

ARTIKEL

KLASIFIKASI KEPOPULERAN ARTIKEL *WEBSITE* SORETOTTE.NET DENGAN MENGGUNAKAN METODE *NAIVE BAYES*



Oleh:

BAYU WARDANI

13.1.03.02.0074

Dibimbing oleh :

1. M. Rizal Arief, ST., M.Kom
2. Intan Nur Farida, M.Kom

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS NUSANTARA PGRI KEDIRI
2017**

**SURAT PERNYATAAN**
ARTIKEL SKRIPSI TAHUN 2017

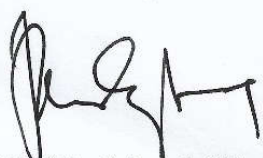


Yang bertanda tangan di bawah ini:

Nama Lengkap : Bayu Wardani
NPM : 13.1.03.02.0074
Telepon/HP : 082231557217
Alamat Surel (Email) : bayuwardani51@gmail.com
Judul Artikel : Klasifikasi Kepopuleran Artikel *Website* Soretotte.net
Dengan Menggunakan Metode *Naive Bayes*
Fakultas – Program Studi : Teknik Informatika
Nama Perguruan Tinggi : Universitas Nusantara PGRI Kediri
Alamat Perguruan Tinggi : Jl. KH. Ahmad Dahlan No.76, Mojoroto, Kota Kediri,
Jawa Timur 64112

Dengan ini menyatakan bahwa :

- artikel yang saya tulis merupakan karya saya pribadi (bersama tim penulis) dan bebas plagiarisme;
- artikel telah diteliti dan disetujui untuk diterbitkan oleh Dosen Pembimbing I dan II.

Demikian surat pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari ditemukan ketidaksesuaian data dengan pernyataan ini dan atau ada tuntutan dari pihak lain, saya bersedia bertanggungjawab dan diproses sesuai dengan ketentuan yang berlaku.

Mengetahui		Kediri, 09 Agustus 2017
Pembimbing I  <u>M. Rizal Arlef, ST., M.Kom</u> NIDN. 0716027504	Pembimbing II  <u>Intan Nur Farida, M.Kom</u> NIDN. 0704108701	Penulis,  <u>Bayu Wardani</u> 13.1.03.02.0074

KLASIFIKASI KEPOPULERAN ARTIKEL *WEBSITE* SORETOTTE.NET DENGAN MENGGUNAKAN METODE *NAIVE BAYES*

Bayu Wardani

13.1.03.02.0074

Fakultas Teknik – Teknik Informatika

Bayuwardani51@gmail.com

M. Rizal Arief, ST., M.Kom dan Intan Nur Farida, M.Kom

UNIVERSITAS NUSANTARA PGRI KEDIRI

ABSTRAK

Penelitian ini berdasarkan hasil pengamatan dan pengalaman peneliti, bahwa penilaian artikel populer sebuah *website* hanya di dasari dari jumlah *view* setiap artikel, jika hanya melihat dari jumlah *view* kurang tepat bila artikel tersebut disebut artikel populer, oleh sebab itu perlu ditambahkan sebuah parameter lain seperti, jumlah *share* Facebook, letak halaman Google, jumlah komentar, jumlah kata, serta lama *publish*.

Penelitian ini menggunakan metode kuantitatif dengan subyek penelitian adalah *website* *soretotte.net*, menggunakan pemrograman berbasis *web* dengan PHP dan HTML serta CSS.

Kesimpulan dari penelitian ini adalah (1) Berdasarkan hasil pengujian dengan 5 data artikel, dapat disimpulkan bahwa sistem ini dapat membantu dan mempermudah pemilik *website* dalam memperoleh informasi mengenai kualitas artikel. (2) Dengan adanya satu kesalahan dari 5 data testing yang dilakukan sistem, maka nilai keefektifan sistem klasifikasi kepopuleran artikel dengan menggunakan metode *Naive Bayes* ini sebesar 75%. (3) Aplikasi ini juga dapat digunakan sebagai alat bantu pembuatan artikel yang disukai pembaca dengan cara melihat artikel yang memiliki keputusan populer.

KATA KUNCI : Artikel, Klasifikasi, *Naive Bayes*, Data Mining

I. LATAR BELAKANG

Sekarang ini, sebuah *website* merupakan sesuatu hal yang lumrah banyak media yang tadinya dimuat secara offline seperti berita, buku dan lain-lain sekarang dimuat secara online dalam bentuk *website* untuk mempermudah akses para pembaca, contohnya saja Jawa Pos yang dulunya diterbitkan secara offline sekarang sudah berubah menjadi online meskipun terbitan offline masih dijalankan. Sebenarnya bukan Jawa Pos saja yang melakukan hal tersebut

ada banyak juga media lain yang melakukan hal yang sama, dalam beberapa tahun ini saja banyak sebuah *website* dengan berbagai niche sendiri lahir yang membuat persaingan dalam menarik seorang pembaca serta pembuatan artikel menjadi lebih sulit.

Artikel sendiri merupakan jantung dari suatu *website* khususnya *website* yang berisi berita, tutorial, maupun hiburan. Semakin banyak artikel yang populer maka akan semakin baik juga kualitas *website*

tersebut serta meningkatkan kepuasan pembaca.

Namun kebanyakan website mengukur kepopuleran artikel mereka hanya dengan melihat jumlah view seperti halnya website yang akan menjadi tempat penelitian, website tersebut adalah soretotte.net yang merupakan sebuah website berisikan banyak sekali artikel tentang game, anime, dan hal menarik lainnya. Akan tetapi apabila hanya dengan melihat jumlah view kurang tepat bila disebut artikel populer, hal ini terjadi karena tidak adanya sebuah aplikasi yang dapat membantu menilai sebuah artikel dengan penambahan parameter lain serta para pemilik website juga tidak mengetahui parameter lain yang dapat meningkatkan penilaian dalam menentukan artikel populer, untuk menyelesaikan masalah ini diperlukan sebuah parameter yang juga sudah sering digunakan dalam penilaian artikel populer atau tidak, parameter tersebut adalah jumlah view artikel, jumlah share Facebook, letak halaman Google, jumlah komentar, jumlah keseluruhan kata, dan juga lamanya artikel tersebut dipublish, serta diperlukan juga aplikasi pengklasifikasian kepopuleran artikel yang dapat membantu pemilik website maupun penulis website agar mudah mengetahui apa yang mereka tulis populer atau tidak.

Dalam hal ini ada sebuah metode yang sangat cocok digunakan untuk mengklasifikasikan kepopuleran sebuah artikel yaitu metode Naive Bayes yang memang telah sering digunakan untuk memecahkan permasalahan yang berhubungan dengan proses klasifikasi.

Diharapkan dengan adanya penelitian ini akan dapat membantu para pemilik website untuk menentukan artikel mereka populer atau tidak, serta dapat mempermudah dalam pembuatan artikel.

II. METODE

A. *Naive Bayes*

Bayes merupakan teknik prediksi probabilistik sederhana yang berdasar pada penerapan teorema Bayes (atau aturan Bayes) dengan asumsi independensi atau ketidaktergantungan yang kuat (naif).

Prediksi Naive Bayes didasarkan pada teorema bayes dengan formula umum sebagai berikut (Prasetyo, 2012) :

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)}$$

Keterangan :

$P(H|E)$ = Probabilitas akhir bersyarat (*conditional probability*) suatu hipotesis H terjadi jika diberikan (*evidence*) E terjadi.

$P(E|H)$ = Probabilitas sebuah E terjadi akan memengaruhi hipotesis H.

$P(H)$ = Probabilitas awal (*priori*) hipotesis H terjadi tanpa memandang hipotesis apapun.

$P(E)$ = Probabilitas awal (*priori*) bukti E terjadi tanpa memandang bukti yang lain.

B. Simulasi Perhitungan

Sebelum melakukan proses pengklasifikasian dengan menggunakan metode naive bayes. Tentukan parameter yang digunakan seperti pada tabel 1.

Tabel 1 : Parameter Naive Bayes

Parameter (satuan)	Jenis Spesifikasi
1. Jumlah <i>view</i> (<i>page view</i>)	1. 0<50 5.201<250 2. 51<100 6.251<300 3.101<150 7.301<350 4.151<200 8.351<400 9. 401<450 10. 451<500 11.500++
2. Jumlah <i>Share</i> Facebook (<i>share</i>)	1. 1<3 3. 8<1 8. 23<26 2. 4<7 4. 12< 9. 26++
3. Halaman Google (<i>google</i>)	1. 1 3. 3 5. 5 2. 2 4. 4 6. 6++
4..Jumlah Komentar (<i>komentar</i>)	1. 0 3. 2<3 3.6<7 2. 1 2.4<5 4. 7 ++
5.Jumlah Keseluruhan Kata (<i>kata</i>)	1. 0<100 3. 151<200 2. 101<150 4. 201<250 5. 300++
6. Lama <i>Publish</i> (<i>hari</i>)	1. 1 3. 3 5. 5 7. 8+ 2. 2 4. 4 6. 7

Dalam perhitungan *naive bayes* data training digunakan sebagai bahan acuan untuk dilakukan analisis sehingga mendapatkan keputusan yang terbaik. Sebagai contoh dalam penelitian ini membutuhkan 10 data training sebagai berikut :

Tabel 1 : Data Training

No	Kriteria Kepopuleran Artikel						Keputusan
	A	B	C	D	E	F	
1	51<100	1<3	3	101<150	0	3	Tidak
2	101<150	4<7	2	151<200	0	4	Populer
3	51<100	1<3	1	101<150	1	1	Populer
4	201<250	1<3	2	300+	2<3	7	Populer
5	500++	8<11	1	300+	6+	7	Populer
6	51<100	1<3	5	300+	0	3	Tidak
7	201>250	4<7	1	151<200	2<3	5	Populer
8	51<100	4<7	1	300+	0	6	Tidak
9	101<150	1<3	2	101<150	2<3	7	Tidak
10	401<450	4<7	2	300+	4<5	3	Populer

Jika diketahui sebuah artikel memiliki spesifikasi sebagai berikut :

1. Jumlah *view* sebanyak 401<450 *page view*.
2. Berada dihalaman pertama Google.
3. Memiliki jumlah komentar 2<3 komentar.
4. Memiliki Jumlah kata sebanyak 300+ kata.
5. Jumlah *Share* Facebook sebanyak 1<3 kali share.
6. Dan waktu lama *publish* selama 3 hari.

Langkah selanjutnya adalah menghitung masing-masing probabilitas parameter seperti berikut :

$$P(\text{Keputusan} \mid \text{Populer}) = \frac{6}{10}$$

$$P(\text{Keputusan} \mid \text{Tidak}) = \frac{4}{10}$$

$$P(A = \text{"view"} \mid \text{Populer}) = \frac{1}{6}$$

$$P(A = \text{"view"} \mid \text{Tidak}) = \frac{0}{4}$$

$$P(B = \text{"sharefb"} \mid \text{Populer}) = \frac{2}{6}$$

$$P(B = \text{"sharefb"} \mid \text{Tidak}) = \frac{3}{4}$$

$$P(C = \text{"Google"} \mid \text{Populer}) = \frac{3}{6}$$

$$P(C = \text{"Google"} \mid \text{Tidak}) = \frac{1}{4}$$

$$P(E = \text{"Komentar"} \mid \text{Populer}) = \frac{2}{6}$$

$$P(E = \text{"Komentar"} \mid \text{Tidak}) = \frac{1}{4}$$

$$P(D = \text{"Jumlah Kata"} \mid \text{Populer}) = \frac{3}{6}$$

$$P(D = \text{"Jumlah Kata"} \mid \text{Tidak}) = \frac{2}{4}$$

$$P(F = \text{"Lama Publish"} \mid \text{Populer}) = \frac{1}{6}$$

$$P(F = \text{"Lama Publish"} \mid \text{Tidak}) = \frac{2}{4}$$

Keterangan :

1. A = Jumlah *View*.
2. B = Jumlah *Share* Facebook.
3. C = Letak Halaman Google.
4. E = Jumlah Komentar.
5. D = Jumlah Kata.
6. F = Lama *Publish*.

Langkah selanjutnya setiap variabel yang memiliki *Class* kualitas yang sama dikalikan seperti pada tabel dibawah :

Parameter						Class	Nilai
A	B	C	D	E	F		
$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{1}{6}$	$\frac{6}{10}$	2,58
$\frac{0}{4}$	$\frac{3}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{2}{4}$	$\frac{4}{10}$	2,25

Dengan Hasil akhir keputusan Populer bernilai 2,58 sedangkan Keputusan tidak Populer bernilai 2,25 maka hasil dari artikel yang dites memiliki keputusan populer karena nilai populer lebih besar dari pada nilai tidak populer.

III. HASIL DAN KESIMPULAN

Berdasarkan permasalahan yang telah dibahas dan diselesaikan maka terdapat beberapa kesimpulan yaitu :

1. Berdasarkan hasil pengujian dengan 5 data artikel, dapat disimpulkan bahwa sistem ini dapat membantu dan mempermudah pemilik website dalam memperoleh informasi mengenai kualitas artikel.
2. Dengan adanya satu kesalahan dari 5 data testing yang dilakukan sistem, maka nilai ke efektifan sistem klasifikasi kepopuleran artikel dengan menggunakan metode Naive Bayes ini sebesar 75%.
3. Aplikasi ini juga dapat digunakan sebagai alat bantu pembuatan artikel yang disukai pembaca dengan cara melihat artikel yang memiliki keputusan populer.

IV. DAFTAR PUSTAKA

Ariandi, Dio. Fithruasari, Kartika. 2015. Kl-
asifikasi Berita Indonesia
Menggunakan Metode Naive Bayes
Classifier dan Support Vector
Machine dengan confix Stripping
Stemmer. Jurnal Sains dan Seni ITS,
Volume 4.

Erlinasari, Y. 2014. *Aplikasi Pengklasifikasi Love Bird Menggunakan Metode Naive Bayes*. Kediri: Teknik Informatika UN PGRI Kediri.

Nugraha, P. A., Saptono, R., & Sulisty, M. E. 2013. Perbandingan Metode Probabilistik Naive Bayesian Classifier dan Jaringan Syaraf Tiruan Learning Vector Quantization dalam Kasus Klasifikasi Penyakit Kandungan. *Jurnal ITSMART*.

Prasetyo, E. 2012. *Data Mining - Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta: Andi.

Saleh, A. (2015). Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. *Citec Journal*, Volume 2.

Sibero, Alexander. 2013. Web Programing Power Pack Buku Web Programing Terlengkap. Penerbit MediaKom : Yogyakarta.

Sugiyono, P. D. 2010. *Metode Penelitian Kuantitatif, Kualitatif Dan R&D*. Bandung: Alfabeta.