

SKRIPSI
ANALISIS SENTIMEN TWITTER TERHADAP PERNIKAHAN DI USIA
MUDA MENGGUNAKAN METODE SUPPORT VECTOR MACHINE
(SVM)



RADEN ISNAWAN ARGY ARYASATYA

NIM: 195410257

PROGRAM STUDI INFORMATIKA
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA

2023

SKRIPSI

**ANALISIS SENTIMEN TWITTER TERHADAP PERNIKAHAN DI USIA
MUDA MENGGUNAKAN METODE SUPPORT VECTOR MACHINE
(SVM)**

Diajukan sebagai salah satu syarat untuk menyelesaikan studi



Disusun Oleh

RADEN ISNAWAN ARGY ARYASATYA

NIM: 195410257

PROGRAM STUDI INFORMATIKA

PROGRAM SARJANA

FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA

YOGYAKARTA

2023

PERNYATAAN KEASLIAN SKRIPSI

Dengan ini saya menyatakan bahwa naskah skripsi ini belum pernah diajukan untuk memperoleh gelar Sarjana Komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 3 Agustus 2023



Raden Isnawan Argi Aryasatya

NIM: 195410257

HALAMAN PERSEMBAHAN

Alhamdulillah rabbil'alamin. Puji dan syukur penulis ucapkan kepada Allah SWT yang telah memberikan berkah dan karunia-Nya sehingga penulis dapat dengan lancar mengerjakan skripsi ini. Karya tulis ini dipersembahkan kepada:

1. Bambang Hanggoro Heru Nurcahyo dan Wernita Tampubolon selaku orang tua penulis yang selalu memberikan kasih sayang dan dukungan kepada penulis di segala kondisi.
2. Fidella Azra Daniswara selaku adik kandung penulis yang telah mengajari banyak hal tentang tren di media sosial dan pengetahuan umum lainnya kepada penulis.
3. Dr. Bambang Purnomosidi P. D. P., Akt., S.Kom. MMSI. selaku dosen sekaligus paman penulis yang merupakan seorang *role model*, contoh, dan teladan bagi penulis dalam proses menggapai cita-cita.
4. Seluruh keluarga besar saya yang selalu memberikan dukungan dan semangat kepada saya untuk meraih kesuksesan.
5. Almarhum kakek dan almarhumah nenek saya yang sekarang sudah tenang di surga.
6. Sahabat – sahabat saya dari SMP Negeri 1 Yogyakarta lulusan tahun 2015 dan SMA Negeri 5 Yogyakarta lulusan tahun 2018 yang selalu hadir saat saya mengajak mereka untuk berkumpul dan saling tertawa bersama.

7. Sahabat – sahabat saya dari Politeknik Keuangan Negara STAN angkatan 2018 yang telah memberikan kenangan - kenangan paling indah dalam hidup saya.
8. Teman – teman dari Universitas Teknologi Digital Indonesia, terutama teman – teman dari kelas IF-3 dan IF-5 yang telah menerima saya ke dalam lingkungan pertemanan mereka dan membuat saya merasa merasa tidak kesepian saat pertama kali menginjakkan kaki di kampus tercinta.

HALAMAN MOTTO

“Life shrinks or expands in proportion to one's courage.”

- Anais Nin -

“To persevere is important for everybody. Don't give up, don't give in. There is always an answer to everything.”

- Louis Zamperini -

“With great power, comes great responsibility.”

- Uncle Ben (Spider-Man, 2002) -

KATA PENGANTAR

Senantiasa penulis ucapkan puji dan syukur atas rahmat dan kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya, sehingga skripsi berjudul “Analisis Sentimen Twitter terhadap Pernikahan di Usia Muda menggunakan Metode Support Vector Machine (SVM)” dapat disusun dan diselesaikan dengan baik.

Skripsi ini disusun sebagai salah satu syarat untuk mendapatkan gelar strata-1 (S1) di program studi Informatika, Fakultas Teknologi Informasi, Universitas Teknologi Digital Indonesia. Dalam pengerjaan skripsi ini, penulis mendapatkan banyak dukungan dan bantuan dari beberapa pihak. Dalam kesempatan ini, penulis mengucapkan terima kasih kepada:

1. Bapak Ir. Totok Suprawoto, M.M., M.T., selaku Rektor Universitas Teknologi Digital Indonesia.
2. Bapak Ir. Muhammad Guntara, M.T., selaku Dekan Fakultas Teknologi Informasi.
3. Ibu Dini Fakta Sari, S.T, M.T., selaku Ketua Program Studi Informatika Universitas Teknologi Digital Indonesia.
4. Ibu Maria Mediatrix Sebatubun, S.Kom., M.Eng., selaku dosen pembimbing yang telah meluangkan waktunya untuk memberikan bimbingan selama berlangsungnya kegiatan penyusunan skripsi.
5. Kedua orang tua tercinta dan seluruh keluarga yang selalu mendukung saya.

6. Ibu Ariesta Damayanti, S.Kom., M.Cs. dan Ibu Sari Iswanti, S.Si., M.Kom., selaku dosen penguji yang telah memberikan arahan dan koreksi bagi penulis dalam menyusun skripsi.
7. Bapak Drs. Tri Prabawa, M.Kom., selaku dosen pembimbing PKL yang telah memberi ilmu kepada penulis tentang cara membuat karya tulis yaitu laporan PKL sehingga penulis dapat menyusun skripsi dengan lancar.
8. Seluruh dosen di Universitas Teknologi Digital Indonesia yang telah memberikan ilmunya kepada penulis selama perkuliahan di kampus ini.
9. Seluruh civitas akademika Universitas Teknologi Digital Indonesia yang telah banyak memberi bantuan dan dukungan selama penulis menempuh studi dan menyelesaikan skripsi di Universitas Teknologi Digital Indonesia.

Penulis menyadari bahwa masih banyak kekurangan dalam penyusunan skripsi ini. Oleh karena itu, penulis mengucapkan mohon maaf atas kesalahan yang terdapat di dalam skripsi ini. Penulis juga mengharapkan pembaca untuk menyampaikan kritik dan saran kepada penulis skripsi ini sebagai sarana introspeksi diri bagi penulis dalam membuat karya tulis di masa yang akan datang. Penulis berharap laporan ini dapat bermanfaat bagi para pembaca. Terima kasih.

Yogyakarta, 3 Agustus 2023



Raden Isnawan Argi Aryasatya

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN	v
HALAMAN MOTTO	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR	xiv
DAFTAR TABEL.....	xvi
INTISARI.....	xvii
ABSTRACT.....	xviii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Ruang Lingkup	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	5

BAB 2 TINJAUAN PUSTAKA DAN DASAR TEORI	7
2.1 Tinjauan Pustaka	7
2.2 Dasar Teori	11
2.2.1 Twitter	11
2.2.2 Snsrape	12
2.2.3 Python	12
2.2.4 Data Mining	13
2.2.5 Machine Learning	14
2.2.6 Analisis Sentimen	16
2.2.7 Text Preprocessing	16
2.2.8 Pelabelan Data.....	18
2.2.9 Ekstraksi Fitur	19
2.2.10 Support Vector Machine	21
2.2.11 Evaluasi Performansi	24
BAB 3 METODE PENELITIAN.....	26
3.1 Bahan/Data	26
3.1.1 Kebutuhan Input.....	26
3.1.2 Kebutuhan Proses.....	26
3.1.3 Kebutuhan Output	27
3.2 Peralatan	27

3.2.1	Kebutuhan Perangkat Lunak	27
3.2.2	Kebutuhan Perangkat Keras	27
3.3	Prosedur Pengumpulan Data	28
3.4	Analisis dan Rancangan Sistem.....	28
3.4.1	Block Diagram	28
3.4.2	Flowchart Support Vector Machine	31
3.4.3	Perancangan Antarmuka	34
BAB 4 IMPLEMENTASI DAN PEMBAHASAN SISTEM		38
4.1	Implementasi	38
4.1.1	Pengambilan Data	38
4.1.2	Preprocessing	39
4.1.3	Labeling.....	45
4.1.4	Ekstraksi Fitur	48
4.1.5	Implementasi SVM	49
4.1.6	Evaluasi Performansi	50
4.2	Uji Coba dan Pembahasan Sistem.....	54
4.2.1	Antarmuka Aplikasi Web.....	54
4.2.2	Hasil Uji Coba.....	57
BAB 5 PENUTUP		60
5.1	Kesimpulan.....	60

5.2	Saran.....	60
	DAFTAR PUSTAKA	62
	LAMPIRAN	

DAFTAR GAMBAR

Gambar 2.1 Tahapan <i>Knowledge Discovery Database</i> (Ependi & Putra, 2019)..	14
Gambar 2.2 Skema <i>Artificial Intelligence</i> dan <i>Machine Learning</i> (Roihan, et al., 2020)	15
Gambar 2.3 <i>Text Preprocessing</i>	17
Gambar 2.4 Support Vector Machine (Parapat, et al., 2018).....	21
Gambar 2.5 Pemetaan <i>Input Space</i> Berdimensi Dua dengan Pemetaan ke Dimensi Tinggi (Rahutomo et al., 2018).....	23
Gambar 3.1 Block Diagram	28
Gambar 3.2 Flowchart SVM.....	31
Gambar 3.3 Visualisasi <i>Hyperplane</i>	34
Gambar 3.4 Desain Halaman Beranda	35
Gambar 3.5 Desain Halaman Confusion Matrix.....	36
Gambar 3.6 Desain Halaman Classification Report	36
Gambar 3.7 Desain Halaman Sebaran Data.....	37
Gambar 4.1 Kode <i>Scraping</i> Bulan Maret.....	38
Gambar 4.2 Kode <i>Scraping</i> Bulan April.....	39
Gambar 4.3 Kode <i>Case Folding</i>	39
Gambar 4.4 Kode <i>Data Cleansing</i>	40
Gambar 4.5 Kode <i>Tokenizing</i>	41
Gambar 4.6 File Normalisasi	42
Gambar 4.7 Kode Normalisasi.....	42
Gambar 4.8 Kode <i>Filtering</i>	43

Gambar 4.9 Kode <i>Stemming</i>	44
Gambar 4.10 Kode Penggabungan Kata	45
Gambar 4.11 Google Spreadsheet.....	46
Gambar 4.12 Hasil Pelabelan dengan TextBlob	46
Gambar 4.13 <i>Bar Chart Labeling</i>	47
Gambar 4.14 <i>Pie Chart Labeling</i>	47
Gambar 4.15 Kode <i>Split</i> Data Latih dan Data Uji	48
Gambar 4.16 Kode <i>CountVectorizer</i>	48
Gambar 4.17 Kode Pelatihan dan Pengujian SVM.....	49
Gambar 4.18 Kode Akurasi SVM.....	49
Gambar 4.19 Kode <i>Confusion Matrix</i>	50
Gambar 4.20 Visualisasi <i>Confusion Matrix</i>	51
Gambar 4.21 Kode <i>Classification Report</i>	52
Gambar 4.22 <i>Classification Report</i>	53
Gambar 4.23 Halaman Beranda	55
Gambar 4.24 Halaman Confusion Matrix	56
Gambar 4.25 Halaman Classification Report.....	56
Gambar 4.26 Halaman Sebaran Data	57
Gambar 4.27 Input Teks dengan Hasil Klasifikasi Negatif	58
Gambar 4.28 Input Teks dengan Hasil Klasifikasi Netral	58
Gambar 4.29 Input Teks dengan Hasil Klasifikasi Positif.....	59

DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka	9
Tabel 2.2 Contoh DTM	20
Tabel 2.3 Contoh Penerapan <i>n-gram</i>	20
Tabel 2.4 <i>Confusion Matrix</i>	24
Tabel 3.1 Tabel Contoh Masukan Data Klasifikasi SVM	32
Tabel 3.2 Tabel Hasil Analisis Klasifikasi SVM	34
Tabel 4.1 Tabel <i>Case Folding</i>	39
Tabel 4.2 Tabel <i>Data Cleansing</i>	40
Tabel 4.3 Tabel <i>Tokenizing</i>	41
Tabel 4.4 Tabel Normalisasi	42
Tabel 4.5 Tabel <i>Filtering</i>	44
Tabel 4.6 Tabel <i>Stemming</i>	45
Tabel 4.7 Tabel Penggabungan Kata	45
Tabel 4.8 Tabel Hasil Pelabelan dengan TextBlob	47
Tabel 4.9 Tabel Perbandingan Akurasi Tiga Kernel SVM	50
Tabel 4.10 Tabel <i>Confusion Matrix</i>	51

INTISARI

Fenomena pernikahan di usia muda banyak terjadi di Indonesia. Pada tahun 2021, Komnas Perempuan mencatat ada 59.709 kasus pernikahan dini yang mendapat dispensasi. Hal itu membuktikan bahwa masih banyak masyarakat Indonesia yang tidak mengetahui atau tidak mengikuti aturan UU No. 1 tahun 1974 yang mengatur tentang umur minimal menikah yaitu 19 tahun. Topik ini pun selalu menjadi perbincangan hangat di Twitter. Ada netizen yang beropini baik, ada yang beropini jelek, dan ada pula yang netral. Hal tersebut dapat dianalisis menggunakan analisis sentimen.

Analisis sentimen dapat digunakan untuk menganalisis pendapat netizen dan mengelompokkannya menjadi tiga kategori yaitu positif, netral, dan negatif. Penelitian ini menggunakan metode *Support Vector Machine* (SVM) untuk melakukan klasifikasi yang dilakukan dengan garis pembatas (*hyperplane*) yang memisahkan kelas-kelas yang ada pada kumpulan data netizen Twitter yang memiliki opini tentang nikah muda.

Jumlah data yang digunakan sebanyak 4000 data yang diambil selama bulan Maret dan April 2023. Data dibagi dengan rasio 80:20 dimana 3200 data digunakan untuk data latih dan 800 data digunakan sebagai data uji. Penelitian ini menghasilkan akurasi *Linear Support Vector Machine* sebesar 87,375% dengan nilai presisi 82% untuk sentimen negatif, 40% untuk sentimen netral, dan 91% untuk sentimen positif. Selain itu juga dihasilkan nilai *recall* 64% untuk sentimen negatif, 44% untuk sentimen netral, dan 95% untuk sentimen positif. Terakhir, *f1-score* untuk sentimen negatif adalah 72%, lalu 42% untuk sentimen netral, kemudian 93% untuk sentimen positif.

Kata kunci: *analisis sentimen, nikah muda, support vector machine, twitter*

ABSTRACT

The phenomenon of marriage at a young age often occurs in Indonesia. In 2021, Komnas Perempuan stated that there were 59,709 cases of early marriage that received dispensations. This proves that there are still many Indonesian people who do not know or do not follow the UU No. 1 tahun 1974 which regulates the minimum age for marriage, which is 19 years old. This topic has always been a hot conversations topic on Twitter. There are netizens who have good opinions, there are those who have bad opinions, and there are those who are neutral. This can be analyzed using sentiment analysis.

Sentiment analysis can be used to analyze the opinions of netizens and group them into three categories, which are positive, neutral and negative. This study uses the Support Vector Machine (SVM) method to do the classification which is done with a dividing line (hyperplane) that separates the classes in the data set of Twitter netizens who have opinions about young marriage.

The data used in this research are 4000 data retrieved in March and April 2023. The data is divided by a ratio of 80:20, where 3200 data are used for training data and 800 data are used as testing data. This research produces a Linear Support Vector Machine accuracy of 87,375% with a precision value of 82% for negative sentiment, 40% for neutral sentiment, and 91% for positive sentiment. In addition, a recall value of 64% for negative sentiment, 44% for neutral sentiment, and 95% for positive sentiment were also generated. Finally, the f1-score for negative sentiment is 72%, then 42% for neutral sentiment, then 93% for positive sentiment.

Keywords: *sentiment analysis, support vector machine, twitter, young marriage*