

**KLASIFIKASI JUDUL BERITA BAHASA INDONESIA
MENGUNAKAN METODE SELEKSI FITUR CHI-SQUARE DAN
ALGORITMA SUPPORT VECTOR MACHINE (SVM)**

Diajukan Sebagai Syarat
Untuk Menyelesaikan Pendidikan Program Strata-1
Pada Jurusan Teknik Informatika



Oleh :
Nazifah Suci Aqillah
NIM : 09021382025129

**Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2024**

LEMBAR PENGESAHAN SKRIPSI

KLASIFIKASI JUDUL BERITA BAHASA INDONESIA
MENGUNAKAN METODE SELEKSI FITUR CHI-SQUARE
DAN ALGORITMA SUPPORT VECTOR MACHINE (SVM)

Oleh :

NAZIFAH SUCI AQILLAH
NIM : 09021382025129

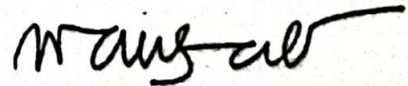
Palembang, 13 Mei 2024

Pembimbing I



Alvi Syahrini Utami, M.Kom
NIP. 197812222006042003

Pembimbing II



M. Naufal Rachmatullah, M.T
NIP. 19921201202231008

Mengetahui,
Ketua Jurusan Teknik Informatika




Dr. M. Fachrurrozi, S.Si., M.T.
NIP. 198005222008121002

TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari Jum'at tanggal 26 April 2024 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : NAZIFAH SUCI AQILLAH

NIM : 09021382025129

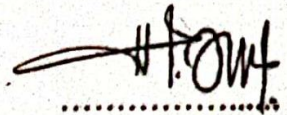
Judul : KLASIFIKASI JUDUL BERITA BAHASA INDONESIA
METODE SELEKSI FITUR CHI-SQUARE DAN ALGORITMA SUPPORT
VECTOR MACHINE (SVM).

dan dinyatakan LULUS.

1. Ketua

Annisa Darmawahyuni, M.Kom.

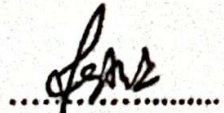
NIP. 199006302023212044



2. Penguji I

Desty Rodiah, M.T.

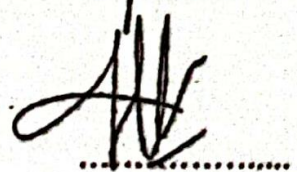
NIP. 198912212020122011



3. Pembimbing I

Alvi Syahrini Utami, M.Kom.

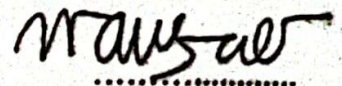
NIP. 197812222006042003



4. Pembimbing II

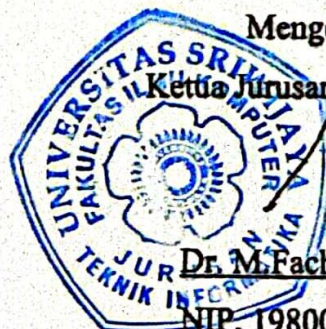
M.Naufal Rachmatullah, M.T.

NIP. 19921201202231008



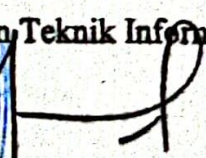
Mengetahui,

Ketua Jurusan Teknik Informatika



Dr. M. Fachrurrozi, S.Si., M.T.

NIP. 198005222008121002



HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Nazifah Suci Aqillah

NIM : 09021382025129

Program Studi : Teknik Informatika Bilingual

Judul : Klasifikasi Judul Berita Bahasa Indonesia Menggunakan Metode Seleksi Fitur Chi-Square dan Algoritma Support Vector Machine

Hasil Pengecekan Software iThenticate/Turnitin : 17%

Menyatakan bahwa Laporan Penelitian saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan penelitian ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, 13 Mei 2024



METERAI
TEMPEL

902C1AKX814021890

Nazifah Suci Aqillah
09021382025129

MOTTO DAN PERSEMBAHAN

“Berilmulah sebelum kamu berbicara, beramal, dan beraktivitas” - HR.Bukhari

Kupersembahkan karya tulis ini kepada :

- Orang Tua, Saudara/i, dan Keluargaku
- Dosen Pembimbing
- Teman-teman seperjuangan
- Fakultas Ilmu Komputer
- Universitas Sriwijaya

ABSTRACT

News is an important source of information in society, and classification of news headlines is a challenge for organizing this information. This research aims to develop software that is able to classify Indonesian news headlines into 3 categories using the Support Vector Machine method with Chi-Square Feature Selection. The SVM method is known as an effective classification algorithm, while chi-square feature selection is used to select important words in news headlines that differentiate between categories. These selected words are then represented using TF-IDF weighting and used as features to train the SVM model. The data used is a collection of Indonesian language news titles in the categories EDU, FINANCE, SPORT. The research stages include text pre-processing, feature extraction, TFIDF weighting, SVM model training, and classification performance evaluation. The results of classification research using the value $C=10$ show that by applying a combination of chi-square feature selection with the SVM algorithm, the level of classification accuracy of Indonesian news titles decreases by around 3% compared to without feature selection. This research shows that chi-square feature selection with a linear kernel combined with the SVM algorithm is less effective for classification results.

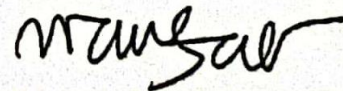
Keywords: Support Vector Machine, Chi-Square, classification accuracy, News Title

Supervisor I,

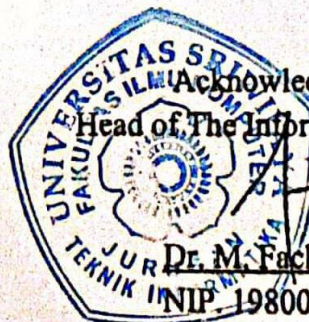


Alvi Syahrini Utami, M.Kom
NIP. 197812222006042003

Supervisor II,



M. Naufal Rachmatullah, M.T
NIP. 19921201202231008



Acknowledged by,
Head of The Informatics Engineering Department

Dr. M. Fachrurrozi, S.Si., M.T.
NIP. 198005222008121002

ABSTRAK

Berita merupakan sumber informasi yang penting dalam masyarakat, dan klasifikasi judul berita menjadi suatu tantangan untuk mengorganisir informasi tersebut. Penelitian ini bertujuan untuk mengembangkan perangkat lunak yang mampu mengklasifikasikan judul berita bahasa Indonesia ke dalam 3 kategori menggunakan metode Support Vector Machine dengan Seleksi Fitur Chi-Square. Metode SVM dikenal sebagai algoritma klasifikasi yang efektif sedangkan seleksi fitur chi-square digunakan untuk memilih kata-kata penting dalam judul berita yang membedakan antar kategori. Kata-kata terpilih ini kemudian direpresentasikan menggunakan pembobotan TF-IDF dan digunakan sebagai fitur untuk melatih model SVM. Data yang digunakan adalah kumpulan judul berita bahasa Indonesia dengan kategori EDU, FINANCE, SPORT. Tahapan penelitian meliputi pra pemrosesan teks, ekstraksi fitur, pembobotan TF-IDF, pelatihan model SVM, dan evaluasi performa klasifikasi. Hasil penelitian klasifikasi menggunakan nilai $C=10$ menunjukkan bahwa dengan menerapkan kombinasi seleksi fitur chi-square dengan algoritma SVM justru mengalami penurunan tingkat akurasi klasifikasi judul berita bahasa Indonesia sekitar 3% dibandingkan tanpa seleksi fitur. Penelitian ini menunjukkan bahwa kurang efektifnya seleksi fitur chi-square dengan kernel linear yang dikombinasikan dengan algoritma SVM untuk hasil klasifikasi.

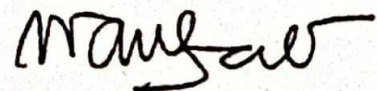
Kata kunci : *Support Vector Machine, Chi-Square, akurasi klasifikasi, Judul Berita*

Pembimbing I,



Alvi Syahini Utami, M.Kom
NIP. 197812222006042003

Pembimbing II,



M. Naufal Rachmatullah, M.T
NIP. 19921201202231008

Mengetahui,
Ketua Jurusan Teknik Informatika

Dr. M. Fachrurrozi, S.Si., M.T.
NIP. 198005222008121002



DAFTAR ISI

LEMBAR PENGESAHAN	ii
TANDA LULUS UJIAN KOMPREHENSIF.....	iii
HALAMAN PERNYATAAN.....	iv
MOTTO DAN PERSEMBAHAN.....	v
ABSTRACT.....	vi
ABASTRAK.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xiii
BAB I PENDAHULUAN	I-1
1.1 Pendahuluan.....	I-1
1.2 Latar Belakang Masalah.....	I-1
1.3 Rumusan Masalah.....	I-3
1.4 Tujuan Penelitian.....	I-3
1.5 Manfaat Penelitian.....	I-4
1.6 Batasan Masalah.....	I-4
1.7 Sistematika Penelitian.....	I-4
1.8 Kesimpulan.....	I-6
BAB II KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan.....	II-1
2.2 Landasan Teori.....	II-1
2.2.1 Text Classification.....	II-1
2.2.2 Preprocessing.....	II-1
2.2.3 TF-IDF.....	II-2
2.2.4 Support Vector Machine.....	II-3
2.2.5 Chi-Square Feature Selection.....	II-8
2.2.6 Evaluation.....	II-9
2.2.7 Rational United Process.....	II-10
2.3 Penelitian Terdahulu.....	II-12
2.4 Kesimpulan.....	II-14
BAB III METODOLOGI PENELITIAN.....	III-1
3.1 Pendahuluan.....	III-1
3.2 Pengumpulan Data.....	III-1
3.2.1 Jenis dan Sumber Data.....	III-1
3.3 Tahapan Penelitian.....	III-3
3.3.1 Pembagian Data Uji dan Data Latih.....	III-3
3.3.2 Menentukan Kerangka Kerja Penelitian.....	III-4
3.3.3 Menentukan Alat Bantu Penelitian.....	III-8

3.3.4 Menentukan Kriteria Penelitian.....	III-9
3.3.5 Melakukan Uji Penelitian.....	III-9
3.3.6 Melakukan Analisis Hasil Pengujian dan Kesimpulan Penelitian....	III-10
3.4 Metode Pengembangan Perangkat Lunak.....	III-10
3.4.1 Fase Insepsi.....	III-10
3.4.2 Fase Elaborasi.....	III-10
3.4.3 Fase Konstruksi.....	III-11
3.4.4 Fase Transisi.....	III-11
3.5 Kesimpulan.....	III-11
BAB IV PENGEMBANGAN PERANGKAT LUNAK.....	IV-1
4.1 Pendahuluan.....	IV-1
4.2 Fase Insepsi.....	IV-1
4.2.1 Pemodelan Bisnis.....	IV-1
4.2.2 Kebutuhan Fungsional dan Non-Fungsional.....	IV-2
4.2.3 Analisis dan Desain.....	IV-3
4.2.3.1 Analisis Kebutuhan Perangkat Lunak.....	IV-3
4.2.3.2 Analisis Data.....	IV-4
4.2.3.3 Analisis Text Preprocessing.....	IV-4
4.2.3.4 Analisis Proses Klasifikasi.....	IV-9
4.2.3.5 Analisis Hasil Klasifikasi.....	IV-24
4.2.4 Implementasi.....	IV-25
4.3 Fase Elaborasi.....	IV-32
4.3.1 Pemodelan Bisnis.....	IV-32
4.3.2 Perancangan Data.....	IV-32
4.3.3 Perancangan Antarmuka.....	IV-32
4.3.4 Kebutuhan.....	IV-36
4.3.5 Diagram Aktivitas.....	IV-37
4.3.6 Diagram Sequence.....	IV-39
4.4 Fase Konstruksi.....	IV-42
4.4.1 Kebutuhan Sistem.....	IV-42
4.4.2 Diagram Kelas.....	IV-43
4.4.3 Implementasi.....	IV-43
4.4.3.1 Implementasi Kelas.....	IV-43
4.4.3.2 Implementasi Antarmuka.....	IV-44
4.5 Fase Transisi.....	IV-46
4.5.1 Pemodelan Bisnis.....	IV-47
4.5.2 Rencana Pengujian.....	IV-47
4.5.3 Implementasi.....	IV-48
4.5.3.1 Pengujian Use Case Memasukkan Data.....	IV-48

4.5.3.2 Pengujian Use Case Melakukan Proses Klasifikasi Menggunakan Algoritma SVM.....	IV-48
4.5.3.3 Pengujian Use Case Melakukan Proses Klasifikasi Menggunakan Algoritma SVM dan Seleksi Fitur Chi-Square.....	IV-49
4.6 Kesimpulan.....	IV-50
BAB V HASIL DAN ANALISIS PENELITIAN.....	V-1
5.1 Pendahuluan.....	V-1
5.2 Data Hasil Penelitian.....	V-1
5.2.1 Konfigurasi Percobaan.....	V-1
5.2.2 Hasil Pengujian.....	V-1
5.3 Analisis Hasil Pengujian Secara Keseluruhan.....	V-6
5.4 Kesimpulan.....	V-13
BAB VI KESIMPULAN DAN SARAN.....	VI-1
6.1 Pendahuluan.....	VI-1
6.2 Kesimpulan.....	VI-1
6.3 Saran.....	VI-2
DAFTAR PUSTAKA.....	xiv
LAMPIRAN.....	xv

DAFTAR TABEL

Tabel II-1 Tabel Kontingensi Term dan Class.....	II-8
Tabel II-2 Confusion Matrix.....	II-10
Tabel III-1 Contoh Data Terkumpul.....	III-2
Tabel III-2 Data yang Mengalami Proses Cleaning.....	III-5
Tabel III-3 Data yang Mengalami Proses Case Folding.....	III-6
Tabel III-4 Data yang Mengalami Proses Tokenizing.....	III-6
Tabel III-5 Rancangan Tabel Data Hasil Confusion Matrix.....	III-9
Tabel IV-1 Kebutuhan Fungsional.....	IV-2
Tabel IV-2 Kebutuhan Non-Fungsional.....	IV-3
Tabel IV-3 Contoh Data Awal.....	IV-5
Tabel IV-4 Hasil Cleaning Data.....	IV-6
Tabel IV-5 Hasil Case Folding.....	IV-7
Tabel IV-6 Data yang Mengalami Tokenizing.....	IV-8
Tabel IV-7 Hasil Pembobotan TF-IDF.....	IV-9
Tabel IV-8 Hasil Pembobotan Kata TF-IDF.....	IV-11
Tabel IV-9 Perhitungan Bobot Nilai Chi-Square.....	IV-17
Tabel IV-10 Hasil Seleksi Fitur Chi-Square.....	IV-20
Tabel IV-11 Hasil Ekstraksi Fitur TF-IDF.....	IV-20
Tabel IV-12 Tabel Definisi Aktor.....	IV-26
Tabel IV-13 Tabel Definisi Use Case.....	IV-26
Tabel IV-14 Tabel Skenario Proses Data.....	IV-27
Tabel IV-15 Tabel Skenario Usecase Klasifikasi SVM.....	IV-28
Tabel IV-16 Tabel Skenario Usecase Klasifikasi SVM+CS.....	IV-30
Tabel IV-17 Tabel Implementasi Kelas.....	IV-44
Tabel IV-18 Rencana Uji Use Case Proses Data.....	IV-47
Tabel IV-19 Rencana Uji Use Case Klasifikasi Menggunakan Support Vector Machine.....	IV-47
Tabel IV-20 Rencana Uji Use Case Melakukan Klasifikasi dengan SVM+Chi- Square.....	IV-48
Tabel IV-21 Pengujian Use Case Proses Data.....	IV-48
Tabel IV-22 Pengujian Use Case Klasifikasi dengan SVM.....	IV-49
Tabel IV-23 Pengujian Use Case Klasifikasi dengan SVM+Chi-Square.....	IV-50
Tabel V-1 Confusion Matrix SVM dengan $C = 0,1$	V-2
Tabel V-2 Confusion Matrix SVM dengan $C = 0,5$	V-2
Tabel V-3 Confusion Matrix SVM dengan $C = 1$	V-2
Tabel V-4 Confusion Matrix SVM dengan $C = 5$	V-3
Tabel V-5 Confusion Matrix SVM dengan $C = 10$	V-3
Tabel V-6 Hasil Evaluasi Klasifikasi SVM tanpa seleksi fitur Chi-Square.....	V-3

Tabel V-7 Confusion Matrix SVM dengan $C = 0,1$	V-4
Tabel V-8 Confusion Matrix SVM dengan $C = 0,5$	V-4
Tabel V-9 Confusion Matrix SVM dengan $C = 1$	V-4
Tabel V-10 Confusion Matrix SVM dengan $C = 5$	V-5
Tabel V-11 Confusion Matrix SVM dengan $C = 010$	V-5
Tabel V-12 Hasil Evaluasi Klasifikasi SVM dengan seleksi fitur Chi-Square....	V-5
Tabel V-13 Perbandingan Nilai Accuracy Data Train dan Data Testing.....	V-10
Tabel V-14 Perbandingan nilai accuracy tiap kelas pada data training	V-11
Tabel V-15 Perbandingan nilai accuracy tiap kelas pada data testing.....	V-12

DAFTAR GAMBAR

Gambar II-1 Pemetaan data ke ruang dimensi lebih tinggi.....	II-6
Gambar II-2 Struktur Dimensi RUP.....	II-11
Gambar III-1 Tahapan Penelitian.....	III-3
Gambar III-2 Kerangka Kerja Penelitian.....	III-4
Gambar IV-1 Diagram Use Case.....	IV-25
Gambar IV-2 Rancangan Interface Halaman Utama.....	IV-33
Gambar IV-3 Rancangan Interface Halaman Hasil Preprocessing.....	IV-34
Gambar IV-4 Rancangan Interface Halaman Evaluasi Model dan Confusion Matrix.....	IV-35
Gambar IV-5 Rancangan Interface Halaman Hasil Klasifikasi.....	IV-36
Gambar IV-6 Diagram Aktivitas Proses Data.....	IV-37
Gambar IV-7 Diagram Aktivitas Melakukan Proses Klasifikasi Menggunakan Algoritma SVM.....	IV-38
Gambar IV-8 Diagram Aktivitas Melakukan Proses Klasifikasi Menggunakan Algoritma SVM dan Seleksi Fitur Chi-Square.....	IV-39
Gambar IV-9 Sequence Diagram Proses Data.....	IV-40
Gambar IV-10 Sequence Diagram Klasifikasi SVM.....	IV-41
Gambar IV-11 Sequence Diagram Klasifikasi SVM + CS.....	IV-42
Gambar IV-12 Class Diagram.....	IV-43
Gambar IV-13 Interface Halaman Utama.....	IV-45
Gambar IV-14 Interface Halaman Preprocessing.....	IV-45
Gambar IV-15 Interface Halaman Evaluasi Model dan Confusion Matrix....	IV-46
Gambar IV-16 Interface Halaman Klasifikasi.....	IV-46
Gambar V-1 Grafik perbandingan nilai akurasi SVM dengan SVM+CS.....	V-7
Gambar V-2 Grafik perbandingan nilai precision SVM dengan SVM+CS.....	V-8
Gambar V-3 Grafik perbandingan nilai recall menggunakan SVM dengan SVM+CS.....	V-8
Gambar V-4 Grafik perbandingan nilai F-measure menggunakan SVM dengan SVM+CS.....	V-9

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pada bab ini akan dijelaskan mengenai latar belakang, rumusan masalah, tujuan penulisan, manfaat penelitian, batasan penelitian, dan sistematika penulisan. Secara keseluruhan, skripsi ini menjelaskan mengenai bagaimana pengklasifikasian judul berita bahasa Indonesia menggunakan metode *Support Vector Machine* dan menggunakan penyeleksian fitur *Chi-Square*.

1.2 Latar Belakang Masalah

Berita atau *news* merupakan laporan yang berlangsung cepat mengenai sebuah fakta dan ide baru yang menarik dan berfungsi sebagai alat pencari informasi di seluruh negeri. Menurut Djuroto (2005:46) berita berasal dari bahasa *Sansekerta* yang memiliki arti terjadi. Berita merupakan sebuah kejadian atau peristiwa yang sudah bisa dipastikan sudah terjadi di suatu tempat dan sudah disiarkan atau diterbitkan dalam bentuk cetak maupun media online. Berita digunakan untuk memberikan sebuah informasi kepada masyarakat luas dengan konsep berita sebagai laporan yang dianggap paling cepat yang dapat berupa rekaman, fakta objektif, interpretasi, gambar, ramalan, minat insani, dan sensasi. Didalam susunan berita tentunya ada hal yang terpenting yang di buat semenarik mungkin

yang dirangkum dalam kata judul berita. Judul berita merupakan label yang berada diatas atau kepala berita dan memberikan informasi yang relevan terhadap isi berita kepada pembaca. Judul berita berfungsi sebagai perwakilan dari isi berita kepada para pembaca (Delfariyadi,dkk., 2022).

Support Vector Machine (SVM) merupakan algoritma dengan metode pembelajaran yang dapat menyelesaikan masalah klasifikasi *linear* maupun klasifikasi *non-linear*. Klasifikasi teks pada berita online menggunakan SVM pernah dijadikan penelitian oleh Liliana dkk (2011). Data yang dipakai adalah berita bahasa Indonesia yang di ambil dari www.kompas.com dengan 4 kategori dan 180 berita. Penelitian dilakukan dengan mencari pengaruh SVM terhadap klasifikasi berita online Indonesia serta kombinasi parameter terbaik di dalamnya yaitu parameter C dan gamma. Dari hasil penelitian ini, ditemukan bahwa akurasi terbaik dicapai ketika parameter C bernilai 110 dan gamma 1 yaitu akurasi sebesar 91%.

Penelitian lain dilakukan oleh Shahi & Pant (2018) dengan membandingkan 3 metode klasifikasi yaitu *Naive Bayes*, SVM, dan *Neural Networks* pada berita Nepal. Hasil dari penelitian ini diperoleh SVM dengan kernel RBF menjadi yang tertinggi dengan akurasi 74,65%. Namun, dalam proses pengklasifikasian teks, komputasi penggunaan metode SVM memakan waktu yang cukup lama tergantung pada banyaknya *fitur* yang digunakan sehingga diperlukan sebuah metode untuk menyeleksi *fitur* yang

dapat mengurangi jumlah *fitur* agar meningkatkan efisiensi dan performa dalam pengklasifikasian teks.

Pada penelitian yang dilakukan oleh Rahmad dan Pribadi (2015) mengenai pemilihan *feature* dengan *Chi-Square* untuk klasifikasi berita mengungkapkan bahwa *Chi-Square* merupakan sebuah *feature selection* yang mampu menghilangkan banyak *fitur* yang tidak relevan tanpa mengurangi tingkat akurasi sehingga mampu mengoptimalkan kinerja dari metode klasifikasi. Penelitian mengenai seleksi fitur pada klasifikasi berita pernah dilakukan juga oleh (Alshalabi dkk., 2013). Penelitian ini membandingkan seleksi *fitur Chi-Square* dengan *Information Gain*, hasil penelitian ini adalah seleksi fitur *Chi-Square* menjadi yang terbaik dengan nilai f-measure 96,14%.

Metode *Support Vector Machine* dan algoritma *Chi-Square Feature Selection* digunakan dalam penelitian klasifikasi judul berita bahasa indonesia berdasarkan hasil referensi penelitian sebelumnya.

1.3 Rumusan Masalah

Berdasarkan latar belakang, rumusan masalah dari penelitian berikut :

1. Bagaimana *Support Vector Machine* dapat digunakan untuk mengklasifikasikan judul berita bahasa indonesia pada 3 kategori?
2. Bagaimana dampak pemilihan seleksi fitur *Chi-Square* terhadap klasifikasi dengan metode *Support Vector Machine*?

1.4 Tujuan Penelitian

Tujuan dari penelitian yang dilakukan adalah :

1. Menghasilkan software yang dapat mengklasifikasi judul berita bahasa indonesia menjadi 3 kategori menggunakan metode *Support Vector Machine* dengan Seleksi Fitur *Chi-Square*.
2. Mengetahui pengaruh penggunaan metode seleksi fitur *Chi-Square* pada klasifikasi judul berita bahasa indonesia dengan *Support Vector Machine*.

1.5 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini adalah :

1. Mendapatkan hasil klasifikasi judul berita bahasa indonesia menggunakan *Support Vector Machine* dengan *Chi-Square*
2. Hasil penelitian dapat dijadikan sebagai referensi penelitian terkait

1.6 Batasan Masalah

Berikut batasan penelitian :

1. Data yang digunakan adalah data yang sudah dilabeli.
2. Data teks berita yang digunakan adalah 1170 judul berita bahasa Indonesia.
3. Klasifikasi dibagi menjadi 3 kategori: *edu,finance,sport*.
4. Fungsi *kernel* SVM yang digunakan dalam penelitian ini adalah *kernel* Linear.

1.7 Sistematika Penulisan

Sistematika penulisan yang digunakan pada penelitian ini mengikuti standar operasional penulisan tugas akhir Fakultas Ilmu Komputer Universitas Sriwijaya, yakni:

BAB I. PENDAHULUAN

Bab ini membahas tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan yang digunakan dalam penyusunan laporan akhir ini.

BAB II. KAJIAN LITERATUR

Bab ini menjelaskan mengenai landasan teori yang digunakan dalam menunjang penelitian. Pada bab ini memuat mengenai literatur dan penelitian terkait sebelumnya yang berkaitan dengan penelitian ini, seperti definisi berita, klasifikasi teks, *Support Vector Machine*, *Chi-Square*, dan beberapa tinjauan literatur dari penelitian terkait.

BAB III. METODOLOGI PENELITIAN

Bab ini akan menjelaskan mengenai tahapan-tahapan atau proses yang dilakukan selama penelitian seperti metode pengumpulan data hingga metode dalam perancangan perangkat lunak. Setiap tahapan penelitian akan dijelaskan secara rinci sesuai dengan kerangka kerja yang telah ditetapkan.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Bab ini akan membahas mengenai perancangan perangkat lunak mulai dari analisis kebutuhan perangkat lunak hingga pengujian pada perangkat lunak guna mengevaluasi pengembangan perangkat lunak.

BAB V. HASIL DAN ANALISIS PENELITIAN

Bab ini memaparkan hasil penelitian berdasarkan langkah dan metode yang telah direncanakan sebelumnya. analisis tersebut diberikan sebagai dasar kesimpulan yang akan diambil dari penelitian ini.

BAB VI. KESIMPULAN DAN SARAN

Bab ini memaparkan kesimpulan dari penelitian yang dilakukan berdasarkan uraian pada bab-bab sebelumnya dan memuat saran yang diharapkan dapat membuat sistem lebih baik lagi kedepannya.

1.8 Kesimpulan

Dengan uraian yang telah dijelaskan pada sub bab sebelumnya, penelitian ini akan membahas mengenai Klasifikasi Judul Berita Bahasa Indonesia menggunakan metode *Support Vector Machine* dan Seleksi Fitur *Chi-Square*

DAFTAR PUSTAKA

- Alfando., Hayami, R. 2023. Klasifikasi Teks Berita Berbahasa Indonesia Menggunakan Machine Learning dan Deep Learning: Studi Literatur.
- Alshalabi, H., Tiun, S., Omar, N., & Albared, M. 2013. Experiments on the Use of Feature Selection and Machine Learning Methods in Automatic Malay Text Categorization.
- Amalia., Lydia, M. S., Fadillah, S. D., & Huda, M. 2018. Perbandingan Metode Klaster dan Preprocessing Untuk Dokumen Berbahasa Indonesia.
- Ardiansyah, A. D., Abrianto, B. D., Andrai, B. R. E., Sani, A., & Nawaningtyas, N. 2023. Implementasi Algoritma Support Vector Machine Untuk Klasifikasi Komentar Spam Pada Instagram.
- Delfariyadi, F., Helen, A., Yuliawati, S. 2022. Klasifikasi Sentimen Judul Berita Pemberitaan COVID-19 Tahun 2021 pada Media DetikHealth.
- Dwiasnati, S., Devianto, Y. 2021. Optimasi Prediksi Bencana Banjir menggunakan Algoritma SVM untuk penentuan Daerah Rawan Banjir.
- Hendriyanto, M. D., Sari, B. N. 2022. Penerapan Algoritma K-Nearest Neighbor dalam Klasifikasi Judul Berita Hoax.
- Liliana, D. Y., Hardianto, A., & Ridok, M. 2011. Indonesian News Classification Using Support Vector Machine.
- Nasution, M. R. A., Hayaty, M. 2019. Perbandingan Akurasi dan Waktu Proses

Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter.

Rahmad, A. N. 2015. Pemilihan Feature dengan Chi-Square dalam Algoritma Naïve Bayes untuk Klasifikasi Berita.

Sani, D. A., Sarwani, M. Z. 2022. Koreksi Jawaban Esai Berdasarkan Persamaan Makna Menggunakan Fasttext dan Algoritma Backpropagation.

Santoso, E. B., Nugroho, A. 2019. Analisis Sentimen Calon Presiden Indonesia 2019 Berdasarkan Komentar Publik di Indonesia.

Shahi, T. B., & Pant, A. K. 2018. Nepal News Classification Using Naïve Bayes, Support Vector Machines and Neural Networks.