

**PREDIKSI TRAFIK INTERNET HARIAN  
MENGUNAKAN ALGORITMA *NAIVE BAYES*  
*CLASSIFICATION***



**SKRIPSI**

**diajukan untuk memenuhi salah satu syarat  
memperoleh gelar sarjana komputer**

Oleh:

**FITRI INDAH NUR AINI**

**NIM. 2016.69.04.0065**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS YUDHARTA PASURUAN**

**2020**

## PERNYATAAN PENULIS

JUDUL : PREDIKSI TRAFIK INTERNET  
HARIAN MENGGUNAKAN  
ALGORITMA *NAIVE BAYES*  
*CLASSIFICATION*

NAMA : FITRI INDAH NUR AINI

NIM : 2016.69.04.0065

“Saya menyatakan dan bertanggung jawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Komputer saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut.”



Fitri Indah Nur Aini  
Penulis



# UNIVERSITAS YUDHARTA PASURUAN FAKULTAS TEKNIK

Kantor Pusat :

Jl. Yudharta No. 07 (Pesantren Ngalah) Sengonagung Purwosari Pasuruan Telp./ Fax. 0343-611186  
e-mail. fakultasteknik@yudharta.ac.id

## SURAT KETERANGAN BEBAS PLAGIASI

Nomor : 0398/S9/FT.UYP/11/09/2020

Yang bertanda tangan dibawah ini:

Nama : Misbach Munir, ST., MT  
NIP.Y : 0690201015  
Jabatan : Dekan Fakultas Teknik

Dengan ini menerangkan bahwa skripsi atas nama mahasiswa :

Nama : Fitri Indah Nur Aini  
NIM : 2016669040065  
Prodi : Teknik Informatika  
Judul Skripsi : Prediksi Traffic Internet Harian Menggunakan Algoritma Naive Bayes Classification  
Hasil Plagiasi : 17%

Demikian surat keterangan ini kami buat untuk digunakan sebagaimana mestinya.

Pasuruan, 14 September 2020  
Dekan Fakultas Teknik



Misbach Munir, ST., MT.  
0690201015

## PERSETUJUAN SKRIPSI

JUDUL : PREDIKSI TRAFIK INTERNET HARIAN  
MENGUNAKAN ALGORITMA *NAIVE*  
*BAYES CLASSIFICATION*

NAMA : FITRI INDAH NUR AINI

NIM : 2016.69.04.0065

Skripsi ini telah diperiksa dan disetujui

Pasuruan, 06 Agustus 2020

Ketua Program Studi



**M. Imron Rosada, M.Kom**

NIP.Y 0690215121

Pembimbing

**M. Faishol Amrulloh, M.Kom**

NIP.Y 0691709007

## PENGESAHAN SKRIPSI

JUDUL : PREDIKSI TRAFIK INTERNET HARIAN  
MENGUNAKAN ALGORITMA *NAIVE  
BAYES CLASSIFICATION*


NAMA : FITRI INDAH NUR AINI

NIM : 2016.69.04.0065

Skripsi ini telah diujikan dan dipertahankan di depan Dewan Penguji pada Sidang Skripsi tanggal 12 Agustus 2020. Menurut pandangan kami, Skripsi ini memadai dari segi kualitas untuk tujuan penganugerahan gelar Sarjana Komputer (S.Kom)

Pasuruan, 24 Agustus 2020

Pembimbing,

  
**M. Faishol Amrulloh, M.Kom**

NIP.Y 0691709007

Penguji Utama,

  
**Arif Faizin, M.Kom**

NIP.Y 0691709008

Kepada

  
**M. Imron Kosadi, M.Kom**

NIP.Y 0690213121

Penguji Anggota

  
**Walidini Syaiful, H.M.Kom**

NIP.Y: 0691709006

Dekan Fakultas Teknik,

  
**Misbach Munir, ST., MT.,**

NIP. Y 0690201015

**Skripsi ini kutujukan kepada  
Ayahanda, Ibunda, adik tercinta,  
dan Suami tersayang**

## **Motto**

**Mulailah dari tempatmu berada  
Gunakan apa yang kau punya  
Lakukan apa yang kau bisa**  
*(Arthur Ashe)*

## **ABSTRACT**

*In this study strong data exploration was carried out on daily internet data traffic generated at the JW.net ISP (Internet Service Provider) for a period of 12 consecutive months for each day for a period of 1 year, uploading and downloading internet data traffic on JW.NET Sengonagung Purwosari is monitored and recorded correctly using software: Hostspot Manager Mikrotik and Winbox.*

*In this proposal research I raise a problem that contrains ISP(Internet Service Provider), that is the difficulty of increasing consumer demand because every day the bandwidth is always changing, so it's a little difficult to determine and estimate the bandwidth to be effective evenly in accordance with consumer needs for the future (the next period). From some of the problems above, so it is proposed to use the Naive Bayes algorithm, which is quite widely applied to calculate the probability value and an evaluation using Confusion Matrix in order to obtain accurate prediction result. Then implemented in the R-studio programing language. From the results of the calculation using the Naive Bayes method which is then evaluated using the Confusion Matrix, the resulting high level of accuracy is 99,49% which 80% trining data and 20% testing data.*

*Keyword : Naive Bayes Algorithm, internet traffic, prediction, ISP (Internet Service Provider)*



# **PREDIKSI TRAFIK INTERNET HARIAN MENGUNAKAN ALGORITMA NAIVE BAYES CLASSIFICATION**

Fitri Indah Nur Aini  
Program Studi Teknik Informatika, Universitas Yudharta  
Pasuruan

## **ABSTRAK**

Dalam penelitian ini eksplorasi data yang kuat dilakukan pada lalu lintas data internet harian yang dihasilkan di ISP (*Internet Service Provider*) JW.net untuk jangka waktu 12 bulan berturut-turut untuk setiap hari selama periode 1 tahun, lalu lintas pengunggahan dan pengunduhan data internet di JW.net Sengonagung Purwosari dimonitor dan catat dengan benar menggunakan perangkat lunak : *Manager Hostspot Mikrotik, winbox*

Dalam proposal ini saya mengangkat sebuah masalah yang menjadi kendala penyedia ISP (*Internet Service Provider*), yaitu sulitnya permintaan konsumen yang semakin banyak karena setiap hari bandwidth itu selalu berubah ubah, jadi sedikit sulit untuk menentukan dan memperkirakan quotanya agar efektif dalam trafik lalu lintas internetnya sesuai dengan kebutuhan konsumen untuk kedepannya (periode selanjutnya), maka dibutuhkan metode yang efektif dan sederhana. Dataset

kuantitatif dengan informasi terperinci disediakan sebagai bahan pelengkap data proposal ini untuk kemudahan penelitian dan validasi, untuk setiap bulan graph statistik dari lalu lintas pengunduhan dan pengunggahan data internet harian disajikan dalam bentuk tabel. Dari beberapa masalah diatas, sehingga diusulkan untuk menggunakan algoritma *Naive Bayes*, yang cukup banyak diterapkan untuk menghitung nilai probabilitas dan dilakukan evaluasi menggunakan Confusion Matrix agar memperoleh hasil prediksi yang akurat. Kemudian diimplementasikan pada bahasa pemrograman *R-studio*. Dari hasil implementasi perhitungan menggunakan model Naive Bayes yang kemudian di evaluasi hasil menggunakan Confusion Matrix dihasilkan tingkat akurasi yang tinggi yaitu 99,49% dengan jumlah data trining 80% dan data testing 20%.

Kata Kunci – algoritma *Naive Bayes*, Trafik Internet ,  
Prediksi, ISP (*Internet Service Provide*)

## **KATA PENGANTAR**

Alhamdulillah segala puji syukur dipanjatkan kehadirat Allah SWT, yang telah melimpahkan nikmat baik berwujud iman dan islam, dan selalu rahmat, taufik, hidayah dan ianayah - Nya sehingga penulis bisa menyelesaikan laporan Proposal skripsi ini yang dilaksanakan di UNIVERSITAS YUDHARTA PASURUAN sebagai salah satu syarat kelulusan dalam program studi S1.

Sholawat serta salam semoga tetap tercurahkan kepada junjungan kita baginda Rasulullah Muhammad SAW, yang menunjukkan jalan kebenaran dan keselamatan, yakni ajaran islam yang menjadi rahmat bagi seluruh umat manusia.

Selama pengerjaan Skripsi ini penulis telah banyak mendapat bimbingan, masukan, motivasi dan arahan dari berbagai pihak. Oleh karena itu, penulis menyampaikan ucapan terima kasih dan penghargaan setinggi – tingginya kepada :

1. KH.Sholeh Bahrudin, selaku Pembina Yayasan Darut Taqwa yang selalu memberikan do'a restunya.
2. Bapak Dr. H. Kholid Murtadlo, SE., MT. selaku Rektor Universitas Yudharta Pasuruan yang telah mengarahkan dan memberikan motivasi kepada penulis.

3. Bapak Misbach Munir, ST., MT., selaku dekan Fakultas Teknik Universitas Yudharta Pasuruan.
4. Wike Adhi Anggono, ST, Selaku Tata Usaha Fakultas Teknik Unifersitas Yudharta Pasuruan.
5. Bapak Muhammad Imron Rosadi, M.Kom. selaku Ketua Program Studi Informatika.
6. Bapak Muhammad Faishol Amrulloh, M.kom., selaku dosen Pembimbing yang telah memberikan banyak arahan kepada penulis.
7. Kedua orang tua saya yang dengan restunya, do'anya, harapan-harapan serta pengorbanannya menjadikan saya untuk tidak menyerah dalam penyelesaian Skripsi ini
8. Teman-teman mahasiswa informatika angkatan 2016 yang telah mendukung serta memberikan support kepada penulis.
9. Sahabat – sahabat saya yang telah mensupport dan mendo'akan untuk kesuksesan pengerjaan skripsi saya
10. ISP JW.NET yang mau menjadi mitra penelitian skripsi saya
11. Semua pihak yang terlibat baik secara langsung maupun tidak langsung pada proses penulisan Proposal Skripsi ini.

Dengan segala kerendahan hati penulis menyadari bahwa Skripsi ini jauh dari kata sempurna, sehingga kritik dan saran sangat diperlukan demi tercapainya hasil yang lebih baik. Harapan penulis semoga Skripsi ini dapat bermanfaat bagi penulis dan bagi pembaca. Amin.

Pasuruan , Agustus 2020

**Penulis**

## DAFTAR ISI

PERNYATAAN PENULIS .....	ii
PERSETUJUAN SKRIPSI .....	iii
PENGESAHAN SKRIPSI .....	iv
Motto .....	v
<i>ABSTRACT</i> .....	vii
ABSTRAK .....	viii
KATA PENGANTAR .....	x
DAFTAR ISI .....	xiii
DAFTAR TABEL .....	xvi
BAB I PENDAHULUAN .....	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian .....	3
1.4 Manfaat Penelitian .....	4
1.5 Batasan Masalah .....	4
1.6 Sistematika Penulisan .....	5
BAB II TINJAUAN PUSTAKA	
2.1 Penelitian Terkait .....	7
2.2 Data Mining .....	18
2.3 Algoritma .....	19

2.4 Pengertian Naive Bayes .....	19
2.4.1 Persamaan Metode Naive Bayes.....	21
2.4.2 Pengukuran Kinerja Klasifikasi.....	27
2.5 Bandwidth .....	28
2.6 Mikrotik .....	29
2.7 Prediksi .....	30
2.8 Evaluasi RMSE .....	31
2.9 Traffic .....	31
2.10 Internet .....	32
2.11 Jaringan .....	32
2.12 Klasifikasi .....	32
2.13 Bahasa Pemrograman Python .....	33
2.14 Confusion Matrix .....	34
2.15 R-Studio .....	35

### BAB III METODE PENELITIAN

3.1 Kerangka Pemikiran .....	37
3.2 Analisa Kebutuhan .....	38
3.3 Metode Penelitian .....	42
3.4 Metode Pemilihan Data .....	43
3.5 Metode Pengumpulan Data .....	46
3.6 Metode Klasifikasi Data .....	46

3.7 Transformasi Data .....	47
3.8 Pengukuran Kinerja Model Klasifikasi .....	47
3.9 Perancangan Penelitian .....	47
3.10 Model/metode yang diusulkan .....	47
<b>BAB IV HASIL DAN PEMBAHASAN</b>	
4.1 Kebutuhan Perangkat Sistem .....	51
4.2 Kebutuhan Perangkat Keras ( <i>hardware</i> ).....	51
4.3 Kebutuhan Perangkat Lunak .....	52
4.4 Perhitungan Manual .....	52
4.4.1 Proses Perhitungan Manual .....	52
<b>BAB V KESIMPULAN DAN SARAN</b>	
5.1 KESIMPULAN .....	67
5.2 SARAN .....	67
DAFTAR PUSTAKA .....	68
LAMPIRAN .....	73



## DAFTAR ATABEL

Tabel 2.1 Penelitian terdahulu .....	12
Tabel 2.2 Confusion matrix .....	34
Tabel 3.1 Dataset spesifikasi dan atribut .....	43
Tabel 4.1 Dataset .....	53
Tabel 4.2 Atribut data sampel .....	54
Tabel 4.3 Data testing .....	55
Tabel 4.4 Perhitungan mean dan stdev atribut .....	55
Tabel 4.5 Data testing .....	56
Tabel 4.6 Data testing .....	57
Tabel 4.7 Data testing .....	58
Tabel 4.8 Data testing .....	59
Tabel 4.9 Data testing .....	60
Tabel 4.10 Data testing .....	61
Tabel 4.11 Data testing .....	62
Tabel 4.12 Confusion matrix pada naive bayes .....	63
Tabel 5.1 Hasil akurasi .....	67

## DAFTAR GAMBAR

Gambar 2.1 Tahapan data mining .....	18
Gambar 2.2 Alur metode naive bayes .....	24
Gambar 2.3 Mikrotik RB 750 .....	29
Gambar 2.4 Simbol Phyton .....	33
Gambar 2.5 R-Studio dekstop .....	36
Gambar 3.1 Kerangka pemikiran .....	37
Gambar 3.2 Tahapan penelitian .....	39
Gambar 3.3 Model yang diusulkan .....	48
Gambar 3.4 Teknik naive bayes .....	50
Gambar 4.2 Hasil perhitungan confusion matrix pada R-studio .....	66