
Penerapan *Data Mining* Untuk Klasifikasi Penyebab Kematian Menggunakan Algoritma *Support Vector Machine*

Nengah Widya Utami, Ni Nyoman Saptiari

STMIK Primakara, Indonesia^{1,2}

Email corresponding author: widya@primakara.ac.id

ABSTRAK

Kematian adalah hal yang paling memberikan perasaan cemas terutama bagi setiap manusia. Adanya suatu bencana merupakan suatu penyebab terjadinya kematian dimana bencana dapat diklasifikasikan yaitu bencana alam, bencana non alam dan penyakit, dan bencana sosial. Penelitian ini bertujuan untuk membantu mengetahui bencana apakah yang paling berdampak pada menurunnya jumlah penduduk Indonesia, dengan mengetahui ini akan memberikan persiapan dini dalamantisipasi bencana di Indonesia. Metode SVM (*Support Vector Machine*) digunakan pada penelitian ini untuk analisis data mining dalam klasifikasi penyebab kematian di Indonesia. Hasil dari analisa yang telah diteliti menggunakan pendekatan SVM (*Support Vector Machine*) menggunakan data sampling sebanyak 81% dan sample pengganti sebanyak 20 instances dengan perulangan sebanyak 100 kali jadi hasil dari penelitian ini yaitu penyebab yang paling tinggi dikarenakan terjadinya bencana non alam dan penyakit yang mencapai hingga 0,9 persen.

Kata Kunci: Data Mining, Support Vector Machine, Klasifikasi

PENDAHULUAN

Suatu negara pasti akan memperhitungkan jumlah penduduknya dimana untuk menstabilkan antara luas wilayah dan jumlah penduduk. Perhitungan Negara yang diambil adalah angka kelahiran dan angka kematian. Kematian atau mortalitas adalah hal pasti akan dialami oleh semua makhluk hidup. Setiap kematian juga memberi efek yang berbeda-beda pada setiap individu [1]. Adanya perbedaan kecemasan secara umum dengan kecemasan dengan kematian. Dimana kecemasan kepada sebuah kematian adalah kondisi individu yang tidak menyenangkan dengan adanya rasa gelisah, sulit tidur, was-was, dan tidak nyaman akibat hal yang belum terjadi yaitu terlepasnya jiwa dari raga [2]. Adanya kematian alami yang bisa diantisipasi bencana non alam seperti AIDS, ANTRAKS, Avian Influenza dan penyakit berat lainnya. Untuuk kematian yang tidak dapat diantisipasi adalah sebuah bencana alam seperti angin puting beliung, angina topan, angin siklon dan bencana alam yang tidak dapat diketahui sebelumnya.

SVM singkatan dari *Support Vector Machine* memiliki prinsip dasar klasifikasi linier yang dapat dipisahkan, tapi SVM juga dikembangkan agar dapat bekerja pada

masalah problem non-linier [3]. Pada awalnya SVM dikembangkan untuk masalah klasifikasi antara dua kelas [4]. Bisa dikatakan ada beberapa keuntungan SVM menurut jurnal [5], yaitu: a. Ruang Input Dimensi Tinggi klasifikasi teks harus memiliki banyak fitur kemungkinan jumlah fitur lebih dari 1000, b) Ruang Dokumen Vektor, meskipun representasi dimensi tinggi, pada beberapa dokumen vector hanya mengandung beberapa elemen non-zero. Selain itu masalah kategorisasi teks yaitu linier terpisah.

Metode SVM (*Support Vector Machine*) digunakan pada penelitian ini untuk analisis data mining dalam klasifikasi penyebab kematian di Indonesia. Pada penelitian sebelumnya menunjukkan dimana SVM dapat digunakan dalam pengelompokan data berdasarkan atribut tertentu [6]. Dalam penelitian ini SVM digunakan untuk pengelompokan penyebab kematian di Indonesia berdasarkan bencana yang terjadi untuk menentukan bencana yang paling berdampak atas tingginya kematian di Indonesia. Berdasarkan permasalahan diatas, peneliti ingin melakukan klasifikasi penyebab kematian berdasarkan bencana yang terjadi. Penelitian ini bertujuan untuk membantu mengetahui bencana apakah yang paling berdampak pada menurunnya jumlah penduduk Indonesia, dengan mengetahui ini akan memberikan persiapan dini dalam antisipasi bencana di Indonesia.

METODE PENELITIAN

Tahapan-tahapan penelitian ini digunakan untuk memberi tujuan atau arah pada setiap tahap yang akan dilakukan yang meliputi:

1) Persiapan

Pada tahap ini dilakukan identifikasi terhadap masalah yang ditemukan, tinjauan pustaka, dan pengungkapan tujuan dari penelitian

2) Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data, implementasi algoritma SVM, dan pengujian.

a. Meneliti Data

Data penyebab kematian diambil dari halaman web kaggle.com untuk sumber dari data set adalah pusdatin.kemendes.go.id.

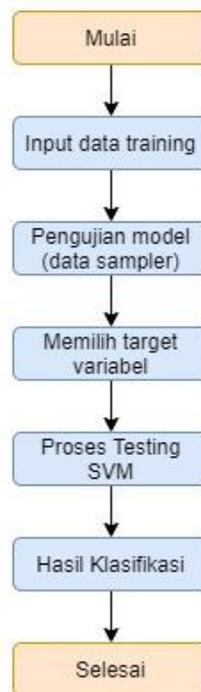
b. Membersihkan Data

Terdapat data yang kurang jelas dimana dilakukan pengurangan. Dalam penelitian ini, dari akumulasi 990 jumlah data yang ada. Peneliti mengambil 200 data dari tahun 2000 sampai 2019 setelah itu dilakukan pembersihan dan mendapatkan data sebanyak 163 data yang digunakan.

3) Implementasi SVM

Data yang didapat dari tahap diatas akan diimplementasikan ke dalam tools data *mining* yaitu menggunakan *orange* data *mining*. Dalam implementasi dilakukan langkah-langkah pengimputan data kedalam *tools* dan hasil nantinya akan memperlihatkan klasifikasi dari tipe bencana yang

menyebabkan kematian di Indonesia. Berikut adalah tahapan yang dilakukan dalam implementasi algoritma SVM.



Gambar 1. Tahapan Implementasi Algoritma SVM

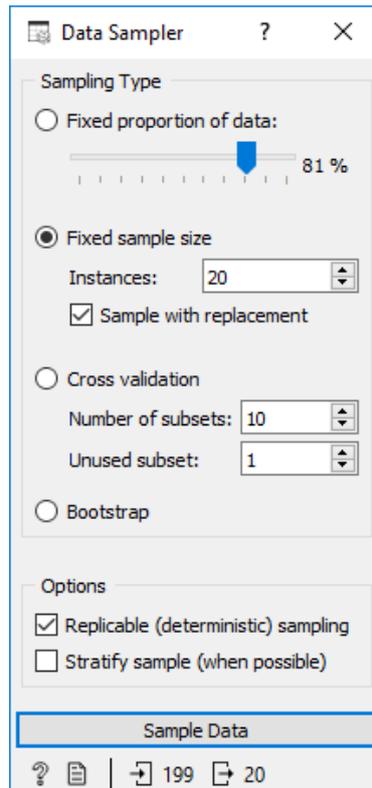
HASIL DAN PEMBAHASAN

Support Vector Machine

Data yang digunakan adalah 163 data dimana data ini adalah data penyebab kematian di Indonesia yang dikelompokkan 3 penyebab yaitu:

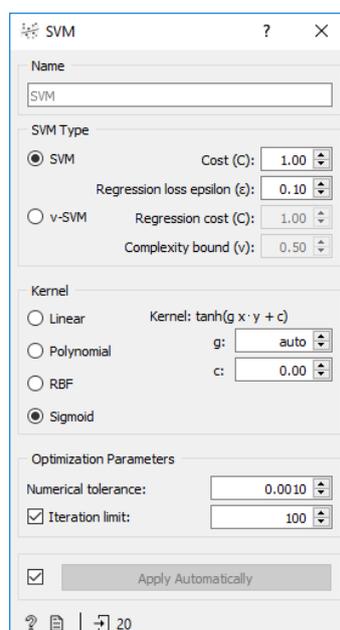
- 1) Bencana Alam (angin puting beliung, angin topan, angin siklon)
- 2) Bencana Non Alam dan Penyakit(AIDS, ANTRAKS, Avian Influenza)
- 3) Bencana Sosial (Aksi terror dan Sabotase)

Dibawah ini adalah implemtasi tool data mining yaitu orange. Pada gambar dibawah dijelaskan peneliti menggunakan data sampling sebanyak 81% data, untuk sample pengganti sebanyak 20 *instances*. Untuk *number subset* digunakan 10 dan *unused subset* 1.



Gambar 2. Data sampler

Parameter uji coba yang digunakan pada penelitian ini adalah $Cost (C) = 1$, Regression Loss Epsilon (ϵ) = 0,10. *Kernel* yang digunakan yaitu *sigmoid*. Jumlah perulangan dilakukan sebanyak 100 kali.



Gambar 3. Parameter SVM

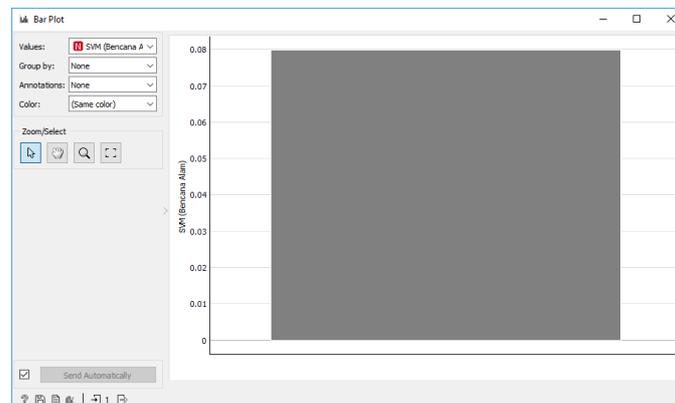
Hasil Clustering

Hasil Clustering menggunakan pendekatan SVM .Pengujian dilakukan menggunakan data sampling sebanyak 81% dan sample pengganti sebanyak 20 instances dengan perulangan sebanyak 100 kali. Gambar dibawah ini adalah prediksi menggunakan pendekatan SVM.

	SVM	Total Kematian	Kasus	Typ	Te
22	0.10 → Bencana Non Alam dan Penyakit	539	AIDS	Bencana Non ...	2006
23	0.10 → Bencana Non Alam dan Penyakit	635	AIDS	Bencana Non ...	2006
24	0.12 → Bencana Non Alam dan Penyakit	498	AIDS	Bencana Non ...	2007
25	0.12 → Bencana Non Alam dan Penyakit	498	AIDS	Bencana Non ...	2007
26	0.12 → Bencana Non Alam dan Penyakit	498	AIDS	Bencana Non ...	2007
27	0.12 → Bencana Non Alam dan Penyakit	498	AIDS	Bencana Non ...	2007
28	0.12 → Bencana Non Alam dan Penyakit	788	AIDS	Bencana Non ...	2007
29	0.13 → Bencana Non Alam dan Penyakit	933	AIDS	Bencana Non ...	2008
30	0.13 → Bencana Non Alam dan Penyakit	933	AIDS	Bencana Non ...	2008
31	0.13 → Bencana Non Alam dan Penyakit	933	AIDS	Bencana Non ...	2008
32	0.13 → Bencana Non Alam dan Penyakit	711	AIDS	Bencana Non ...	2008
33	0.15 → Bencana Non Alam dan Penyakit	464	AIDS	Bencana Non ...	2009
34	0.15 → Bencana Non Alam dan Penyakit	464	AIDS	Bencana Non ...	2009
35	0.15 → Bencana Non Alam dan Penyakit	331	AIDS	Bencana Non ...	2009
36	0.16 → Bencana Non Alam dan Penyakit	693	AIDS	Bencana Non ...	2010
37	0.16 → Bencana Non Alam dan Penyakit	979	AIDS	Bencana Non ...	2010
38	0.18 → Bencana Non Alam dan Penyakit	597	AIDS	Bencana Non ...	2011
39	0.34 → Bencana Non Alam dan Penyakit	614	AIDS	Bencana Non ...	2019
40	0.29 → Bencana Non Alam dan Penyakit	-	Aksi Teror dan ...	Bencana Sosial	2015
41	0.31 → Bencana Non Alam dan Penyakit	13	Aksi Teror dan ...	Bencana Sosial	2016
42	0.33 → Bencana Non Alam dan Penyakit	-	Aksi Teror dan ...	Bencana Sosial	2017
43	0.34 → Bencana Non Alam dan Penyakit	18	Aksi Teror dan ...	Bencana Sosial	2018
44	0.35 → Bencana Non Alam dan Penyakit	3	Aksi Teror dan ...	Bencana Sosial	2019
45	0.16 → Bencana Non Alam dan Penyakit	3	Angin Kencang	Bencana Alam	2008

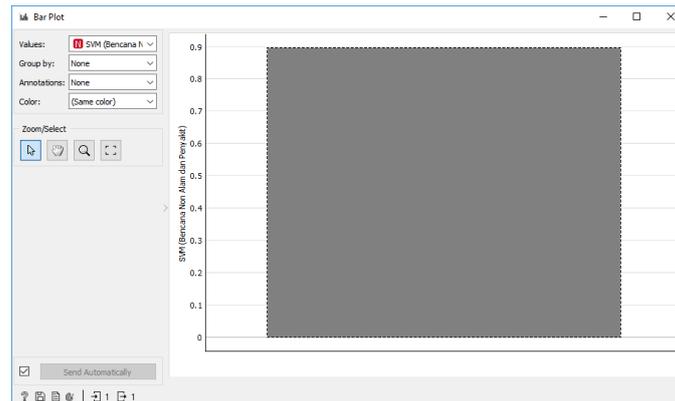
Gambar 4. Prediksi SVM

Untuk hasil dari klasifikasi penyebab bencana di Indonesia diakibatkan oleh bencana alam yang hasil presentasi nya hanya 0 sampai 0,8.



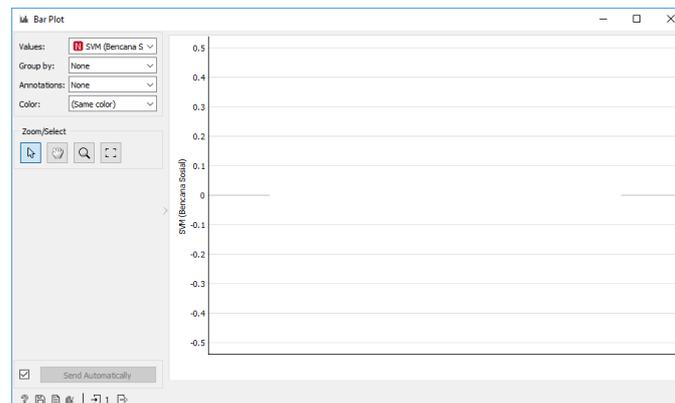
Gambar 5. Klasifikasi penyebab kematian di indonesia Karena bencana alam

Untuk hasil klasifikasi penyebab bencana di Indonesia karena adanya bencana non alam dan penyakit mencapai hasil paling tinggi yaitu 0 sampai 0,9.



Gambar 6. Klasifikasi penyebab bencana kematian di Indonesia karena bencana non alam dan penyakit

Untuk klasifikasi penyebab kematian di Indonesia dikarenakan bencana sosial hasil presentasi hanya pada titik 0 saja.



Gambar 7. Klasifikasi penyebab kematian di Indonesia karena bencana social

KESIMPULAN

Berdasarkan hasil penelitian diatas dapat diambil kesimpulan bahwa data awal dimiliki 990 dan dianalisa data hanya 200 mendapatkan data bersih sebanyak 163 dari data didapat adanya bencana alam, bencana non alam, dan bencana social yang dapat menyebabkan kematian di Indonesia. Data 165 ini di analisa untuk mengetahui bencana apa yang paling berpengaruh terhadap penyebab tingkat terjadinya kematian di Indonesia. Hasil dari analisa yang telah diteliti menggunakan pendekatan SVM (Support Vector Machine) menggunakan data sampling sebanyak 81% dan sample pengganti sebanyak 20 instances dengan perulangan sebanyak 100 kali jadi hasil dari penelitian ini yaitu penyebab yang paling tinggi dikarenakan terjadinya bencana non alam dan penyakit yang mencapai hingga 0,9 persen.

DAFTAR PUSTAKA

- Y. D. Astuti, "Kematian Akibat Bencana dan Pengaruhnya Pada Kondisi Psikologis," *J. Ilmu-Ilmu Sos.*, vol. 30, no. 66, pp. 363–376, 2007, [Online]. Available: <https://journal.uui.ac.id/Unisia/article/view/2680/2459>.
- E. Widianti, "Tingkat Kecemasan terhadap Kematian pada ODHA," *J. Keperawatan Soedirman*, vol. 12, no. 3, p. 199, 2018, doi: 10.20884/1.jks.2017.12.3.758.
- D. Alita, Y. Fernando, and H. Sulistiani, "Implementasi Algoritma Multiclass SVM Pada Opini Publik Berbahasa Indonesia Di Twitter," *J. TEKNOKOMPAK*, vol. 14, no. 2, p. 86, 2020, [Online]. Available: <https://ejurnal.teknokrat.ac.id/index.php/teknokompak/article/view/792>.
- P. A. Octaviani, Yuciana Wilandari, and D. Ispriyanti, "Penerapan Metode Klasifikasi Support Vector Machine (SVM) pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang," *J. Gaussian*, vol. 3, no. 8, pp. 811–820, 2014, [Online]. Available: [http://download.portalgaruda.org/article.php?article=286497&val=4706&title=PENERAPAN METODE KLASIFIKASI SUPPORT VECTOR MACHINE \(SVM\) PADA DATA AKREDITASI SEKOLAH DASAR \(SD\) DI KABUPATEN MAGELANG](http://download.portalgaruda.org/article.php?article=286497&val=4706&title=PENERAPAN METODE KLASIFIKASI SUPPORT VECTOR MACHINE (SVM) PADA DATA AKREDITASI SEKOLAH DASAR (SD) DI KABUPATEN MAGELANG).
- E. Indrayuni, "Analisa Sentimen Review Hotel Menggunakan Algoritma Support Vector Machine Berbasis Particle Swarm Optimization," *J. Evolusi Vol. 4 Nomor 2 - 2016*, vol. 4, no. 2, pp. 20–27, 2016.
- A. Handayanto, K. Latifa, N. D. Saputro, and R. R. Waliansyah, "Analisis dan Penerapan Algoritma Support Vector Machine (SVM) dalam Data Mining untuk Menunjang Strategi Promosi," *JUITA J. Inform.*, vol. 7, no. 2, p. 71, 2019, doi: 10.30595/juita.v7i2.4378.