

**METODE SUPPORT VECTOR REGRESSION MENGGUNAKAN KERNEL  
POLYNOMIAL UNTUK AUTO ESSAY SCORING BAHASA INDONESIA**

**SKRIPSI**

Diajukan untuk memenuhi sebagian persyaratan mencapai gelar  
Sarjana Komputer

**DISUSUN OLEH :**

**RIZKY DWIKI FAISAL RAMADHAN**

**1811102441086**



**PROGRAM STUDI S1 TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR  
SAMARINDA  
2022**

**Metode Support Vector Regression Menggunakan Kernel  
Polynomial untuk Auto Essay Scoring Bahasa Indonesia**

**SKRIPSI**

Diajukan untuk memenuhi sebagian persyaratan mencapai gelar  
Sarjana Komputer

**Disusun Oleh :**

**Rizky Dwiki Faisal Ramadhan**

**1811102441086**



**PROGRAM STUDI S1 TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR  
SAMARINDA  
2022**

## HALAMAN PENGESAHAN

**METODE SUPPORT VECTOR REGRESSION MENGGUNAKAN KERNEL  
POLYNOMIAL UNTUK AUTO ESSAY SCORING BAHASA INDONESIA**

**DISUSUN OLEH :**

**RIZKY DWIKI FAISAL RAMADHAN**

**1811102441086**

Telah melaksanakan ujian skripsi dan dinyatakan lulus,


Pada tanggal 2 Juli 2022

Dosen Pembimbing



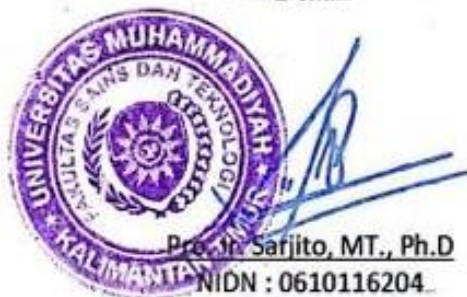
Naufal Azmi Verdikha, S.Kom., M.Eng  
NIDN : 1114048801

Penguji



Rofilde Hasudungan, S.Kom., M.Sc  
NIDN : 1107048601

Dekan



Prof. Dr. Sarjito, MT., Ph.D  
NIDN : 0610116204

Ketua Program Studi



Asslia Johar Latipah, S.Kom., M.Cs  
NIDN : 1124098902

## PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

**Nama mahasiswa** : Rizky Dwiki Faisal Ramadhan

**NIM** : 1811102441086

**Konsentrasi** : Sistem Cerdas

Menyatakan bahwa Skripsi dengan judul berikut:

“METODE SUPPORT VECTOR REGRESSION MENGGUNAKAN KERNEL  
POLYNOMIAL UNTUK AUTO ESSAY SCORING BAHASA INDONESIA”

Dosen Pembimbing : Naufal Azmi Verdikha, S.Kom., M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar Sarjana Komputer, baik di Universitas Muhammadiyah Kalimantan Timur (UMKT) maupun di Perguruan Tinggi lainnya
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas Muhammadiyah Kalimantan Timur (UMKT)
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi Muhammadiyah.

Samarinda, 2 Juli 2022

Yang Menyatakan,



The image shows a handwritten signature in black ink over a yellow postage stamp. The stamp is a 2000 Rupiah meterai tempel (postage meter) with the Garuda Pancasila emblem and the text 'METERAI TEMPEL' and '2000'. The stamp number is CA7DAAKX089889854.

Rizky Dwiki Faisal Ramadhan

1811102441086

## PRAKATA

Alhamdulillah dengan mengucapkan puji syukur kehadirat Allah SWT, karena dengan petunjuk-Nya skripsi ini dengan judul Automatic Essay Scoring dalam Bahasa Indonesia Menggunakan FastText dalam Natural Language Processing dapat penulis selesaikan dengan tepat waktu, sebagai salah satu persyaratan untuk menyelesaikan tugas perkuliahan dan untuk mencapai syarat kelulusan Universitas Muhammadiyah Kalimantan Timur Samarinda.

Segala kekurangan dalam penulisan skripsi ini sangat disadari oleh penulis, namun atas bantuan moril maupun materil dari berbagai pihak dapatlah terpenuhi segala kelengkapan dalam penulisannya. Untuk itu dalam kesempatan ini patutlah penulis menyampaikan ucapan terimakasih serta penghargaan yang setinggi-tingginya kepada berbagai pihak yang secara langsung maupun tidak langsung telah memberikan bantuannya.

1. Bapak Prof.Dr.Bambang Setiaji selaku rektor Universitas Muhammadiyah Kalimantan Timur, dan Ir.Waluyo Adi Siswanto, M.Eng,PhD selaku Dekan Fakultas Sains dan Teknologi Universitas Muhammadiyah Kalimantan Timur yang telah berkenan memberikan kesempatan kepada penulis untuk menyelesaikan studi di Universitas Muhammadiyah Kalimantan Timur.
2. Ibu asslia Johar Latipah,S.Kom., MCs selaku ketua program studi Teknik Informatika.
3. Bapak Naufal Azmi Verdikha,S.Kom., M.Eng selaku dosen pembimbing yang telah memberikan bimbingan masukan dan saran-saran yang sangat membantu dalam menyempurnakan skripsi ini.
4. Rudiman, S.Kom., M.Sc selaku dosen wali yang telah membantu penulis dalam mengikuti dan menyelesaikan studi di Fakultas Sains dan Teknologi Universitas Muhammadiyah Kalimantan Timur.
5. Seluruh dosen prodi Teknik informatika dan seluruh dosen Universitas muhammadiyah Kalimantan Timur yang telah banyak memberikan bekal ilmu

pengetahuan kepada penulis sehingga dapat menyelesaikan penulisan skripsi sebagai salah satu syarat untuk memperoleh gelar sarjana komputer.

6. Orang tua dan saudara tercinta yang selalu memberikan kasih sayang, do'a, nasehat, serta atas kesabarannya yang luar biasa menghadapi dalam setiap langkah hidup penulis, yang merupakan anugerah terbesar dalam hidup.
7. Rekan-rekan seperjuangan prodi Teknik Informatika Universitas Muhammadiyah Kalimantan Timur yang telah banyak memberikan bantuan pemikiran dan motivasi hingga penulis dapat menyelesaikannya.

Samarinda, 08 Februari 2022



Penulis

## ABSTRAK

*Automated Essay Scoring* (AES) merupakan salah satu pemanfaatan penggunaan *Natural Language Processing* (NLP) pada bidang Pendidikan yang memberikan solusi pada permasalahan belum tersedianya sistem yang dapat melakukan pengoreksian essay secara otomatis berbahasa Indonesia.

Penelitian ini bertujuan untuk membandingkan penggunaan kernel *polynomial* dengan penelitian terdahulu yang menggunakan kernel RBF. Data yang dimasukan akan melalui proses ekstraksi fitur menggunakan TF-IDF dan algoritma *Support Vector Regression* (SVR) dengan kernel *polynomial*. *Root Mean Squared Error* (RMSE) digunakan untuk membandingkan hasil penggunaan kernel, dimana semakin kecil hasilnya maka semakin baik. Hasil evaluasi RMSE menunjukkan hasil 2,225 berarti terdapat peningkatan nilai RMSE dengan penggunaan kernel *polynomial* sebesar 0,059 dibandingkan dengan penelitian terdahulu.

Kata kunci : *Automated Essay Scoring*; *Natural Language Processing*; Bahasa Indonesia; *Support Vector Regression*.

## **ABSTRACT**

*Automated Essay Scoring (AES) is one of the uses of Natural Language Processing (NLP) in the field of Education which provides a solution to unavailability of a system that can automatically essay scoring in Indonesian language.*

*This study aims to compare the use of a polynomial kernel with previous studies using the RBF kernel. The data entered will go through a feature extraction process using TF-IDF and the Support Vector Regression (SVR) algorithm with a polynomial kernel. Root Mean Squared Error (RMSE) is used to compare the results of using the kernel, where the smaller the result, the better. The results of the RMSE evaluation show the results of 2,225, which means that there is an increase in the RMSE value with the use of a polynomial kernel of 0,059 compared to previous studies.*

*Keywords : Automated Essay Scoring; Natural Language Processing; Indonesian language; Support Vector Regression.*



## DAFTAR ISI

PERNYATAAN KEASLIAN SKRIPSI .....	iii
PRAKATA.....	v
ABSTRAK .....	vii
<i>ABSTRACT</i> .....	viii
DAFTAR ISI .....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR .....	xii
LAMPIRAN .....	xiii
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Tujuan Penelitian.....	2
BAB 2 TINJAUAN PUSTAKA.....	4
2.1 Auto Essay Scoring.....	4
2.2 TF-IDF.....	7
2.4 Support Vector Regression.....	10
2.5 Fungsi Kernel Pada SVR .....	11
2.6 Root Mean Square Error.....	11
BAB 3 METODOLOGI PENELITIAN .....	13
3.1 Tahapan Penelitian .....	13
3.2 Dataset.....	14
3.3 Ekstraksi Fitur TF-IDF .....	15
3.4 Data Split .....	16

3.5	Regresi SVR dan Evaluasi .....	16
3.6	Perbandingan Kinerja .....	16
BAB 4 HASIL DAN PEMBAHASAN .....		17
4.1	Input Data .....	17
4.2	Ekstraksi Fitur Menggunakan TF-IDF .....	19
4.3	Data Split .....	19
4.4	Prediksi Menggunakan SVR dengan Kernel Polynomial.....	21
4.5	Evaluasi Menggunakan RMSE.....	22
4.6	Perbandingan Kinerja .....	22
BAB 5 PENUTUP .....		25
5.1	Kesimpulan .....	25
5.2	Saran.....	25
DAFTAR PUSTAKA.....		26
LAMPIRAN .....		29

## DAFTAR TABEL

Tabel 2.1 Penelitian Terkait .....	5
Tabel 2.2 Contoh data .....	7
Tabel 2.3 TF pada dokumen .....	7
Tabel 2.4 Hasil IDF .....	8
Tabel 2.5 Hasil TF-IDF .....	8
Tabel 3.1 Data penelitian .....	14
Tabel 4.1 Hasil input data.....	17
Tabel 4.2 Rincian data .....	18
Tabel 4.3 Rincian data split y.....	20
Tabel 4.4 Perbandingan nilai prediksi dan nilai aktual .....	22
Tabel 4.6 Perbandingan data split y.....	23
Tabel 4.7 Perbandingan evaluasi RMSE .....	23

## DAFTAR GAMBAR

Gambar 2.1 Ilustrasi SVR (Yudhawan, 2020).....	10
Gambar 2.2 Fungsi kernel (Ningrum, 2018).....	11
Gambar 3.1 Tahapan penelitian.....	13
Gambar 4.1 Distribusi data .....	18
Gambar 4.2 Output TF-IDF .....	19
Gambar 4.3 Grafik rincian data split y .....	20
Gambar 4.4 Data $y_{test}$ .....	21
Gambar 4.5 Perbandingan evaluasi RMSE .....	24

## LAMPIRAN

Lampiran 1 Riwayat Hidup .....	30
Lampiran 2 Source code dataset dan hasilnya .....	31
Lampiran 3 Ekstraksi fitur dan hasilnya .....	32
Lampiran 4 Data split dan hasilnya .....	33
Lampiran 5 Regresi SVR dan evaluasi serta hasilnya .....	34
Lampiran 6 Surat Ijin Penggunaan Data .....	35
Lampiran 7 Surat Keterangan Tidak Ui Validitas.....	36
Lampiran 8 Lembar Bimbingan Skripsi.....	37
Lampiran 9 Hasil Uji Plagiarisme .....	39