

PENERAPAN ALGORITMA *SUPPORT VECTOR MACHINE* (SVM) DALAM PEMILIHAN BEASISWA: STUDI KASUS SMK YAPIMDA

LUKMAN

Lkmanaja51@gmail.com

(021) 94319769

Program Studi Teknik Informatika, Fakultas Teknik, Matematika dan IPA
Universitas Indraprasta PGRI

Abstrak. Dalam memajukan Bangsa dan Negara dibutuhkan generasi yang pintar dan cerdas. Salah satu factor yang penting adalah pendidikan, namun banyak siswa-siswi yang memiliki kemampuan dan potensi yang besar tidak bisa melanjutkan sekolah karena tidak mampu secara financial tetapi banyak juga siswa-siswi yang mampu yang mendapat beasiswa. Dalam lingkungan pendidikan terutama sekolah seharusnya ada beberapa peraturan atau klasifikasi dalam menentukan siswa-siswi yang mendapat beasiswa. Oleh karena itu dalam penelitian ini dilakukan algoritma Support Vector Machine (SVM) yang di aplikasikan terhadap data siswa-siswi yang mendapat beasiswa. Penelitian ini bertujuan untuk mengukur tingkat akurasi dari algoritma SVM, dalam pemilihan calon penerima beasiswa di SMK YAPIMDA Jakarta. Dari hasil pengujian dengan mengukur kinerja Algoritma tersebut menggunakan metode pengujian Cross Validation, Confusion Matrix dan Kurva ROC, diketahui bahwa algoritma Support Vector Machine (SVM) memiliki nilai accuracy paling tinggi yaitu 85.82% .

Kata Kunci: Beasiswa, pendidikan, Support Vector Machine (SVM)

Abstract. In advance of the Nation takes generations of smart and intelligent. One important factor is education, but many students who have the ability and great potential can not attend school because they can not financially but also many students who are capable of receiving a scholarship. In the educational environment, especially schools there should be some rules or classification in determining the students who received scholarships. Therefore, in this study the algorithm Support Vector Machine (SVM) which is applied to the data the students who got beasiswa. Penelitian aims to measure the level of accuracy of SVM algorithm, in the selection of scholarship recipients in vocational YAPIMDA Jakarta. From the test results to measure the performance of the algorithm using Cross Validation testing methods, Confusion Matrix and ROC curve, it is known that the algorithm Support Vector Machine (SVM) has the highest value of accuracy is 85.82%.

Keywords: Scholarship, Education, Support Vector Machine (SVM)

PENDAHULUAN

Pendidikan merupakan pondasi bagi sebuah pendidikan, kita bisa melihat jepang ketika negaranya hancur lembur oleh bom atom, maka pondasi yang utama adalah pendidikan, mereka mengutamakan pendidikan agar suatu saat dapat membangun negaranya lebih maju dan meraka berhasil melalukannya. Di negara berkembang seperti indonesia, pendidikan juga di prioritaskan oleh pemerintah tetapi pemisah antara kota dan desa seta miskin dan kaya jauh berbeda. Sekolah di perkotaan sangat bagus sedangkan di pedesaan jauh dari layak. Hal tersebut berakibat pendidikan yang tidak merata terutama bagi masyarakat yang tidak mampu. Solusi yang baik adalah dengan beasiswa bagi masyarakat yang membutuhkan, namun banyak beasiswa yang tidak tepat sasaran

Tahap *business understanding* disebut juga sebagai tahap *research understanding*, dalam tahap ini ditentukan tujuan dan requirement secara detail pada keseluruhan penelitian, merumuskan masalah data mining, dan menyiapkan strategi awal untuk mencapai tujuan yang telah ditentukan.

2. Tahap *Data Understanding* (Tahap Pemahaman Data)

Pada tahap ini mulai dilakukan proses pengumpulan data, menganalisis data, mengevaluasi kualitas data, dan memilih subset yang mungkin mengandung pola yang ditindaklanjuti

3. Tahap *Data Preparation* (Data Persiapan)

Pada tahap ini data akhir yang akan digunakan pada tahap berikutnya mulai disiapkan, memilih kasus dan variabel yang sesuai dengan analisis yang akan dilakukan, melakukan transformasi pada variabel tertentu jika diperlukan, dan membersihkan data mentah sehingga siap untuk digunakan sebagai alat pemodelan

4. Tahap *Modelling* (Tahap Pemodelan)

Pada tahap ini memilih dan menerapkan teknik pemodelan yang tepat, mengatur kalibrasi model untuk mendapatkan hasil yang optimal. Pada tahap ini juga dapat diterapkan beberapa teknik yang berbeda untuk permasalahan data mining yang sama, dan jika diperlukan proses dapat kembali ke tahap *data preparation* untuk menjadikan data kedalam bentuk yang sesuai dengan spesifikasi kebutuhan data mining tertentu.

5. Tahap *Evaluation* (Tahap Evaluasi)

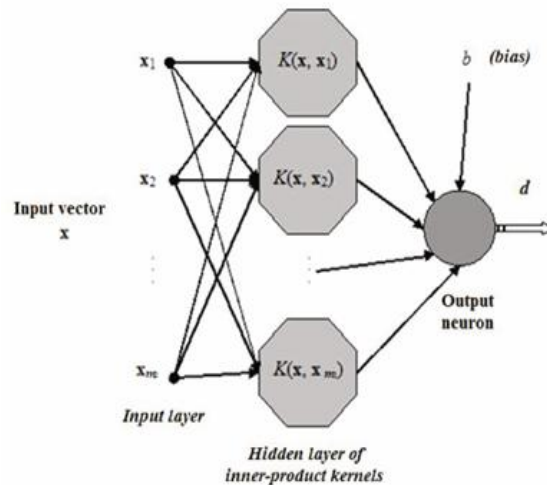
Setelah tahap *modelling* selesai dilakukan, model tersebut harus dievaluasi untuk melihat kualitas dan efektivitas sebelum disebarkan untuk digunakan. Pada tahap ini juga ditentukan apakah model dapat mencapai tujuan yang telah ditetapkan pada tahap pertama, apakah terdapat permasalahan penting dari penelitian yang tidak tertangani dengan baik, dan Mengambil keputusan berkaitan dengan penggunaan hasil dari data mining.

6. Tahap *Deployment*

Terbentuknya model tidak menandakan telah terselesaikannya proyek, perlu untuk menggunakan model yang dibuat sesuai dengan tujuan bisnis. Contoh sederhana dari tahap ini adalah pembuatan laporan, sedangkan contoh yang lebih kompleksnya yaitu dengan menerapkan proses data mining pada departemen lain secara paralel.

Support Vector Machine (SVM)

SVM adalah suatu teknik yang baru (1995) untuk melakukan prediksi, baik dalam kasus klasifikasi maupun regresi yang sangat populer pada saat ini. Svm berada dalam satu kelas dengan ANN dalam hal fungsi dan kondisi permasalahan yang biasa diselesaikan. Keduanya masuk dalam kelas supervised learning. Baik para ilmuwan ataupun praktisi telah banyak menerapkan teknik ini dalam menyelesaikan masalah-masalah nyata dalam kehidupan sehari-hari. Baik dalam masalah gene exception analysis, financial, cuaca hingga pada bidang kedokteran. terbukti dalam banyak implementasi, svm member hasil yang lebih baik dari ANN, terutama dalam solusi yang dicapai. ANN menemukan solusi berupa local optimal sedangkan SVM menemukan solusi yang global optimal



Gambar 2. Arsitektur *Support Vector Machine* (SVM) [2]

METODE

Penelitian Eksperimental

Penelitian eksperimental merupakan penelitian yang bersifat uji coba, memanipulasi dan mempengaruhi hal-hal yang terkait dengan seluruh variabel atau atribut.

Sampling/Metode Pemilihan Sampel

Populasi dalam penelitian ini merupakan siswa SMK YAPIMDA untuk kelas XI tahun ajaran 2011/2012. Responden atau sample dalam penelitian ini di ambil dari pihak-pihak yang terkait dalam pemilihan penerima beasiswa di SMK YAPIMDA Jakarta Selatan. Penentuan sampel dilakukan dengan teknik *proposive* (teknik penarikan sampel bertujuan) yaitu penarikan sampel dengan memiliki target dan tujuan tertentu dalam hal ini menerapkan model pengambilan keputusan dalam pemilihan 1 dari 2 buah algoritma berdasarkan metode Data mining dengan tujuan membantu dalam pengambilan keputusan pihak sekolah dalam memilih penerima beasiswa.

Metode Pengumpulan Data

Pengumpulan data dilakukan untuk mendapatkan informasi dan data yang diperoleh yang berhubungan dengan penelitian ini. Untuk mengumpulkan data dan informasi tersebut, dilakukan metode pengumpulan data sebagai berikut:

1. Pengumpulan data primer

Metode ini digunakan untuk mendapatkan data primer, yaitu dengan cara melakukan wawancara langsung dengan pihak Sekolah SMK YAPIMDA.

2. Pengumpulan data sekunder

Data sekunder dikumpulkan dengan mengamati data, membaca, mempelajari dan mengutip dari buku literatur, serta sumber-sumber yang berhubungan erat dengan penelitian ini.

Instrument Penelitian

1. Penelitian ini menggunakan Data sekunder berupa data siswa yang digunakan sebagai instrumentasi guna memperoleh data dalam proses penentuan beasiswa.

2. Data disajikan dalam bentuk Tabulasi model dan variabel masing- masing sebanyak 324 siswa terdiri dari siswa kelas XI tahun ajaran 2011/2012.

3. Perangkat lunak yang digunakan untuk menganalisis adalah *Rapid Miner 5.2*. dan *Graphical User Interface* (GUI) untuk menguji *role* algoritma terpilih adalah JAVA.4.5

Teknis Analisis dan Pengujian Data

Teknik Analisis data menggunakan Data Kuantitatif berupa kaidah-kaidah matematika terhadap anda atau numerik. Analisa dilakukan melalui data nilai rata-rata hasil tes potensi akademik dan nilai rata-rata raport semester 1 dan 2 menggunakan pengujian pada algoritma yaitu algoritma Support Vector Machine (SVM) . Rule yang diperoleh dari kedua algoritma tersebut kemudian diuji dengan confusion matrix dan kurva Receiver Operating Characteristic (ROC) .

Analisis, interpretasi dan implikasi penelitian

Pengelompokan Data

Untuk pemilihan penerima beasiswa, didapat data dari SMK YAPIMDA Jakarta sebanyak 324 data siswa, direduksi menjadi 272 yang terdiri dari 10 atribut. Dimana 9 atribut predictor dan 1 atribut hasil. Penelitian ini bertujuan untuk menentukan akurasi kelayakan pemberian beasiswa yang dibandingkan dengan menggunakan metode MultiLayer Perceptron. Setelah itu membandingkan nilai akurasi kedua metode tersebut.

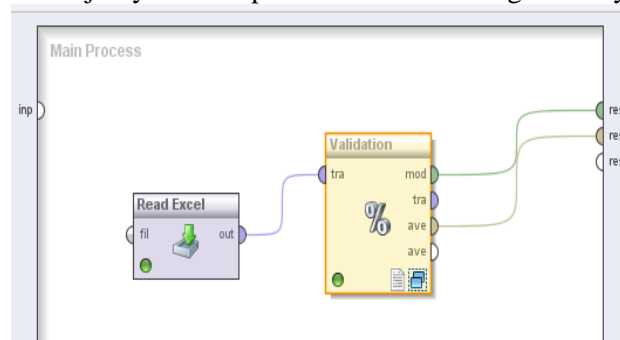
HASIL DAN PEMBAHASAN

Penyusunan Model *Data Mining*

Model data mining dalam penelitian ini dibuat dengan menggunakan perangkat lunak Rapid Miner 5.2. Algoritma berbasis jaringan syaraf tiruan (artificial neural network) yang berupa algoritma Support Vector Machine (SVM) dijalankan dengan menggunakan perangkat lunak tersebut. Model yang dihasilkan dijelaskan sebagai berikut:

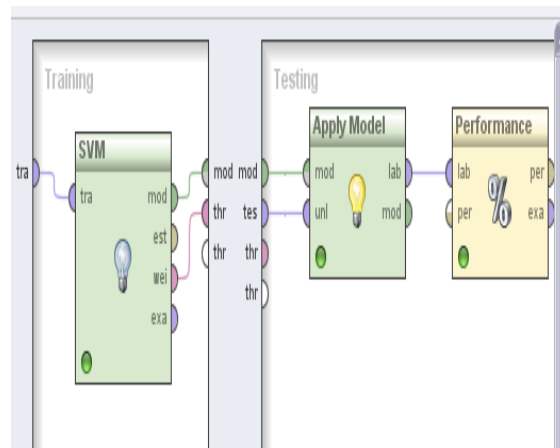
Algoritma *Support Vector Machine* (SVM)

pembuatan model algoritma Neural Network diawali dengan pembacaan file data (Read Excel). Data training disimpan dalam satu file Excell 2003. Lalu data di transform ke bentuk numeric, Langkah selanjutnya adalah penentuan model algoritmanya.



Gambar 3. Model Algoritma *Support Vector Machine* (SVM)-1

Langkah selanjutnya adalah penentuan model algoritmanya, dalam hal ini adalah *Support Vector Machine* (SVM). Pada tahapan ini, diatur nilai C dan epsilon yang digunakan dalam algoritma.



Gambar 4. Model Support Vector Machine (SVM) -2

Analisis dan Pengujian Model Data Mining

Model algoritma yang dihasilkan oleh perangkat lunak adalah *Rapid Miner 5.2* dievaluasi dan diuji dengan *confusion matrix* dan kurva *Receiver Operating Characteristic (ROC)* dapat dihasilkan juga dari *Rapid Miner 5.2*.

Model algoritma *Support Vector Machine (SVM)*

Confusion matrix dan kurva *Receiver Operating Characteristic (ROC)* yang dihasilkan dari model algoritma *Support Vector Machine (SVM)* dapat dijelaskan dibawah ini.

a. Confusion matrix

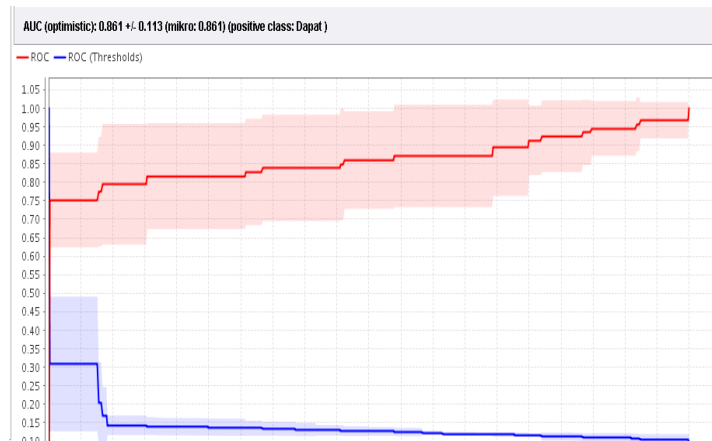
Data *Confusion matrix* untuk model algoritma *Support Vector Machine (SVM)* dapat dilihat pada tabel sebagai berikut.

Tabel 1. *Confusion matrix* Model Algoritma *Support Vector Machine (SVM)*

| accuracy: 83.98% +/- 5.00% (mikro: 83.94%) | | | |
|--|------------|------------|-----------------|
| | true Tidak | true Dapat | class precision |
| pred. Tidak | 125 | 35 | 78.12% |
| pred. Dapat | 0 | 58 | 100.00% |
| class recall | 100.00% | 62.37% | |

Dari tabel di atas dapat dilihat bahwa data *confusion matrix* tersebut menghasilkan tingkat akurasi 83.94%.

a. Kurva *Receiver Operating Characteristic (ROC)* Selain *confusion matrix*, kurva *Receiver Operating Characteristic (ROC)* dihasilkan oleh *Rapid Miner 5.2*. Kurva tersebut dapat dilihat di bawah ini.



Gambar 5. Kurva *Receiver Operating Characteristic* (ROC) Model Algoritma *Support Vector Machine* (SVM)

Dari gambar tersebut dapat diketahui bahwa nilai *The Area Under Curve* (AUC) model algoritma *Support Vector Machine* (SVM) adalah 0,861. Hal ini menunjukkan bahwa dalam penelitian ini, model algoritma *Support Vector Machine* (SVM) mencapai *excellent classification*.

Tabel tersebut menunjukkan bahwa model algoritma *Support Vector Machine* (SVM) menghasilkan tingkat akurasi 83.94% dan nilai *The Area Under Curve* (AUC) 0.861. Oleh karena itu, dalam penelitian ini, model algoritma *Support Vector Machine* (SVM) adalah model algoritma yang dapat digunakan dengan baik. Selanjutnya, Role bmodel algoritma *Support Vector Machine* (SVM) ini akan digunakan untuk pembuatan *Graphical User Interface* (GUI).

***Graphical User Interface* (GUI)**

Dengan menggunakan role model algoritma *Support Vector Machine* (SVM), dibangun suatu aplikasi dengan nama system beasiswa. *Graphical User Interface* (GUI) yang sudah dibangun dapat dilihat pada gambar 6.

Parameter yang digunakan pada gambar tersebut adalah Nama, Kelas,, Rangkaing, Nilai Rapot, Pekerjaan Orang Tua, Jumlah Tanggungan Oraang Tua dan Transportasi ke Sekolah. Nama diisikan sebagai identitas siswa data. Pengisian parameter tersebut dilakukan dengan menekan tombol . Untuk menampilkan hasil (*Result*) dilakukan dengan menekan tombol . Sedangkan untuk ke luar dari sistem ini dilakukan dengan menekan tombol .

Temuan dan Interpretasi Penelitian Berdasarkan penelitian yang telah dilakukan, terdapat beberapa temuan penelitian. Model algoritma *Support Vector Machine* (SVM) menghasilkan tingkat akurasi 83,98% Model algoritma *Support Vector Machine* (SVM) menghasilkan nilai *The Area Under Curve* (AUC) 0,861. Dari temuan-temuan tersebut terlihat bahwa algoritma *Support Vector Machine* (SVM) dapat digunakan untuk tugas *data mining* klasifikasi.

The screenshot shows a web-based graphical user interface titled "Klasifikasi Siswa". The interface is dark blue with white text and input fields. It is organized into several sections:

- Pilih Kelas:** Contains a "Nama" field with the value "joko" and a "Kelas" dropdown menu showing "XII MDM1".
- Input Data Orang Tua:** Contains four dropdown menus: "Rangking" (2-4), "Nilain Raport" (>80), "Kerajinan" (Baik), and "Keperbadian" (Baik).
- Other Data:** Contains four more dropdown menus: "Pekerjaan Orang Tua" (Tetap), "Penghasilan Orang Tua" (>2000000), "Jumlah Tanggungan Orang Tua" (2-4), and "Transportasi Ke Sekolah" (Sepeda).
- Bottom Bar:** Features four buttons: "Hasil" (with a document icon), "Input" (with a folder icon), "Selesai" (with a red X icon), and a "Dapat" label.

Gambar 6. *Graphical User Interface (GUI) Sistem Beasiswa*

Implikasi Penelitian

Penelitian ini mempunyai beberapa implikasi. Hal itu akan dijelaskan sebagai berikut.

1. Implikasi aspek sistem

Dikarenakan aplikasi beasiswa ini adalah program yang memerlukan perangkat yang baik, pihak sekolah harus menyediakan infrastruktur yang lebih baik sehingga, dalam penerapannya akan berjalan dengan maksimal. Pihak lembaga harus meningkatkan kualitas sistem, meng-*upgrade* perangkat keras terutama komputer yang digunakan dalam program ini. Peningkatan *memory*, *processor*, media penyimpanan dan software aplikasi, diharapkan dapat meningkatkan unjuk kerja dan stabilitas sistem.

2. Implikasi aspek manajerial

Pihak sekolah meningkatkan strategi dan keaktifan guru untuk memotifasi murid agar bersaing dalam mendapatkan beasiswa. Pihak lembaga harus meningkatkan kualitas informasi, kualitas sistem dan kualitas pelayanan, sehingga informasi mengenai beasiswa dapat segera diketahui. Pihak sekolah perlu mengadakan studi banding ke instansi lain yang telah berhasil menyelenggarakan beasiswa sebagai bahan masukan untuk meningkatkan kualitas murid-murid SMK Yapimda.

3. Implikasi aspek penelitian lanjut

Disarankan bagi penelitian selanjutnya akan dapat dilakukan di intitusi pemerintahan maupun pada masyarakat luas. Hasil penelitian ini diharapkan dapat dikembangkan pada penelitian selanjutnya dengan penambahan jumlah sampel dan variable. Diharapkan hasil penelitian ini dapat dikembangkan dalam penelitian lanjutan dengan menggunakan model atau pendekatan yang masih relevan dengan kasusnya. Hasil penelitian ini juga dapat diharapkan dapat dijadikan sebagai masukan untuk peningkatan kualitas siswa, sehingga tercapai proses belajar mengajar yang optimal.

PENUTUP

Simpulan

Dari uraian yang telah dijelaskan pada bab-bab sebelumnya, dapat disimpulkan hal-hal sebagai berikut.

1. Dengan menggunakan tugas *data mining* klasifikasi, algoritma yang berbasis jaringan syaraf tiruan (*artificial neural network*), yaitu: algoritma *Support Vector Machine* (SVM) dapat diterapkan untuk menentukan beasiswa pada sekolah SMK YAPIMDA .
2. Penerapan algoritma berbasis jaringan syaraf tiruan (*artificial neural network*) yang berupa algoritma *Support Vector Machine* (SVM) menghasilkan keputusan yang akurat sesuai dengan pengujiannya.

DAFTAR PUSTAKA

- Bramer, Max. 2007. *Principles of Data Mining*, Springer, London.
- Gorunescu. 2011. *Data Mining: Concepts, Models, and Techniques*, Verlag Berlin Heidelberg: Springer.
- Han, Jiawei and Kamber, M. 2007. *Data Mining: Concepts and Techniques*, University of Illinois at Urbana-Champaign.
- Maimon, Oded and Rokach, Lior. 2010. *Data Mining and Knowledge Discovery Handbook*, Springer, York.
- Larose, D. T. 2009. *Discovering Knowledge in Data*, New Jersey: John Willey & Sons, Inc, 2005.
- Vercellis *The Business Intelligence*, Politecnico di milano, Italy.
- Witten, I. H., Frank, E., & Hall, M. A. 2011. *Data Mining: Practical Machine Learning and Tools.*, Morgan Kaufmann Publisher, Burlington.