atau bukan menggunakan dataset yang berisi data informasi mengenai produk *Self Declare* BPJPH tahun 2022 dari BPJPH. Dalam penelitian ini, data dibagi menjadi data latih dan data uji dengan perbandingan 80% data latih dan 20% data uji. Metode yang digunakan dalam penelitian ini yaitu algoritma *Naive Bayes*. Hasil dari penelitian adalah model algoritma *Naive Bayes* dapat digunakan dalam melakukan klasifikasi produk *Self Declare*. Hasil evaluasi model Algoritma *Naive Bayes* dengan perbandingan data latih dan data uji 80%:20% menghasilkan *accuracy* sebesar 0,94 (94%), *precision* sebesar 0,99 (99%), *recall* sebesar 0,94 (94%), dan *F1-Score* sebesar 0,96 (96%).

Kata Kunci: BPJPH, *Self Declare*, Klasifikasi, *Naive Bayes*

ABSTRACT

The Halal Product Guarantee Administering Body (BPJPH) is a work unit under the Ministry of Religion which was formed in accordance with the mandate of Law Number 33 of 2014 concerning Halal Product Guarantee. In order to succeed in the target of 10 million halal certified products, BPJPH, Ministries/Agencies, Regional Governments, and other BPJPH partners are here to provide strengthening for micro and small business actors through the Free Halal Certification (SEHATI) program. A total of 1 million MSE registrant quotas are prepared to obtain free halal certificates using a self-declaration mechanism. Self Declare is a statement of the halal status of Micro and Small Business (UMK) products by the business actor himself. At the product verification stage carried out by BPJPH, the Self Declare product classification process is still carried out by BPJPH staff. This method still takes a long time and is susceptible to errors such as inspection errors by BPJPH staff. The aim of this research is to classify Self Declare products or not using a dataset containing information data regarding the 2022 BPJPH Self Declare products from BPJPH. In this research, the data is divided into training data and test data with a ratio of 80% training data and 20% test data. The method used in this research is the Naive Bayes algorithm. The results of the research are that the Naive Bayes algorithm model can be used to classify Self Declare products. The results of the evaluation of the Naive Bayes Algorithm model with a comparison of training data and test data of 80%:20% produce an accuracy of 0,94 (94%), precision of 0,99 (99%), recall of 0,94 (94%), and F1-Score of 0,96 (96%).

Keywords: BPJPH, *Self Declare*, Classification, *Naive Bayes*

KATA PENGANTAR

Syukur Alhamdulillah penulis panjatkan atas kehadiran Allah Subhanahu Wataala. Atas perlindungan dan rahmat-Nya penulis dapat menyelesaikan tugas akhir ini dengan judul “Penerapan Algoritma *Naive Bayes* Untuk Klasifikasi Produk *Self Declare* (Studi Kasus : Badan Penyelenggara Jaminan Produk Halal)”.

Penulisan tugas akhir ini dilakukan dalam rangka memenuhi salah satu syarat untuk memperoleh gelar sarjana pada Program Studi S1 Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta. Pada kesempatan ini perkenankanlah penulis menyampaikan ucapan terima kasih yang tulus kepada semua pihak yang telah membantu penulis dalam menyelesaikan penulisan tugas akhir ini. Berikut beberapa pihak yang ingin penulis ucapkan terima kasih kepada:

1. Orang tua yang selalu memberi dukungan kepada penulis selama mengerjakan tugas akhir ini.
2. Ibu Erly Krisnanik, S.Kom., M.M. selaku Plt Dekan Fakultas Ilmu Komputer.
3. Ibu Anita Muliawati, S.Kom., M.TI selaku Kepala Jurusan S1 Sistem Informasi, Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.
4. Bapak Rio Wirawan, S.Kom., MMSI. dan Bambang Triwahyono, S.Kom., M.Si. selaku dosen pembimbing skripsi.
5. Bapak Andhika Octa Indarso, S.Kom., MMSI selaku dosen pembimbing akademik.
6. Bapak Muhammad Yanuar Arief, S.Kom selaku Kepala Subbagian Sistem Informasi & Humas di BPJPH.
7. Teman-teman jurusan S1 Sistem Informasi angkatan 2019 yang memberikan dukungan kepada penulis.

Penulis menyadari bahwa masih terdapat banyak kekurangan di dalam penulisan dan penyusunan tugas akhir ini, oleh karena itu penulis mengharapkan kritik dan saran yang membangun dari pembaca demi perbaikan tugas akhir ini. Semoga tugas akhir ini dapat bermanfaat bagi para pembaca pada umumnya dan bagi penulis pada khususnya.

Jakarta, 21 Oktober 2023

Penulis

DAFTAR ISI

ABSTRAK	vi
ABSTRACT	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan masalah	3
1.3 Ruang Lingkup Penelitian	3
1.4 Tujuan dan Manfaat Penelitian	3
1.4.1 Tujuan Penelitian	3
1.4.2 Manfaat Penelitian	4
1.5 Luaran yang Diharapkan	4
1.6 Sistematika Penulisan	4
BAB II LANDASAN TEORI	6
2.1 Produk	6
2.2 Produk <i>Self Declare</i>	6
2.3 Data Mining	7
2.4 Pre-Processing	8
2.4.1 Data Cleansing	8
2.4.2 Remove Punctuation	8
2.4.3 Case Folding	8
2.4.4 Tokenization	8
2.5 Pembobotan TF-IDF	9
2.6 <i>Synthetic Minority Oversampling Technique (SMOTE)</i>	10
2.7 Klasifikasi	10
2.8 <i>Naive Bayes</i>	10

2.9	<i>Confusion Matrix</i>	11
2.10	Sistem.....	13
2.11	<i>Python</i>	14
2.12	Flask.....	14
2.13	Black Box Testing	15
2.14	Penelitian Terdahulu.....	15
BAB III METODOLOGI PENELITIAN		18
3.1	Alur Penelitian.....	18
3.2	Tahapan Penelitian	18
3.2.1.	Identifikasi Masalah.....	18
3.2.2.	Pengumpulan Data.....	19
3.2.3.	Pre-Processing Data.....	19
3.2.4.	Pembobotan TF-IDF	19
3.2.5.	Pembagian Data.....	19
3.2.6.	Penyeimbangan Data Menggunakan SMOTE	20
3.2.7.	Pemodelan Menggunakan <i>Naive Bayes</i>	20
3.2.8.	Evaluasi Model.....	20
3.2.9.	Perancangan Sistem.....	20
3.2.10.	Pembangunan Sistem	21
3.2.11.	Pengujian Sistem.....	21
3.3	Perangkat Penelitian	21
3.3.1.	Perangkat Keras (Hardware).....	21
3.3.2.	Perangkat Lunak (Software)	22
3.4	Jadwal Penelitian.....	22
BAB IV HASIL DAN PEMBAHASAN		23
4.1	Profil Badan Penyelenggara Jaminan Produk Halal (BPJPH)	23
4.1.1	Visi dan Misi	23
4.1.2	Struktur Organisasi	23
4.1.3	Self Declare.....	24
4.2	Dataset.....	25
4.3	<i>Preprocessing Data</i>	26
4.3.1	<i>Remove Punctuation</i>	26

4.3.2	<i>Case Folding</i>	27
4.3.3	<i>Data Cleansing</i>	28
4.4	Pembobotan TF-IDF	29
4.5	Pembagian Data.....	31
4.6	SMOTE (<i>Synthetic Minority Oversampling Technique</i>)	31
4.7	Implementasi Algoritma <i>Naive Bayes</i>	32
4.8	Evaluasi Model.....	32
4.9	Perancangan Sistem	36
4.9.1	<i>Use Case Diagram</i>	36
4.9.2	<i>Wireframe</i>	36
4.10	Pembangunan Sistem.....	38
4.10.1	Halaman Beranda	39
4.10.2	Halaman Tentang Kami	40
4.10.3	Halaman Prediksi.....	40
4.10.4	Halaman Hasil Prediksi	41
4.11	Pengujian Sistem dengan <i>Blackbox Testing</i>	41
BAB V PENUTUP		43
5.1	Kesimpulan	43
5.2	Saran	43
DAFTAR PUSTAKA		45
DAFTAR RIWAYAT HIDUP		48
LAMPIRAN		49

DAFTAR GAMBAR

Gambar 1. Alur Pendaftaran Program SEHATI.....	7
Gambar 2. <i>Confusion Matrix</i>	12
Gambar 3. Alur Penelitian.....	18
Gambar 4. Struktur Organisasi.....	24
Gambar 5. Dataset	25
Gambar 6. Grafik perbandingan kelas	26
Gambar 7. Kode program penghapusan tanda baca kecuali koma.....	27
Gambar 8. Kode program <i>Case Folding</i>	28
Gambar 9. Kode program penghapusan angka dan spasi berlebihan	29
Gambar 10. Kode program Split data	31
Gambar 11. Kode program oversampling dengan menggunakan SMOTE	31
Gambar 12. Kode program pemodelan algoritma <i>Naive Bayes</i>	32
Gambar 13. Kode program evaluasi model.....	35
Gambar 14. <i>Use Case Diagram</i>	36
Gambar 15. <i>Wireframe</i> Halaman Beranda.....	37
Gambar 16. <i>Wireframe</i> Halaman Tentang Kami.....	37
Gambar 17. <i>Wireframe</i> Halaman Prediksi	38
Gambar 18. <i>Wireframe</i> Halaman Hasil Prediksi	38
Gambar 19. Menjalankan Sistem Klasifikasi Melalui Visual Studio Code.....	39
Gambar 20. Halaman Beranda	39
Gambar 21. Halaman Tentang Kami	40
Gambar 22. Halaman Prediksi.....	40
Gambar 23. Halaman Prediksi yang sudah diisi.....	41
Gambar 24. Halaman Hasil Prediksi.....	41

DAFTAR TABEL

Tabel 1. Penelitian Terdahulu	15
Tabel 2. Jadwal Kegiatan Penelitian.....	22
Tabel 3. Remove Punctuation	27
Tabel 4. Case Folding	28
Tabel 5. <i>Data Cleansing</i>	29
Tabel 6. Perhitungan TF-IDF	30
Tabel 7. Confusion Matrix 80%:20%.....	32
Tabel 8. Confusion Matrix 70%:30%.....	33
Tabel 9. Confusion Matrix 60%:40%.....	34
Tabel 10. Evaluasi Hasil Akurasi Percobaan	35
Tabel 11. <i>Blackbox Testing</i>	42

DAFTAR LAMPIRAN

Lampiran 1 Surat Permohonan Riset	49
Lampiran 2 Dataset.....	50
Lampiran 3 Kode Program app.py.....	64
Lampiran 4 Kode Program Python Google Collab	65
Lampiran 5 Hasil Uji Turnitin.....	71