

**PENGLASIFIKASIAN ARTIKEL ILMIAH MENGGUNAKAN
METODE *NAIVE BAYES CLASSIFIER* (NBC)**

SKRIPSI

**Diajukan untuk Memenuhi Salah Satu Syarat Mencapai Gelar Strata Satu
Program Studi Informatika**



DISUSUN OLEH:

DWI PRAMITA BAGASSANTY BESTARI

NIM. M0512014

**PROGRAM STUDI INFORMATIKA
FAKULTAS MATEMATIKA & ILMU PENGETAHUAN ALAM
UNIVERSITAS SEBELAS MARET
SURAKARTA
2018**

commit to user

SKRIPSI

**PENGLASIFIKASIAN ARTIKEL ILMIAH MENGGUNAKAN
METODE *NAIVE BAYES CLASSIFIER* (NBC)**

Disusun oleh:

DWI PRAMITA BAGASSANTY BESTARI

NIM M0512014

telah disetujui untuk dipertahankan di hadapan Dewan Penguji
pada tanggal:

Pembimbing 1,

Pembimbing 2,



Ristu Saptono, S.Si., M.T.

NIP. 197902102002121001



Rini Anggrainingsih, S.T., M.T.

NIP. 197809092008122002

SKRIPSI

PENGLASIFIKASIAN ARTIKEL ILMIAH MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER* (NBC)

Disusun oleh:

DWI PRAMITA BAGASSANTY BESTARI

NIM M0512014

telah dipertahankan di hadapan Dewan Penguji
pada tanggal:

Susunan Dewan Penguji

1. **Ristu Saptono, S.Si., M.T.** (Ketua)
NIP. 197902102002121001
2. **Rini Anggrainingsih, S.T., M.T.** (Sekretaris)
NIP. 197809092008122002
3. **Afrizal Doewes, S.Kom., M.Sc.** (Anggota)
NIP. 198508312012121004
4. **Heri Prasetyo S.Kom, M.Sc.Eng., Ph** (Anggota)
NIP. 1983030220161001

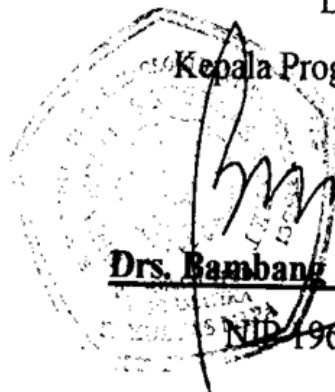


Disahkan oleh

Kepala Program Studi Informatika,

Drs. Bambang Harjito, M.App.Sc., Ph.D.

NIP. 196211301991031002



PENGLASIFIKASIAN ARTIKEL ILMIAH MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER (NBC)*

DWI PRAMITA B. B

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Sebelas Maret

ABSTRAK

Universitas Sebelas Maret telah banyak menerbitkan artikel ilmiah. Mengklasifikasikan banyak artikel sekaligus bukanlah hal yang mudah dilakukan. Semakin banyak artikel yang harus dikelompokkan maka tenaga dan waktu yang dibutuhkan juga semakin banyak. Metode *Naive Bayes Classifier* dapat digunakan untuk mengklasifikasikan artikel ilmiah dalam waktu singkat. *Naive Bayes Classifier* mengklasifikasikan masing-masing artikel berdasarkan bidang ilmu dengan menganalisis judul dan abstraknya. Salah satu metode seleksi fitur *Document Frequency Improved* diterapkan untuk meningkatkan kinerja klasifikasi. Penelitian menggunakan 292 artikel sebagai data latih dan 100 artikel sebagai data uji. Pengujian dilakukan dengan menerapkan 5 nilai *threshold* berbeda dari 1 hingga 2,5 dengan masing-masing nilai *threshold* diujikan sebanyak 5 kali. Hasil terbaik ditunjukkan pada nilai *threshold* 2 dengan nilai rata-rata akurasi, *precision*, *recall*, dan *f-measure* masing-masing sebesar 87,8%, 76,6%, 76,2%, dan 76,0%.

Kata kunci: klasifikasi, *naive bayes classifier*, *document frequency improved*

**ACADEMIC ARTICLES CLASSIFICATION USING
NAIVE BAYES CLASSIFIER (NBC) METHOD**

DWI PRAMITA B. B

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Sebelas Maret

ABSTRACT

Sebelas Maret University has been publishing many academic articles. Classifying many articles at a time is not a simple task. The more articles need to be classified, the more energy and time needed. Naive Bayes Classifier method can be used to classify academic articles in short time. Naive Bayes Classifier classifies each article based on the field of study by analyzing its title and abstract. One of feature selection method Document Frequency Improved (DFM) is implemented for improving the classification performance. This study used of 292 articles as training data and 100 articles as testing data. It tested by applying 5 threshold value from 1 to 2,5 with each threshold executed 5 times. The best results showed at 2 threshold level with the average value of accuracy, precision, recall, and f-measure respectively are 87,8%, 76,6%, 76,2%, and 76,0%.

Keywords: *classification, naive bayes classifier, document frequency improved*

commit to user

KATA PENGANTAR

Assalamu'alaikum Wr. Wb.

Segala puji penulis panjatkan kehadirat Allah SWT atas limpahan nikmat, hidayah dan inayah-Nya sehingga penulis dapat menyelesaikan laporan skripsi yang berjudul “Pengklasifikasian Artikel Ilmiah Menggunakan Metode *Naive Bayes Classifier (NBC)*”.

Penulis menyadari bahwa laporan skripsi ini masih jauh dari kesempurnaan, baik dari segi penulisan maupun materi. Penulis mengucapkan terima kasih kepada semua pihak yang telah meluangkan waktu untuk memberikan bimbingan dan saran sehingga laporan ini dapat berwujud sebagaimana yang diharapkan, terutama kepada :

1. Kedua orang tua dan saudara-saudara saya yang telah memberikan dukungan sehingga penulis dapat menyelesaikan laporan ini.
2. Bapak Ristu Saptono, S.Si.,M.T. dan Ibu Rini Anggrainingsih, S.T.,M.T., selaku Dosen Pembimbing, yang telah memberikan pengarahan selama proses penyusunan laporan ini.
3. Bapak Drs. Bambang Harjito, M.App.Sc., Ph.D., selaku Kepala Program Studi Informatika FMIPA UNS.
4. Sahabat-sahabat saya: Rofiqoh Hasanah, Risalatul Hanifah, Amelia Rahman, Ely Desyanawati, yang telah memberikan dukungan serta bantuan dalam menyelesaikan laporan ini.
5. Teman-teman seperjuangan di Program Studi Informatika angkatan 2012 yang telah berbagi keceriaan dalam pengerjaan laporan ini.

Semoga Allah SWT membalas jerih payah dan pengorbanan yang telah diberikan dengan balasan yang lebih baik. Semoga laporan skripsi ini bermanfaat dan dapat memberikan sumbangsih bagi pihak-pihak yang berkepentingan.

Wassalamu'alaikum Wr.Wb.

commit to user

Surakarta, Juli 2018

Penulis

DAFTAR ISI

| | |
|--------------------------------------|------|
| HALAMAN JUDUL | i |
| HALAMAN PERSETUJUAN | ii |
| HALAMAN PENGESAHAN | iii |
| ABSTRAK..... | iv |
| <i>ABSTRACT</i> | v |
| <i>MOTTO</i> | vi |
| PERSEMBAHAN..... | vii |
| KATA PENGANTAR | viii |
| DAFTAR ISI..... | ix |
| DAFTAR TABEL..... | xi |
| DAFTAR GAMBAR..... | xii |
| DAFTAR LAMPIRAN..... | xiii |
| BAB 1 PENDAHULUAN..... | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah | 3 |
| 1.3 Batasan Masalah..... | 3 |
| 1.4 Tujuan Penelitian | 4 |
| 1.5 Manfaat Penelitian | 4 |
| 1.6 Sistematika Penelitian | 4 |
| BAB 2 TINJAUAN PUSTAKA | 5 |
| 2.1 Landasan Teori..... | 5 |
| 2.1.1 <i>Text Mining</i> | 5 |
| 2.1.2 <i>Feature Selection</i> | 8 |

| | | |
|----------------|---|----|
| 2.1.3 | <i>Naive Bayes Classifier (NBC)</i> | 11 |
| 2.2 | Penelitian Terkait | 13 |
| BAB 3 | METODOLOGI..... | 15 |
| 3.1 | Pengumpulan Data | 15 |
| 3.2 | <i>Text Preprocessing</i> | 16 |
| 3.3 | <i>Feature Selection</i> | 19 |
| 3.4 | <i>Text Classification</i> | 19 |
| 3.5 | Analisis Hasil Klasifikasi | 20 |
| BAB 4 | HASIL DAN PEMBAHASAN | 23 |
| 4.1 | Pengumpulan Data | 23 |
| 4.2 | <i>Text Preprocessing</i> | 26 |
| 4.3 | <i>Feature Selection</i> | 27 |
| 4.4 | <i>Text Classification</i> | 28 |
| 4.5 | Analisis Hasil Klasifikasi | 29 |
| BAB V | KESIMPULAN DAN SARAN | 33 |
| 5.1 | Kesimpulan | 33 |
| 5.2 | Saran..... | 33 |
| DAFTAR PUSTAKA | | 34 |