****

Jurnal Matematika dan Pendidikan Matematika

Website: <http://ejournal.uinib.ac.id/jurnal/index.php/matheduca>

Email: mej.uinibpadang@gmail.com

 Math Educa Journalxx(x)(20xx):xx-xx

**PENERAPAN METODE *SUPPORT VECTOR MACHINE* DALAM KLASIFIKASI INDEKS PEMBANGUNAN MANUSIA DI SUMATERA UTARA**

***IMPLEMENTATION OF SUPPORT VECTOR MACHINE METHOD IN CLASSIFICATION OF HUMAN DEVELOPMENT INDEX IN NORTH SUMATERA***

1Aulia Yusharsah, 2Sajaratud Dur, 3Hendra Cipta

1,2,3Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara Medan, Indonesia

E-mail: 1auliayusharsah0116@gmail.com, 2sajaratuddur@uinsu.ac.id, 3hendracipta@uinsu.ac.id

Received: Month 20xx; Accepted: Month 20xx; Published: Month 20xx

**Abstrak**

Indeks Pembangunan Manusia pada tahun 2020 mengalami perlambatan pertumbuhan sebesar 0,04%. Hal ini disebabkan karena pengeluaran perkapita mengalami penurunan akibat adanya pandemi COVID-19. Tujuan penelitian ini adalah menerapkan suatu metode klasifikasi yang digunakan dalam klasifikasi Indeks Pembangunan Manusia (IPM) di Provinsi Sumatera Utara tahun 2020 dan mengetahui nilai akurasi yang diperoleh. Dalam menerapkan metode *support vector machine* dengan fungsi kernel *Radial Basic Function* (RBF) diperoleh nilai akurasi cukup baik sebesar 79,31% dengan parameter C=1, , ,  dan .

Kata kunci: Indeks Pembangunan Manusia (IPM), *Support Vector Machine*, *Radial Basic Function* (RBF)

***Abstract***

The human development index in 2020 get slowdown in growth of 0.04%. The problem becaused per capita expenditure has decreased due to the COVID-19 pandemic. The purpose of this research is to implementation of a classification method used in the classification of the human development index in North Sumatra Province in 2020 and find out the accuracy value obtained. In implementation of support vector machine method with the Radial Basic Function (RBF) kernel function got the accuracy value is enough good at 79.31% with parameters C=1, , ,  and .

*Keywords: Human Development Index, Support Vector Machine, Radial Basic Function (RBF)*

**Pendahuluan**

Pengertian Indeks Pembangunan Manusia (IPM) adalah bagian dari pengukuran kualitas untuk mengetahui pencapaian hasil pembangunan yang digunakan untuk melihat sejauh mana kualitas pembangunan manusia yang telah diperoleh. (Fuzi, 2017). Oleh karena itu, manusia merupakan kekayaan suatu Negara yang sesungguhnya. IPM terdiri atas tiga komponen utama, yaitu Angka harapan hidup (AHH) diukur dengan usia yang panjang dan kesehatan, angka melek huruf (AMH) diukur dengan pendidikan dengan kemampuan baca tulis dan angka partisipasi pendidikan yang telah ditamatkan atau rata-rata lama sekolah (RLS), serta ekonomi diukur dengan Standar Hidup yang layak dengan pendekatan Produk Domestik Bruto per Kapita pada tingkat konsumsi riil per kapita atau kemampuan daya beli masyarakat. (Kusumah, 2017).

Menurut BPS Sumatera Utara, klasifikasi status pembangunan manusia dibagi menjadi empat, apabila nilai IPM <60 dikategorikan rendah, 60≤IPM<70 dikategorikan sedang, 70≤IPM<80 dikategorikan tinggi dan ≥80 IPM sangat tinggi.

Berdasarkan BPS Sumatera Utara, IPM di Sumatera Utara pada tahun 2020 mencapai 71,77% sedangkan pada tahun 2019 sebesar 71,74%. Dengan pencapaian IPM tersebut, Sumatera Utara berada pada posisi status pembangunan kategori tinggi. IPM Sumatera Utara pada tahun 2020 mengalami peningkatan daripada tahun sebelumnya tetapi mengalami perlambatan pada pertumbuhannya.

Pada tahun sebelumnya IPM mengalami pertumbuhan sebesar 0,79% sedangkan pada tahun 2020 IPM mengalami perlambatan pertumbuhan sebesar 0,04%. Hal yang menyebabkan perlambatan IPM di Sumatera Utara adalah pengeluaran perkapita akibat adanya pandemi COVID-19. Banyak masyarakat Sumatera Utara mengalami penurunan pendapatan sejak adanya COVID-19 sehingga pengeluaran per kapitanya juga mengalami penurunan.

Salah satu metode statistik yang dapat diterapkan untuk melakukan klasifikasi adalah Support Vector Machine (SVM). Pada dasarnya,metode SVM digunakan untuk mengklasifikasikan data yang linier namun telah dikembangkan untuk digunakan dalam bentuk data yang non-linier dengan mengimplementasikan kernel trick. Tujuan dari metode *Support Vector Machine* adalah untuk memperoleh fungsi pemisah (klasifier) yang optimal yang dapat memisahkan dua set data yang berbeda (Fuzi, 2017). Metode ini mempunyai kelebihan, diantaranya yaitu menggunakan *support vector* untuk menentukan jarak agar komputasi yang diperoleh lebih cepat. SVM juga mempunyai tingkat akurasi klasifikasi yang lebih baik dibandingkan dengan metode klasifikasi lainnya, seperti *Naïve Bayes* dan *k-Nearest Neighbour* (KNN) (Sari, 2020).

Penelitian ini menerapkan metode *Support Vector Machine* dalam mengkasifikasikan indeks pembangunan manusia di Provinsi Sumatera Utara dan untuk mengetahui ketepatan hasil akurasinya.

**metode penelitian**

**Jenis Penelitian**

Penelitian yang digunakan adalah penelitian kuantitatif dimana kegiatan pengumpulan, pengolahan, analisis, dan penyajian data bersumber pada banyaknya data yang dilakukan secara objektif untuk memecahkan suatu masalah atau menguji suatu hipotesis untuk mengembangkan prinsip-prinsip umum (Dauli, 2019).

**Waktu dan Tempat Penelitian**

Penelitian ini dilaksanakan di Badan Pusat Statistik (BPS) Kota Medan yang beralamat di Jl.Gaperta N0.311 Kecamatan Medan Helvetia selama enam bulan terhitung sejak Februari 2021 sampai dengan selesai.

**Subjek Penelitian**

Penelitian ini mengambil data sekunder dari data Indeks Pembangunan Manusia (IPM) Provinsi Sumatera Utara tahun 2020 yang meliputi data penyususn IPM yaitu angka harapan hidup, pendidikan yang terdiri dari harapan lama sekolah dan rata-rata lama sekolah serta standar hidup layak yang diukur dengan pengeluaran perkapita.

**Prosedur**

Prosedur penelitian yang dilakukan adalah:

1. Pengumpulan data sekunder
2. Proses pembagian data *training* dan data *testing*
3. Normalisasi data *training* dan data *testing*
4. Menghitung data *training* dengan *Squential Training* yang akan dijadikan sebagai acuan untuk menentukan sebuah inputan data testing, dengan kategori rendah atau tinggi.
5. Menghitng nilai bobot dan nilai bias
6. Menghitung fungsi keputusan klasifikasi
7. Menghitung nilai akurasi
8. Membuat kesimpulan dan saran dari hasil yang diperoleh

**Teknik Pengumpulan Data**

Sumber data yang digunakan dalam penelitian ini adalah data Indeks Pembangunan Manusia Provinsi Sumatera Utara. Data diperoleh dari prusahaan dan websiteBadan Pusat Statistik (BPS) Kota Medan.

**Teknik Analisis Data**

Setelah data IPM diperoleh, tahap selanjutnya adalah pembagian data yakni data dibagi menjadi data *training* dan data *testing* kemudian dilakukan normalisasi data untuk memperkecil data yang yang berskala besar ataupun berskala kecil. Langkah selanjutnya yaitu menerapkan metode *Support Vector Machine* dengan kernel *Radial Basic Function.*

Pertama yaitu melakukan perhitungan *training* data SVM dengan menggunakan *Squential Training*. Hasil perhitungan *training* akan dijadikan sebagai acuan untuk menetukan sebuah inputan data tes IPM, dengan IPM kategori rendah atau tinggi. Kemudian melakukan perhitungan keputusan klasifikasi dengan menghitung nilai bobot dan bias. Setelah diperoleh nilai bobot dan bias, melakukan pengujian terhadap data latih dengan menghitung *dot product* antara data *testing* dan data *training* menggunakan fungsi kernel RBF.

**hasil penelitian dan pembahasan**

**Statistik Deskriptif**

Provinsi Sumatera Utara memiliki 33 Kabupaten/Kota yang terdiri dari 25 Kabupaten dan 8 Kota. Hasil analisis deskriptif IPM beserta komponen-komponennya dapat dilihat pada tabel 1.

**Tabel 1. Statistika Deskriptif Indeks Pembangunan Manusia**

|  |  |  |  |
| --- | --- | --- | --- |
| Komponen IPM | Min | Max | Mean |
| AHH | 69,10 | 73,55 | 69,02 |
| HLS | 62,60 | 14,74 | 13,18 |
| RLS | 12,23 | 11,39 | 9,02 |
| Pengeluaranperkapita | 5830 | 14890 | 10222,67 |

Pada tahun 2020, IPM di Provinsi Sumatera Utara sebesar 71,77 dan dikategorikan tinggi. Berdasarkan gambar 1 Kabupaten/Kota di Provinsi Sumatera Utara tahun 2020 terdapat 20 Kabupaten/Kota berada dalam kategori IPM tinggi dan 13 Kabupaten/Kota berada dalam kategori IPM rendah. Jumlah Kabupaten/Kota Provinsi Sumatera Utara dengan klasifikasi menurut Badan Pusat Statistik (BPS) dapat dilihat pada Gambar 1.

**Gambar 1. Deskriptif IPM**

**Normalisasi**

Pada tahap ini, normalisasi bertujuan untuk memperoleh data dengan ukuran yang lebih kecil yang mewakili data asli tanpa kehilangan karakteristik sendirinya (Wijayanti, 2018)

 (1)

dimana:

 = nilai setelah dinormalisasi

 = nilai sebelum dinormalisasi

 = nilai minimum dari fitur

 = nilai maksimum dari fitur

Dengan menggunakan persamaan (1) diatas, data dengan skala terlalu besar ataupun terlalu kecil akan diubah untuk mendapatkan data dengan ukuran yang lebih kecil dengan nilai 0 sampai 1 yang mewakili data asli.

**Data *Training* dan data *Testing***

 Pada tahap ini dilakukan pembagian data yang diambil dari keseluruhan data IPM Provinsi Sumatera Utara. Metode yang digunakan untuk membagi data *training* dan data *testing* adalah *k-fold* *cross validation*. Data dibagi secara random menjadi beberapa (*k*) bagian selanjutnya dilakukan *training* dengan menggunakan beberapa bagian data dan testing dengan bagian lainnya.

Jumlah *k* yang biasa digunakan adalah 10 (10-*fold* *cross* *validation*) karena memberikan estimasi akurasi yang tidak bias. Kemudian data akan dibagi ke dalam data *training* dan data *testing* dimana data *training* sebsar 90% dan data *testing* sebesar 10% dari 33 data IPM Provinsi Sumatera Utara. Untuk perhitungan manual, data *testing* yang digunakan adalah sebanyak 4 data dan data *training* sebanyak 29 data.

***Squential Training***

Proses *training* dari SVM adalah menggunakan *squential training*. Langkah-langkah dari *squential training*:

1. Menginisialisasi awal untuk nilai ,, dan.

Dalam penelitian ini, nilai , ditentukan masing-masing dari nilai, ,  dan C=1

1. Memasukkan data testing

X1 = Angka harapan Hidup

X2 = Harapan Lama Sekolah

X3 = Rata-Rata Lama Sekolah

X4 = Pengeluaran Per Kapita

Tabel 2 merupakan tabel data *testing* pada langkah sequential training yang sudah dinormalisasi. Dimana 1 adalah label positif (kategori tinggi) dan -1 adalah label negatif (kategori rendah).

**Tabel 2. Tabel Data *Testing***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| No | X1 | X2 | X3 | X4 | Y |
| C1 | 0,5808 | 0,2828 | 0,1873 | 0 | -1 |
| C2 | 0,0611 | 0,1075 | 0,6766 | 0,5852 | -1 |
| C3 | 0,7522 | 0,1952 | 0,8208 | 0,7777 | 1 |
| C4 | 0,9625 | 1 | 1 | 1 | 1 |

1. Menentukan *dot product* setiap data dengan memasukkan fungsi kernel (*k*)

Fungsi kernel yang digunakan adalah fungsi kernel RBF. Pada metode kernel, data tidak direpresentasikan secara individual, melainkan lewat perbandingan antara sepasang data dimana setiap data akan dibandingkan dengan dirinya dan data lainnya.

**Tabel 3. Tabel Perbandingan Data *Testing***

 **Dengan Data Lainnya**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| No | X1 | X2 | X3 | X4 | Y |
| C1 | 0,5808 | 0,2828 | 0,1873 | 0 | -1 |
| C2 | 0,0611 | 0,1075 | 0,6766 | 0,5852 | -1 |
| C3 | 0,7522 | 0,1952 | 0,8208 | 0,7777 | 1 |
| C4 | 0,9625 | 1 | 1 | 1 | 1 |

Simulasi perhitungan perbandingan antara sepasang data dengan data lainnya menggunakan persamaan:

 **  (2)



Semua data *testing* dihitung dengan cara yang sama, sehingga menghasilkan nilai *dot product.*

1. Menghitung *Matriks Hessian* dengan menggunakan persamaan:

  (3)

dimana:

= elemen *Matriks Hessian* ke-ij

= kelas data ke-i

= kelas data ke-j

= batas teoritis yang akan diturunkan.

Pada penelitian ini digunakan nilainya 0,5. Simulasi perhitungan *Matriks Hessian* untuk pasangan data C1 dengan C1 yakni:



1. Mencari nilai *error* dengan menggunakan persamaan:

 (4)

Simulasi perhitungan nilai *error*:



1. Menghitung nilai delta *alpha* menggunakan persamaan:

 (5)

Simulasi perhitungan delta *alpha*:



1. Menghitung nilai  baru menggunakan persamaan:

  (6)

Simulasi perhitungan *alpha* baru:



***Testing Support Vector Machine***

Sebelum nilai bias diperoleh, maka yang pertama dilakukan adalah menghitung nilai bobot (*w*) terlebih dahulu.

(kelas positif) yaitu bobot *dot product* data dengan alpha terbesar di kelas positif.

(kelas negatif) yaitu bobot *dot product* data dengan alpha terbesar di kelas negatif

Perhitungan manual didasarkan pada nilai  terbesar di kelas positif, yaitu 0,0832. Sedangkan nilai  di dasarkan pada niali terbesar dikelas negatif yaitu 0,0889. Berdasarkan nilai tersebut digunakan nilai *dot product* pada kolom 2 dan 3 adalah sebagai berikut:

**Tabel 4. Tabel Nilai *Dot Product* Kolom 2 dan 3**

|  |  |  |
| --- | --- | --- |
|  | C1 | C2 |
| C1 | 0,6432 | 0,4343 |
| C2 | 1 | 0,6927 |
| C3 | 0,7617 | 1 |
| C4 | 0,3948 | 0,6797 |

Simulasi perhitungan nilai bobot (*w*) adalah:





Setelah diperoleh nilai bobot (*w)*, maka nilai biasnya adalah:



Setelah nilai  diperoleh, langkah berikutnya yaitu melakukan pengujian terhadap data *training*. Dengan menghitung *dot product* antara data *testing* dan data *training* menggunakan fungsi kernel RBF.

Simulasi perhitungan *dot product* untuk data ke-1 adalah sebagai berikut:



Kemudian, langkah berikutnya melakukan perhitungan fungsi keputusan dengan menggunakan persamaan sebagai berikut:

 (7)

Simulasi perhitungan fungsi keputusan adalah:



Semua nilai diberi label -1 (kategori rendah) dan  diberi label +1 (kategori tinggi). Dari perhitungan diatas, diperoleh fungsi keputusan bernilai -1 maka data *training* ke-1 di atas termasuk kelas negatif atau kategori rendah. Data sebenarnya pada data *training* ke-1 adalah kategori rendah, setelah dilakukan prediksi dengan menggunakan SVM kernel RBF maka data *training* ke-1 termasuk TN yaitu *True Negatif*.

**simpulan dan saran**

**Simpulan**

Metode *Support Vector Machine* dapat diterapkan dalam pengklasifikasian Indeks Pembangunan Manusia (IPM) Provinsi Sumatera Utara pada tahun 2020. Hasil akurasi yang diperoleh dari klasifikasi Indeks Pembangunan Manusia (IPM) menggunakan metode SVM menghasilkan nilai akurasi cukup baik sebesar 79,31%. Kernel yang digunakan adalah kernel RBF dengan parameter C=1, , ,  dan dengan perbandingan data *training* sebesar 90% dan data *testing* 10%.

**Saran**

Pada penelitian ini, kernel yang digunakan adalah *Radial Basic Function* (RBF) untuk penelitian selanjutnya dapat menggunakan jenis kernel yang lain dan pemilihan parameter yang lebih tepat agar tingkat akurasinya semakin baik.

**Referensi**

Fuzi, F. (2017). *K-Nerest Neighbor* (K-NN) dan *Support Vector Machine*(SVM) untuk Klasifikasi Indeks Pembangunan Manusia (IPM) Provinsi Jawa Tengah. *Jurnal MIPA*, 40(2), 118-124.

Kusumah, D., Budi, W., dan Abdul, M. (2017). Perbandingan Metode *K–Means*dan *Self Organizing MAP* (Studi Kasus: Pengelompokan Kabupaten/Kota Di Jawa Tengah Berdasarkan Indikator Indeks Pembangunan Manusia . *Jurnal Gaussian*, 6(3), 429-437.

Sari, A., Ilhamsyah, A., Shariati, S. (2020). Klasifikasi Kabupaten Tertinggal Di Kawasan Timur Indonesia dengan *Support Vector Machine*. *Jurnal Informatika dan Komputer*, 3(3), 188-195.

Dauli, N. (2019). *Metodologi Penelitian Kuantitatif Beberapa Konsep Dasar Untuk Penulisan Skripsi dan Analisis Data dengan SPPS*. Yogyakarta: Deepublish.

Wiyanti, A., Tanzil, F., Sigit, A. (2018). Penerapan Algoritme *Support Vector Machine* Terhadap Klasifikasi Tingkat Rasiko Pasien Gagal Ginjal. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(10), 3500-3507.