

Penerapan Metode Support Vector Machine (SVM) Guna Menentukan Tingkat Lulus Mahasiswa E-Learning

Sucitra Sahara

Abstract—E learning system is a web -based communication platform that enables learners , without limitation of place and time , to access a variety of learning tools such as discussion forums , assessment , content repositories , and document sharing system . E -Learning can be just as effective as face-to-face in a conventional classroom teaching and learning , if proper teaching techniques and well-organized (Oztekin et al . 2013) . Based on the data processing has been done with membandingkan algorithm and Support Vector Machine Support Vector Machine by using the log data of students . Later in the test to get the accuracy and AUC values of each algorithm so that the highest test results obtained by using support vector machine an accuracy of 81.02% , and AUC value of 0.810 .

Intisari-E-Learning adalah sistem web platform komunikasi berbasis yang memungkinkan peserta didik, tanpa batasan tempat dan waktu, untuk mengakses berbagai alat belajar, seperti forum diskusi, penilaian, repositori konten, dan sistem sharing dokumen. E-Learning bisa sama efektifnya dengan tatap muka dalam pengajaran di kelas konvensional dan belajar, jika teknik mengajar yang tepat dan terorganisir dengan baik (Oztekin et al. 2013). Berdasarkan pengolahan data yang telah dilakukan dengan algoritma membandingkan dan Support Vector Machine Support Vector Machine dengan menggunakan data log dari siswa. Kemudian di tes untuk mendapatkan akurasi dan nilai-nilai AUC setiap algoritma sehingga hasil tes yang tertinggi diperoleh dengan menggunakan mesin dukungan vektor dengan akurasi 81.02%, dan nilai AUC 0.810

I. PENDAHULUAN

Sistem e-learning adalah sebuah platform berbasis web komunikasi yang memungkinkan peserta didik, tanpa batasan tempat dan waktu, untuk mengakses beragam alat pembelajaran, seperti forum diskusi, penilaian, repositori konten, dan sistem sharing dokumen [14].

E-Learning bisa sama efektifnya dengan tatap muka di kelas konvensional mengajar dan belajar, jika teknik pengajarannya tepat dan terorganisir dengan baik [20].

Di Indonesia sendiri terkait dengan permen 24 tahun 2012, pemerintah mengeluarkan aturan tentang penyelenggaraan pendidikan jarak jauh (PJJ). Dimana dalam permen tersebut pembelajaran jarak jauh adalah pendidikan yang peserta didiknya terpisah dari pendidik dan pembelajarannya menggunakan berbagai sumber belajar melalui teknologi informasi dan komunikasi

(kemendiknas.go.id). Hal ini menjadikan e-learning sangat penting untuk mendapat perhatian.

Proses identifikasi perilaku siswa di dalam e-learning sangat penting untuk meningkatkan kegunaan dari system [5]. Kegagalan dan kesuksesan akademi pelaksanaan pembelajaran e-learning merupakan bagian dari mahasiswa yang telah menjadi subyek dari perdebatan. Banyak psikolog pendidikan telah mencoba untuk memahami masalah ini dan kemudian menjelaskannya, dan banyak ahli statistik telah mencoba untuk memprediksi [20]. Platform E-learning menghasilkan file log yang mengumpulkan semua informasi yang tersedia yang kemudian memberi kita kesempatan untuk menerapkan metode data penambangan untuk menemukan pola-pola tersembunyi, asosiasi, dan anomali hadir dalam data pendidikan dan menggunakan pengetahuan ini untuk meningkatkan proses pengambilan keputusan dalam sistem e-learning [20]. Peningkatan ini dapat membawa keuntungan seperti memaksimalkan efisiensi sistem pendidikan, mengurangi tingkat drop-out siswa, meningkatkan tingkat promosi siswa, meningkatkan retensi siswa, meningkatkan tingkat transisi siswa, meningkatkan rasio pendidikan, meningkatkan keberhasilan siswa, meningkatkan hasil belajar siswa, dan mengurangi biaya dari proses system [19].

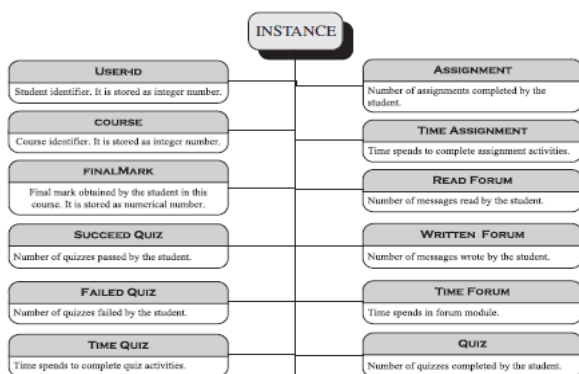
II. TINJAUAN PUSTAKA

Untuk menyusun penulisan ini penulis melakukan tinjauan studi dari berbagai sumber baik buku referensi, jurnal terkait maupun media online terpercaya. Hal ini dilakukan untuk memperkuat landasan dan pedoman dalam melakukan penelitian. Tinjauan studi berikut di jelaskan secara detail dalam bab ini.

1. Model Penelitian Zafra dan Cristobal Romero

Penelitian [19], mempunyai pendekatan baru berdasarkan beberapa contoh pembelajaran diusulkan untuk memprediksi kinerja siswa dan untuk meningkatkan hasil yang diperoleh dengan menggunakan klasik tunggal contoh pembelajaran. Beberapa contoh belajar memberikan representasi yang lebih cocok dan dioptimalkan untuk disesuaikan dengan informasi yang tersedia pada masing-masing siswa dan tentu saja menghilangkan nilai-nilai yang hilang yang membuat sulit untuk menemukan solusi yang efisien ketika belajar dengan metode tradisional. Untuk memeriksa efisiensi dari representasi yang diusulkan, teknik yang paling populer dari suatu metode pembelajaran tradisional yang didasarkan pada contoh tunggal yang didasarkan pada beberapa contoh pembel ajaran.

Penelitian ini menggunakan informasi siswa dari virtual lingkungan belajar di Cordoba Universitas menggunakan Moodle Platform. Platform ini menyimpan tugas yang dilakukan oleh siswa selama tahun akademik dari bulan September sampai bulan Juni, sebelum Ujian Akhir. Dalam rangka untuk mengumpulkan informasi, setiap pengguna dalam sistem diberikan sebuah identifier dan setiap kali dia / dia log ke kerangka, semua gerakan dalam tertentu Situs saja disimpan sehubungan dengan nya / akses nya ke konten dan alat-alat (misalnya, kalender, forum diskusi, email arsip, chatting, kuis, dan tugas). Dalam pekerjaan kami hanya dianggap sebagai informasi tentang kuis, forum dan tugas. informasi ini adalah diolah untuk bekerja dengan algoritma data mining menggunakan dua jenis representasi yang dijelaskan dalam Bagian



Sumber : Zafra et al.,(2011)

Gbr 1. Konsep Penelitian Zafra dan Cristobal Romero

2. Model Penelitian Sembiring dkk.

Salah satu fakta yang signifikan dalam pendidikan tinggi institusi adalah data pendidikan pertumbuhan eksplosif [16]. Data ini meningkat dengan cepat tanpa manfaat kepada manajemen. Tujuan utama dari setiap institusi pendidikan tinggi adalah untuk meningkatkan kualitas keputusan manajerial dan untuk memberikan pendidikan yang berkualitas. Prediksi yang baik dari keberhasilan siswa dalam institusi pendidikan tinggi adalah salah satu cara untuk mencapai tingkat tertinggi kualitas dalam sistem pendidikan tinggi. Ada banyak Model prediksi yang tersedia dengan pendekatan perbedaan siswa kinerja dilaporkan oleh peneliti, namun tidak ada kepastian apakah ada prediktor yang akurat menentukan apakah mahasiswa akan menjadi jenius akademik, drop out, atau seorang pemain rata-rata. Tujuan dari penelitian ini adalah untuk menerapkan. Metode kernel sebagai teknik data mining untuk menganalisis hubungan antara perilaku siswa dan keberhasilan mereka dan untuk mengembangkan model prediksi kinerja siswa. Hal ini dilakukan dengan menggunakan Smooth Support Vector Machine (SSVM) klasifikasi dan kernel k berarti teknik clustering. Itu Hasil

penelitian ini dilaporkan model akademik mahasiswa prediktor kinerja dengan menggunakan faktor psikometri sebagai variabel prediktor.

3. Model Penelitian Yin, Hand dan Cai.

Pendidikan koperasi Sino-asing memainkan peran penting dalam sistem pendidikan Cina modern dan semakin banyak siswa bergabung dengan proyek setiap tahun. Namun, beberapa siswa tidak bisa melakukan baik dalam proyek-proyek kerja sama seperti mengajar dalam gaya barat dan dalam bahasa lain [18]. Oleh karena itu, apa yang jenis siswa sesuai untuk proyek kerjasama Sino-asing? Makalah ini mengusulkan Sino asing analisis pilihan pendidikan koperasi menggunakan mesin dukungan vektor (SVM). Metode ini dapat membantu universitas untuk mencari tahu fitur apa siswa harus memiliki dalam proyek kooperatif dan memiliki jauh-mulai aplikasi dalam manajemen. Hasil percobaan menunjukkan bahwa algoritma yang diusulkan adalah efektif [19].

4. Model Penelitian Gokhan, Erdal dan Arif.

Model penelitian ini bertujuan untuk memprediksi nilai akhir siswa (A, B, C, D dan F) berdasarkan wiki data penggunaan. Data penggunaan disimpan dalam database wiki secara terbatas ketika pengaturan default digunakan. Oleh karena itu penelitian dikembangkan untuk memperluas kemampuan dengan menggunakan data log masuk siswa dan data navigasi. Meskipun pertambahan data pendidikan untuk memprediksi kinerja siswa bukan hal yang baru, namun penelitian dalam bidang ini masih banyak menyimpan banyak misteri. Makalah ini melaporkan hasil perbandingan Acak Forest, Support Vector Machines, Naif Bayes, dan didorong Klasifikasi Pohon untuk mengklasifikasikan siswa guna memprediksi nilai akhir yang diperoleh dalam program sarjana berdasarkan data penggunaan wiki. Kami membandingkan klasifikasi yang berbeda algoritma karena tidak ada satu algoritma tunggal yang memperoleh klasifikasi terbaik akurasi dalam semua kasus dan semua dataset. Menurut temuan dari penelitian ini, SVM melebihi metode lain [2]. Temuan ini menunjukkan bahwa metode data mining dapat membantu peneliti untuk menilai siswa berdasarkan kontribusi individu untuk wiki jika informasi yang diperlukan disimpan dalam database atau log file [2].

5. e-Learning

Menurut [14] e-learning adalah sebuah platform berbasis web komunikasi yang memungkinkan peserta didik, tanpa batasan tempat dan waktu untuk mengakses akses beragam alat pembelajaran, seperti forum diskusi, penilaian, repositori konten, dan sistem sharing dokumen.

Menurut permen 24 tahun 2012 pembelajaran jarak jauh adalah pendidikan yang peserta didiknya terpisah dari pendidik dan pembelajarannya menggunakan berbagai sumber belajar melalui teknologi informasi dan komunikasi (www.kemendiknas.go.id)

6. Data Mining

Menurut [20] data mining adalah proses penggalian pengetahuan yang berguna dan informasi dari kumpulan data. Saat ini data mining telah digunakan di banyak domain aplikasi seperti industri biomedis, ritel dan pemasaran, telekomunikasi, pertambangan web, audit komputer, industri keuangan, obat-obatan dan seterusnya.

7. Klasifikasi

Klasifikasi adalah proses menempatkan obyek atau konsep tertentu kedalam satu set kategori, berdasarkan sifat obyek atau konsep yang bersangkutan [6]. Metode klasifikasi ditujukan untuk pembelajaran fungsi-fungsi berbeda yang memetakan masing-masing data terpilih kedalam salah satu dari kelompok kelas yang telah ditetapkan sebelumnya. Proses klasifikasi didasarkan pada komponen [6].

a. Kelas (*class*)

Variabel dependen dari model yang merupakan kategori variabel yang mewakili label-label yang diletakkan pada obyek setelah pengklasifikasian. Contoh: kelas bintang dan kelas gempa bumi.

b. Prediktor (*predictors*)

Variabel independen dari model yang diwakili oleh karakteristik atau atribut dari data yang diklasifikasikan berdasarkan klasifikasi yang dibuat. Contoh: tekanan darah, status perkawinan dan musim.

c. Dataset pelatihan (*training dataset*)

Merupakan dataset yang berisi dua komponen nilai yang digunakan untuk pelatihan untuk mengenali model yang sesuai dengan kelasnya, berdasarkan prediktor yang ada. Contoh: database penelitian gempa, database badai dan database pelanggan supermarket.

d. Dataset pengujian (*testing dataset*)

Merupakan dataset baru yang akan diklasifikasikan oleh model yang dibangun sehingga dapat dievaluasi hasil akurasi klasifikasi tersebut.

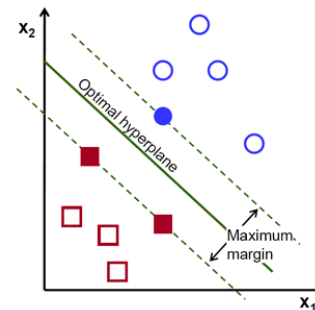
8. Algoritma Support Vector Machine

Menurut [15] *Support Vector Machine* (SVM) adalah metode yang berakar dari teori pembelajaran statistik yang hasilnya sangat menjanjikan untuk memberikan hasil yang lebih baik dari pada metode. SVM juga dapat bekerja dengan baik pada set data berdimensi tinggi, bahkan SVM yang menggunakan teknik kernel harus memetakan data asli dari dimensi asalnya menjadi dimensi lain yang relative lebih tinggi lain.

Support Vector Machine (SVM) didefinisikan sebagai seperangkat metode pembelajaran terkait yang menganalisis data dan mengenali pola, yang kemudian digunakan untuk klasifikasi dan analisis regresi. SVM mengambil satu set data input dan memprediksi untuk setiap masukan yang diberikan, yang berasal dari dua kelas yang kemudian di klasifikasikan dengan mencari nilai *hyperplane* terbaik [19].

Support Vector Machine (SVM) merupakan pembelajaran yang mengarah ke pemrograman kuadratik dengan kendala linear. Berdasarkan minimalisasi risiko prinsip terstruktur, SVM berusaha untuk meminimalkan

batas atas kesalahan generalisasi bukan kesalahan empiris, sehingga model prediksi baru efektif menghindari over-pas masalah. Selain itu, model SVM bekerja di ruang fitur berdimensi tinggi yang dibentuk oleh pemetaan nonlinear dari N-dimensi vektor input x ke dalam ruang fitur K-dimensi ($K > N$) melalui penggunaan fungsi ϕ nonlinear (x) [11].



Sumber : Prasetyo (2012)

Gbr 2. Konsep SVM untuk mencari hyperplane terbaik

Hyperplane (batas keputusan) pemisah terbaik antara kedua kelas dapat ditemukan dengan mengukur margin *hyperplane* tersebut dan mencari titik maksimalnya. Margin adalah jarak antara *hyperplane* tersebut dengan data terdekat dari masing-masing kelas. Data yang paling dekat ini disebut support vector. Garis solid pada gambar di atas menunjukkan *hyperplane* yang terbaik, yaitu yang terletak tepat pada tengah-tengah kedua kelas, sedangkan data lingkaran dan bujur sangkar yang dilewati garis batas margin (garis putus-putus) adalah support vector. Usaha untuk mencari lokasi *hyperplane* ini merupakan inti dari proses pelatihan *Support Vector Machine* (SVM) [15].

III. KERANGKA PEMIKIRAN

Kerangka pemikiran pada penelitian ini terdiri dari beberapa tahap. Permasalahan pada penelitian ini adalah banyaknya atribut sehingga akan mengurangi akurasi dan menambah kompleksitas dari algoritma Support Vector Machine. Penelitian ini dimulai dari adanya masalah dalam klasifikasi siswa dalam pembelajaran online menggunakan

SVM, di mana pengklasifikasi tersebut memiliki kekurangan yaitu sangat sensitif pada fitur yang terlalu banyak, yang mengakibatkan akurasi klasifikasi menjadi rendah. Dataset yang digunakan dalam penelitian ini berupa file log mahasiswa berupa jumlah login, jumlah akses dokumen, jumlah latihan, jumlah forum dan jumlah pesan. Preprocessing yang dilakukan dengan *remove duplicate*. Metode pemilihan fitur yang digunakan pengklasifikasi yang digunakan adalah SVM. Pengujian 10 fold cross validation akan dilakukan, akurasi algoritma akan diukur menggunakan *confusion matrix* dan hasil olahan data dalam bentuk kurva ROC. RapidMiner versi 5.3 digunakan sebagai alat bantu dalam mengukur akurasi data eksperimen.

IV. METODE PENELITIAN

A. Desain Penelitian

Dalam konteks sebuah penelitian, pendekatan metode yang digunakan untuk memecahkan masalah, diantaranya: mengumpulkan data, merumuskan hipotesis atau proposisi, menguji hipotesis, hasil penafsiran, dan kesimpulan yang dapat dievaluasi secara independen oleh orang lain [3]. Sedangkan menurut [4] terdapat empat metode penelitian yang umum digunakan, diantaranya: *Action Research*, *Experiment*, *Case Study*, dan *Survey*. Penelitian ini adalah penelitian eksperimen dengan metode penelitian sebagai berikut:

1. Pengumpulan data

Mengumpulkan data yang tersedia, memperoleh data tambahan yang dibutuhkan, mengintegrasikan semua data kedalam data set, termasuk variabel yang diperlukan dalam proses dan kemudian diseleksi dari data yang tidak sesuai

2. Pengolahan awal data

Pada bagian ini dijelaskan tentang tahap awal persiapan data olah. Pengolahan awal data meliputi proses input data ke format yang dibutuhkan, pengelompokan dan penentuan atribut data, serta pemecahan data (split) untuk digunakan dalam proses pembelajaran (training) dan pengujian (testing).

3. Model yang diusulkan

Pengujian menggunakan Algoritma Support Vector Machine (SVM)

4. Eksperimen dan pengujian model

Untuk eksperimen data penelitian, penulis menggunakan RapidMiner 5 untuk mengolah data. Sedangkan untuk pengujian metode, penulis membuat aplikasi menggunakan bahasa pemrograman PHP 5.4 dengan framework Code Igniter dan menggunakan Sublime Text sebagai editor.

5. Evaluasi dan validasi hasil

Pada bagian ini dilakukan evaluasi dan validasi hasil penerapan terhadap model penelitian yang dilakukan untuk mengetahui tingkat keakurasian model.

B. Pengumpulan Data

Penulis menggunakan data yang di ambil dari database platform e-learning Moodle (Modular Object Oriented Developmental Learning Environment) yang digunakan untuk proses belajar mengajar di Bina Sarana Informatika. Moodle sendiri mempunyai sekitar 280 tabel dalam databasenya. Dalam hal ini moodle di instalasi dengan metode single database. Artinya hanya ada satu database sebagai sumber data penyimpanan. Sehingga kelengkapan dan akurasi data sangat terjamin. .

C. Pengolahan Awal Data

Jumlah data awal yang diperoleh dari pengumpulan data yaitu sebanyak 400 data, namun tidak semua data dapat digunakan dan tidak semua atribut digunakan karena harus melalui beberapa tahap pengolahan awal data (preparation

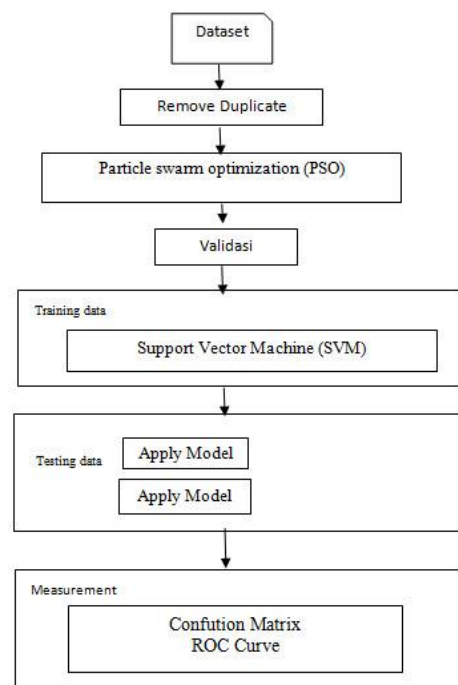
data). Untuk mendapatkan data yang berkualitas, beberapa teknik yang dilakukan sebagai berikut

1. Remove duplicate, menghapus data-data yang bernilai sama atau redukan.
2. Data validation, untuk mengidentifikasi dan menghapus data yang ganjil (outlier/noise), data yang tidak konsisten, dan data yang tidak lengkap (missing value).

Data integration and transformation, untuk meningkatkan akurasi dan efisiensi algoritma. Data yang digunakan dalam penulisan ini bernilai kategorikal. Data ditransformasikan kedalam software Rapidminer.

D. Model yang Diusulkan

Model yang diusulkan pada penelitian ini adalah menggunakan algoritma support vector machine dan algoritma support vector machine, yang terlihat pada gambar 3.



Gbr.3 Model yang diusulkan

Pada Gambar 3. menunjukkan proses yang dilakukan dalam Tahap modeling untuk menyelesaikan penentuan kelulusan mahasiswa e-learning dengan menggunakan dua metode yaitu algoritma support vector machine.

E. Eksperimen dan Pengujian Metode

Penelitian eksperimen melibatkan penyelidikan hubungan kausal menggunakan tes dikendalikan oleh si peneliti itu sendiri. Peneliti melakukan proses ekperimental menggunakan aplikasi RapidMiner 5.3. Sedangkan untuk pengujian model dilakukan menggunakan data file log e-learning Bina Sarana Informatika. Untuk implementasi model selain data training pada aplikasi RapidMiner, peneliti membuat aplikasi pengolahan file log menggunakan

bahasa PHP dan HTML. Spesifikasi perangkat lunak dan perangkat keras sebagai alat bantu dalam penelitian pada Tabel I.

TABEL I
SPESIFIKASI PERANGKAT LUNAK DAN PERANGKAT KERAS

Processor	Intel core i3-2,7Ghz
Memori	2GB
Harddisk	500GB
Sistem Operasi	Windows 8.1
Perangkat Lunak	RapidMiner 5.3
Editor	Sublime text 2
Bahasa Pemrograman	PHP

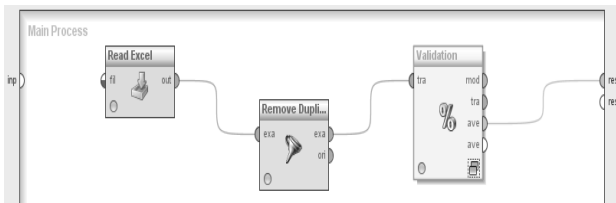
Sumber : Hasil penelitian (2017).

V. HASIL DAN PEMBAHASAN

A. Eksperimen dan Pengujian Metode

1. Metode Support Vector Machine

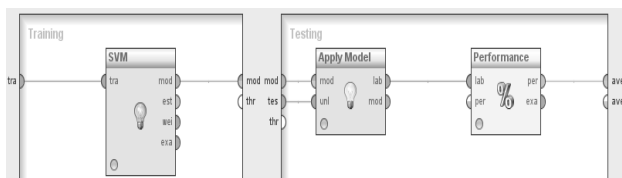
Pada tahap pertama peneliti bereksperiment dengan menguji data pada aplikasi Rapidminer menggunakan algoritma Support Vector Machine. Peneliti menggunakan fitur read excel untuk membaca data uji yang kemudian dihubungkan dengan remove duplicate dan validasi. Berikut gambar dari proses awal yang terdapat dalam frame *Main Proses*:



Sumber : Hasil penelitian (2017)

Gbr.4 Pengujian *K-Fold Cross Validation*

Didalam validasi terbagi dalam dua frame yaitu *Training* dan *Testing*. Pada frame training peneliti menggunakan algoritma Support Vector Machine (SVM) yang kemudian dihubungkan dengan Apply Model dan Performance pada frame Testing. Berikut adalah gambar pengujian algoritma Support Vector Machine menggunakan RapidMiner pada frame *training* dan *testing* :



Sumber : Hasil peneliti (2017)

Gbr 5. Pengujian algoritma *Support Vector Machine*

Berdasarkan eksperimen diatas, data yang peneliti proses penunjukkan hasil akurasi 80,30%. Berikut hasil pengujian dengan Support Vector Machine (SVM).

TABEL II
HASIL PENGUJIAN ALGORITMA *SUPPORT VECTOR MACHINE* (SVM)

accuracy: 81.30% +/- 6.640% (mikro: 81.25%)			
	true L	true T	class precision
pred. L	170	38	81.73%
pred. T	4	12	75.00%
class recall	97.70%	24.00%	

Sumber : Hasil penelitian (2017)

2. Evaluasi dan Validasi Hasil

Hasil pengujian model adalah untuk mengukur tingkat akurasi dan AUC (*Area Under Curve*) dari penentuan kelulusan dengan metode cross validation.

a. Confusion Matrix

Tabel III menunjukkan hasil dari confusion matrix metode support vector machine.

TABEL III HASIL CONFUSION MATRIX UNTUK METODE *SUPPORT VECTOR MACHINE*

Classification	Predicted Class	
Observed Class	170	38
	4	12

Sumber : Hasil penelitian (2017)

Keterangan dari hasil Confusion Matrix untuk Metode Support Vector Machine dapat dilihat dari table di atas. Nilai 170 menunjukkan bahwa hasil prediksi negative dan data sebenarnya negative. Nilai 38 menunjukkan hasil prediksi positif sedangkan nilai yang sebenarnya adalah negatif. Nilai 4 menunjukkan bahwa hasil prediksi negatif sedangkan nilai sebenarnya positif. Untuk yang terakhir nilai 12 menunjukkan bahwa hasil prediksi positif dan nilai sebenarnya positif. Tingkat akurasi yang dihasilkan dengan menggunakan algoritma support vector machine adalah sebesar **81.02%**, dan dapat dihitung untuk mencari nilai accuracy, sensitivity, specificity, ppv, dan npv pada persamaan dibawah ini:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{170 + 12}{170 + 12 + 4 + 38} = 0,8125$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} = \frac{170}{170 + 38} = 0,8173$$

$$\text{Specificity} = \frac{TN}{TN + FP} = \frac{12}{12 + 4} = 0,75$$

$$\text{PPV} = \frac{TP}{TP + FP} = \frac{170}{170 + 4} = 0,9770$$

$$NPV = \frac{TN}{TN + FN} = \frac{12}{12 + 38} = 0,24$$

Hasil perhitungan terlihat pada tabel IV

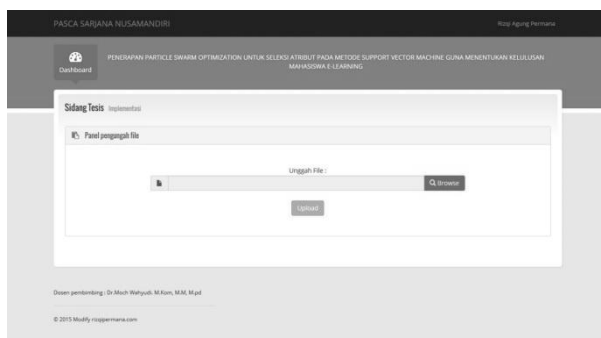
TABEL IV
NILAI ACCURACY, SENSITIVITY, SPECIFICITY, PPV DAN NPV
METODE SVM

	Nilai (%)
Accuracy	81.25
Sensitivity	81.73
Specificity	75
PPV	97.70
NPV	24

Sumber : Hasil penelitian (2017)

B. Penerapan Algoritma Terpilih

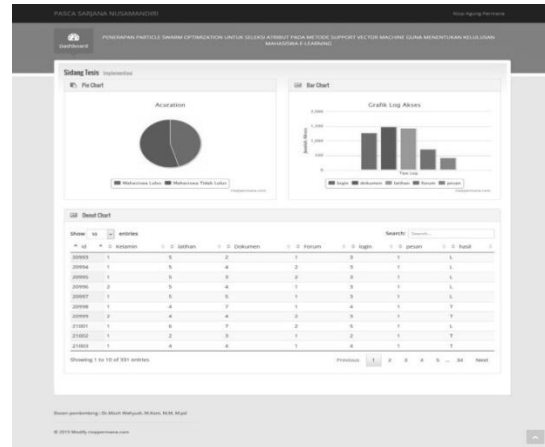
Penerapan dalam aplikasi sudah dapat di pastikan bahwa akan berpedoman pada nilai jumlah latihan sebagai atribut yang paling berpengaruh terhadap kelulusan mahasiswa e-learning. Untuk itu peneliti membuat prototype berupa aplikasi berbasis web untuk memudahkan dalam pemakai. Untuk tahapan penggunaanya, pertama user harus mengunggah file log mahasiswa dalam format Ms.excel 2003-2007 dengan cara mengklik tombol browse pada halaman web. Kemudian pilih tombol unggah maka file log akan di proses oleh palikasi, tunggu beberapa menit karena bisanya proses ini membutuhkan waktu sedikit lama. Estimasi waktu peoses tergantung dari jumlah data yang di upload. Interface dari aplikasi unggah file log tersebut bisa di lihat pada gambar 6.



Sumber : Hasil penelitian (2017)

Gbr 6. Graphical User Interface (GUI) Aplikasi Prediksi Kelulusan Mahasiswa E-learning

Selanjutnya pada bagian bawah terdapat susunan table yang merinci setiap log file persiswa elearning. Pada tampilan ini bisa di atur jumlah data yang ditampilkan dalam satu page, selain itu tampilan ini juga dilengkapi dengan fasilitas pencarian mahasiswa berdasarkan data yang ada. Tampilan output ini dapat dilihat pada gambar 7



Sumber : Hasil penelitian (2015)

Gbr 7. Graphical User Interface (GUI) Aplikasi Prediksi Kelulusan Mahasiswa E-learning

VI. KESIMPULAN

Berdasarkan pengolahan data yang sudah dilakukan dengan membandingkan algoritma Support Vector Machine dengan menggunakan data log mahasiswa. Kemudian di uji untuk mendapatkan nilai accuracy dan AUC dari setiap algoritma sehingga didapat hasil pengujian tertinggi dengan menggunakan support vector dengan nilai accuracy sebesar **81.02%**, dan nilai AUC sebesar **0.810**. Dibandingkan dengan hanya menggunakan Support Vector Machine (SVM) hasil ini lebih tinggi dengan selisih nilai akurasi sebesar 1.77% dan selisih AUC 0.006.

Dari 6 atribut yang digunakan dalam seleksi atribut yang peneliti gunakan yaitu jenis kelamin, jumlah latihan, jumlah login, jumlah akses document ,jumlah forum dan jumlah pesan diperoleh bahwa jumlah latihan sebagai atribut yang paling berpengaruh terhadap hasil kelulusan mahasiswa dengan nilai probabilitas 0.898. Hal ini membuktikan bahwa benar terjadi peningkatan akurasi ketika Support Vector Machine (SVM).

Hasil pengujian ini membuktikan bahwa Support Vector Machine (SVM) memiliki kemampuan generalisasi yang sangat baik untuk memecahkan masalah walaupun dengan sampel yang terbatas. Untuk mendukung penelitian dan sebagai kontribusi penulis membuat prototype aplikasi berbasis web guna mengembangkan dan memudahkan dalam menentukan kelulusan mahasiswa elearning menggunakan bahasa pemrograman PHP berplatform framework Code Igniter versi 2.2. Model prototype yang terbentuk diharapkan dapat membantu menganalisa tingkat kelulusan dengan cepat dengan memperhatikan atribut latihan dalam file log elearning.

UCAPAN TERIMA KASIH

Terima kasih saya sampaikan kepada Tim JSI yang telah meluangkan waktu untuk membuat template ini. Dan

memberikan arahan kepada saya untuk perbaikan paper. Tidak lupa kepada institusi STMIK Antar Bangsa yang telah memberikan waktu untuk menyelesaikan penelitian ini.

REFERENSI

- [1] Abdous, M., He, W., & Yen, C. J. (2012). Using data mining for predicting relationships between online question theme and final grade. *Educational Technology and Society*, 15(3), 77–88.
- [2] Akcapinar, G., Cosgun, E., & Altun, A. (2012). Mining Wiki Usage Data for Predicting Final Grades of Students, 1–6.
- [3] Berndtsson, M., Hansson, J., Olsson, B., & Lundell, B. (2010). *Thesis Guide - A Guide for Students in*
- [4] Dawson, C. W. (2011). *Projects in Computing and Information Systems. Information Systems Journal* (Vol. 2). Retrieved from http://www.sentimentaltoday.net/National_Academy_Press/0321263553.Addison.Wesley.Publishing.Company.Projects.in.Computing.and.Information.Systems.A.Students.Guide.Jun.2005.pdf
- [5] Firdausiah Mansur, A. B., & Yusof, N. (2013). Social learning network analysis model to identify learning patterns using ontology clustering techniques and meaningful learning. *Computers & Education*, 63, 73–86.
- [6] Gorunescu, F. (2011). *Data mining: concepts and techniques. Chemistry &*
- [7] Grosan, C., Abraham, A., & Chis, M. (2006). Swarm Intelligence in Data Mining, 34(2006), 1–20.
- [8] Islam, a. K. M. N. (2016). E-learning system use and its outcomes: Moderating role of perceived compatibility. *Telematics and Informatics*, 33(1), 48–55.
- [9] Kumar, S. A., & Vijayalakshmi, M. N. (2013). Relevance of Data Mining Techniques in Edification Sector, 3(1), 4–6.
- [10] Larose, D. T. (2005). *Discovering Knowledge in Data. New Jersey: John Willey dan Sons, Inc.*
- [11] Li, G., You, J., & Liu, X. (2015). Support Vector Machine (SVM) based prestack AVO inversion and its applications. *Journal of Applied Geophysics*, 120, 60–68.
- [12] Maimon, O., & Lior, R. (2010). *Data Mining and Knowledge Discovery Handbook.*
- [13] Pandey, U. K., & Pal, S. (2011). Data Mining : A prediction of performer or underperformer using classification, 2(2), 5.
- [14] Pintar, R., Jereb, E., Vukovic, G., & Urh, M. (2015). Analysis of Web Sites for e-Learning in the Field of Foreign Exchange Trading. *Procedia - Social and Behavioral Sciences*, 197(February), 245–254.
- [15] Prasetyo, E. (2012). *Data Mining : Konsep dan Aplikasi Menggunakan Matlab. Indonesia: Andi Yogyakarta.*
- [16] Sembiring, S., Zarlis, M., Hartama, D., Ramliana, S., & Wani, E. (2011). Prediction of Student Academic Performance By an Application of Data Mining Techniques. *International Conference on Management and Artificial Intelligence (IPEDR)*, 6, 110–114.
- [17] Vercellis, C. (2011). *Business Intelligence: Data Mining and Optimization for Decision Making (Google eBook). Methods.*
- [18] Yin, H., Jiao, X., Chai, Y., & Fang, B. (2015). Scene classification based on single-layer SAE and SVM. *Expert Systems with Applications*, 42(7), 3368–3380.
- [19] Yin, Y., Han, D., & Cai, Z. (2011). Explore Data Zafra, A., Romero, C., & Ventura, S. (2011). Multiple instance learning for classifying students in learning management

systems. *Expert Systems with Applications*, 38(12), 15020–15031.

- [20] Zafra, A., & Ventura, S. (2012). Multi-instance genetic programming for predicting student performance in web based educational environments. *Applied Soft Computing Journal*, 12(8), 2693–2706.



Sucitra Sahara, M.Kom. Lahir di Purworejo Tahun 1988, Tahun 2012 lulus dari Program Strata Satu (S1) Program Studi Sistem Informasi STMIK Nusa Mandiri Jakarta. Tahun 2015 lulus dari Program Strata Dua (S2) Program Studi Ilmu Komputer STMIK Nusa Mandiri Jakarta. Saat ini bekerja sebagai staf pengajar di STMIK Antar Bangsa. Aktif mengikuti seminar dan menulis artikel ilmiah di Jurnal Sistem Informasi STMIK Antar Bangsa.