

Prediksi Kemampuan Akademik Mahasiswa dengan Metode Support Vector Machine

¹Bangkit Indarmawan Nugroho, ²Nugroho Adhi Santoso, ³Aang Alim Murtopo
^{1,2,3}STMIK YMI Tegal
Tegal, Indonesia

¹efbeterang@gmail.com, ²nugrohoadhisantoso29@gmail.com, ³aang.alim@gmail.com

*Penulis Korespondensi

Diajukan : 16/12/2022

Diterima : 19/12/2022

Dipublikasi : 01/01/2023

ABSTRAK

Indikator yang digunakan untuk mengukur keberhasilan pelaksanaan pendidikan di perguruan tinggi adalah kinerja/kemampuan akademik (*academic performance*) mahasiswa, yang dapat diukur dari prestasi akademik (*academic achievement*) mahasiswa yaitu berdasarkan capaian nilai (*grade*), yang di Indonesia disebut dengan IPK (Indeks Prestasi Kumulatif). Penelitian ini bertujuan untuk membuat sebuah pemodelan dengan menggunakan metode *Support Vector Machine* untuk mengklasifikasikan dan memprediksi kemampuan akademik mahasiswa berdasarkan faktor demografi, *multiple intelligence*, dan *learning style* yang ia miliki. Data yang akan digunakan sebagai *criteria variable* adalah data sekunder berupa nilai IPK mahasiswa yang diperoleh dari Sistem Informasi Akademik (Simak) AMIK YMI Tegal untuk mengidentifikasi data nilai IPK mahasiswa dari tiga angkatan yang berbeda. Proses pengumpulan data (*data collection*) dilakukan menggunakan kuesioner melibatkan 204 responden mahasiswa terdiri dari 4 angkatan yang berbeda. Dalam penelitian ini digunakan fungsi kernel polynomial ordinal-2, dan hasil yang diperoleh menghasilkan nilai akurasi yang cukup tinggi yaitu 92.5% yang menunjukkan hubungan yang signifikan antara faktor demografi, *multiple intelligence*, dan *learning style* terhadap kemampuan akademik mahasiswa yang diklasifikasikan ke dalam dua kelas, potensial dan tidak potensial. Hasil lainnya menunjukkan bahwa di antara ketiga faktor tersebut, faktor *multiple intelligence* memiliki tingkat signifikansi tertinggi terhadap kemampuan mahasiswa.

Kata Kunci: *support vector machine*, demografi, *multiple intelligence*, *learning style*, fungsi kernel

I. PENDAHULUAN

Terdapat banyak faktor yang bisa mempengaruhi kemampuan akademik mahasiswa. Penelitian yang dilakukan oleh (TL et al., 2017) bertujuan untuk meneliti mengenai faktor-faktor yang mempengaruhi prestasi akademik mahasiswa. Prestasi akademik mahasiswa diprestasi oleh dua faktor yaitu faktor internal dan faktor eksternal. Dalam penelitian ini faktor-faktor tersebut antara lain gender, jurusan sekolah, status pernikahan, bekerja, dukungan orang tua, dan keaktifan dalam organisasi. Namun secara garis besar, faktor intelektual mahasiswa menjadi faktor utama yang menentukan kemampuan akademik dibandingkan dengan tiga faktor yang lain: faktor lingkungan, faktor sekolah, dan faktor keluarga mahasiswa yang bersangkutan (Bakare, 1975). *Multiple intelligence* menjadi indikator yang mempengaruhi proses belajar, dan kecerdasan akan dipengaruhi oleh kondisi sosial ekonomi, jenis kelamin, dan tingkat pendidikan orang tua (Beceran, 2010). Faktor cara/pendekatan yang digunakan seseorang dalam kebiasaan proses belajar ternyata juga berdampak kepada kemampuan akademik (*academic performance*). Faktor

ini disebut *learning style*, dan salah satu *tools* yang digunakan untuk mengidentifikasi faktor ini adalah *VARK (Visual, Aural, Read-write, Kinesthetic) Questionnaire* (Mlambo, 2011). Hubungan *interpersonal* dan keterlibatan antara orang tua dengan siswa juga memberi dampak yang signifikan untuk mengoptimalkan kecerdasan yang dimiliki sehingga mampu memaksimalkan kinerja akademiknya (Khajehpour, 2011).

Belakangan ini, metoda *Support Vector Machine (SVM)* menyita banyak perhatian para peneliti maupun praktisi, karena performansinya yang relatif lebih baik dibanding metoda yang sudah ada, terutama untuk kasus nonlinier (Jun, 2021).

Penelitian ini bertujuan untuk membuat sebuah pemodelan dengan menggunakan metode *support vector machine* untuk mengklasifikasikan dan memprediksi kemampuan akademik mahasiswa berdasarkan faktor demografi, *multiple intelligence*, dan *learning style* yang ia miliki. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi kepada institusi perguruan tinggi dalam upaya untuk menghasilkan output lulusan yang berkualitas dengan cara melakukan identifikasi dini terhadap tingkat kecerdasan dan gaya belajar mahasiswanya.

II. STUDI LITERATUR

Penelitian Terdahulu

Telah banyak penelitian yang membahas tentang pemanfaatan SVM untuk memprediksi kemampuan akademik seseorang dengan berbagai varian algoritma pemodelan dan berbagai parameter yang berbeda. Antara lain adalah penelitian yang dilakukan oleh (Oloruntoba & Akinode, 2017) yang mencoba mempelajari hubungan antara profil akademik siswa sebelum diterima kuliah dengan prestasi akademik akhir dia sebagai mahasiswa. Sampel yang digunakan adalah data mahasiswa di salah satu Politeknik Federal di bagian barat daya Nigeria. Profil akademik pra-penerimaan yang digunakan untuk penelitian ini adalah nilai level 'O' (hasil sekolah menengah atas). Prestasi akademik ditentukan dengan Indeks Prestasi Kumulatif (IPK) mahasiswa. Penelitian ini difokuskan pada penggunaan teknik data mining untuk mengembangkan model guna memprediksi performa mahasiswa berdasarkan hasil level 'O' dan IPK mereka pada 3 semester awal. Preprocessing data dilakukan untuk menghilangkan hasil *rusted* dan *expelled*. Hasil diperoleh dengan membandingkan SVM dengan teknik ML lainnya seperti KNN, pohon keputusan, dan regresi linier, di mana hasilnya menunjukkan bahwa SVM mengungguli algoritma ML lainnya. Parameter algoritma SVM (kernel) juga disetel untuk meningkatkan akurasi dan hasil yang diperoleh menunjukkan bahwa kernel RBF dengan penalti ($C=100$) berkinerja terbaik. SVM dan RBF memberikan akurasi pelatihan tertinggi 94% dan akurasi prediksi 97% yang mengungguli teknik ML canggih lainnya seperti KNN, pohon keputusan, dll.

Penelitian lainnya adalah dari (Janan & Ghosh, 2021) yang mencoba menganalisis performa tahun kedua dari mahasiswa universitas terkemuka di Bangladesh. Penelitian ini pada dasarnya berusaha menemukan faktor-faktor yang akan berpengaruh terhadap performa mahasiswa dan setelah itu memprediksi hasil siswa berdasarkan enam variabel yaitu *creating good notes* (membuat catatan dengan bagus), *exam strategy* (strategi ujian), *class tests marks* (nilai ujian kelas), *Personal problems* (masalah pribadi), *Fear of examinations* (kekhawatiran terhadap ujian), dan *Political Involvement* (Keterlibatan politik). Pada penelitian ini SVM digunakan untuk mengklasifikasikan performa siswa. Dataset input (untuk pelatihan dan pengujian) diambil dengan menggabungkan nilai-nilai yang diperoleh dari dua survei yang dilakukan pada mahasiswa dengan memanfaatkan sistem interferensi neuro-fuzzy adaptif (ANFIS). Dengan penerapan model ini diharapkan bisa meningkatkan kemampuan prediksi kinerja siswa dan ketepatannya ketika faktor yang baru maupun peserta survey ditambahkan.

Kemudian (Maya et al., 2020) melakukan penelitian untuk memprediksi masa studi mahasiswa. Analisis prediksi menggunakan SVM digunakan dengan memanfaatkan data historis alumni Fakultas Hukum UMY tahun kelulusan 2015-2019. Penerapan SVM telah memberikan hasil akurasi, presisi, dan recall yang lebih baik. Tingkat akurasi kernel terbaik adalah kernel SVM RBF dengan nilai C optimum 10 dan nilai γ 0,4 dengan akurasi 96,00%.

Kemampuan Akademik



Kata kemampuan akademik diterjemahkan dari bahasa Inggris *academic performance* yang juga memiliki arti prestasi atau kinerja akademik. *Performance* dalam bahasa Inggris memiliki arti pertunjukan, pembuatan, daya guna, prestasi dan hasil. Dari terjemahan tersebut dapat dilihat bahwa *performance* mempunyai penekanan pada kemampuan seseorang dalam melaksanakan tugasnya.

Kemampuan akademik yaitu sejumlah kapasitas yang dimiliki seseorang dalam proses belajar baik itu berupa materi bidang studi, wawasan, berbahasa dan diskusi. Jadi bisa diartikan bahwa kemampuan akademik merupakan suatu kapasitas yang dimiliki seseorang dalam melakukan suatu pembelajaran.

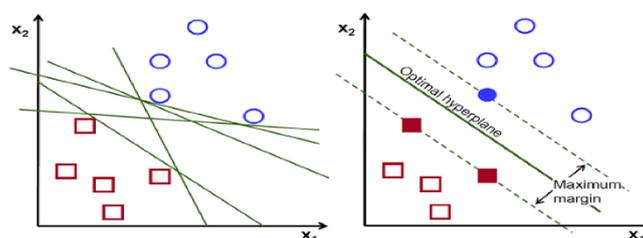
Data Mining

Data mining adalah proses untuk mendapatkan informasi dari gudang basis data yang besar. *Data mining* juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan. Istilah *data mining* kadang disebut juga *knowledge discovery*. Pekerjaan yang berkaitan dengan *data mining* dapat dibagi menjadi empat kelompok (Prasetyo, 2012), yaitu: model prediksi (*prediction modelling*), analisis kelompok (*cluster analysis*), analisis asosiasi (*association analysis*), dan deteksi anomali (*anomaly detection*)

Prosedur yang digunakan dalam proses *data mining* adalah melalui tahapan-tahapan: menentukan masalah dan rumusan hipotesis di mana pada tahap ini akan ditentukan variabel-variabel terkait masalah dan merumuskan hipotesis, dilanjutkan tahap pengumpulan data yaitu berkaitan dengan pengumpulan data dan bagaimana data dihasilkan, tahap *preprocessing* data yaitu pembersihan terhadap outlier, penanganan *missing value*, dan transformasi data. Setelah itu dilanjutkan tahap perkiraan model yaitu pemilihan teknik *data mining* sesuai tugas yang akan dilakukan, dan diakhiri dengan tahap menafsirkan model dan menarik kesimpulan.

Support Vector Machine

Konsep SVM dapat dijelaskan secara sederhana sebagai usaha mencari *hyperplane* terbaik (*maximal margin hyperplane*) yang berfungsi sebagai pemisah dua buah class pada input space (Nugroho, 2003). Seperti ilustrasi Gambar 1 yang memperlihatkan beberapa *pattern* yang merupakan anggota dari dua buah kelas: positif (dinotasikan dengan +1) dan negatif (dinotasikan dengan -1). *Pattern* yang tergabung pada kelas negatif disimbolkan dengan kotak, sedangkan *pattern* pada kelas positif, disimbolkan dengan lingkaran. Proses pembelajaran dalam problem klasifikasi diterjemahkan sebagai upaya menemukan garis (*hyperplane*) yang memisahkan antara kedua kelompok tersebut. Berbagai alternatif garis pemisah (*discrimination boundaries*) ditunjukkan pada Gambar 1(a), dan Gambar 1(b) merupakan *hyperplane* dengan margin yang paling maksimal. *Hyperplane* dengan margin yang maksimal akan memberikan generalisasi yang lebih baik pada metode klasifikasi.

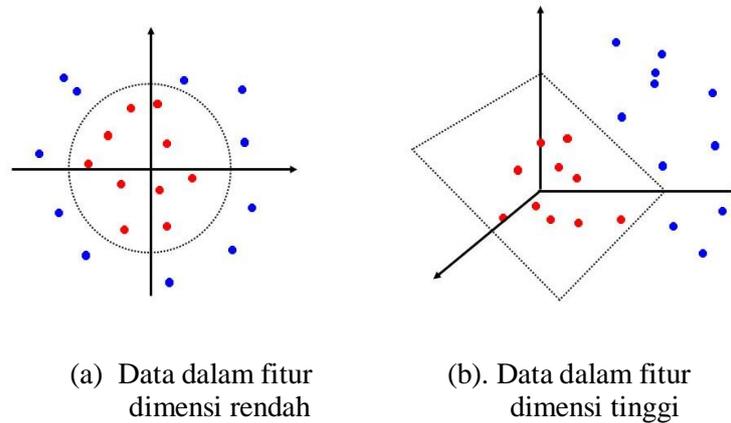


Gambar 1. Prinsip Kerja SVM: Berusaha Menemukan Hyperplane Terbaik Untuk Memisahkan Kelas Negatif dan Positif
Sumber Gambar: asnugroho.net/papers/ikcsvm.pdf

Hyperplane pemisah terbaik antara kedua class dapat ditemukan dengan mengukur margin *hyperplane* tersebut dan mencari titik maksimalnya. Margin adalah jarak antara *hyperplane* tersebut dengan data terdekat dari masing-masing class. Subset data training set yang paling dekat ini disebut sebagai *support vector*. Garis solid pada Gambar 1(b) menunjukkan *hyperplane* yang

terbaik, yaitu yang terletak tepat pada tengah-tengah kedua class, sedangkan titik kotak dan lingkaran yang berada dalam lingkaran hitam adalah support vector. Upaya mencari lokasi hyperplane optimal ini merupakan inti dari proses pembelajaran pada SVM.

Pada kehidupan nyata, hampir seluruh dataset yang akan diklasifikasikan bersifat *nonlinear*. Sedangkan SVM sebenarnya adalah *hyperplane linear* yang hanya bekerja pada data yang bisa dipisahkan secara *linear*. Untuk data yang distribusi kelasnya tidak linear biasanya digunakan pendekatan kernel pada fitur awal *dataset*. Kernel dapat didefinisikan sebagai suatu fungsi yang memetakan fitur data dari dimensi awal (rendah) ke fitur lain yang berdimensi tinggi. Pendekatan ini berbeda dengan metode klasifikasi pada umumnya yang justru mengurangi dimensi awal untuk menyederhanakan proses komputasi dan memberikan akurasi prediksi yang lebih baik (Christianini, N, and Taylor, 2000).



Gambar 2. Dimensi Data

Sumber Gambar: asnugroho.net/papers/ikcsvm.pdf

Untuk pilihan fungsi kernel yang banyak digunakan dalam aplikasi, ditunjukkan pada tabel berikut :

Tabel 1. Fungsi Kernel

Nama Kernel	Definisi Fungsi
Linear	$K(x, y) = x \cdot y$
Polinomial	$K(x, y) = (x \cdot y + c)^d$
Gaussian RBF	$K(x, y) = \exp\left(\frac{-\ x-y\ ^2}{2 \cdot \sigma^2}\right)$
Sigmoid (Tangent Hiperbolik)	$K(x, y) = \tanh(\sigma(x \cdot y) + c)$
Invers Multikuadrik	$K(x, y) = \frac{1}{\sqrt{\ x-y\ ^2 + c^2}}$

Sumber Tabel: asnugroho.net/papers/ikcsvm.pdf

Demografi

Kata demografi berasal dari bahasa Yunani yang berarti “ *Demos* ” adalah rakyat atau penduduk dan “ *Grafein* ” adalah menulis. Jadi Demografi adalah tulisan-tulisan atau karangan-karangan mengenai rakyat atau penduduk.

Tahap pengalaman pra kuliah dan karakteristik latar belakang siswa berpengaruh signifikan terhadap kemampuan akademik mahasiswa. Yang termasuk dalam tahap ini adalah faktor dan kondisi yang mempengaruhi kemungkinan bahwa siswa akan melakukan apa yang diperlukan untuk mempersiapkan dan berhasil di perguruan tinggi. Kuh et al. merinci faktor-faktor tersebut, antara lain pilihan siswa pada saat pendaftaran (*enrollment choices*), persiapan akademis (*academic preparation*) termasuk di dalamnya prestasi akademis pada tingkat pendidikan

sebelumnya, kecerdasan dan kesiapan kuliah (*aptitude and college readiness*), dukungan keluarga dan rekan (*family and peer support*), motivasi belajar, serta demografi (misalnya gender, ras dan kondisi sosial ekonomi) (Mirvald & Tománková, 2018).

Multiple Intelligence

Multiple Intelligences/kecerdasan majemuk merupakan suatu teori kecerdasan yang memperoleh banyak pengakuan akhir-akhir ini. Teori ini dicetuskan oleh Howard Gardner, psikolog dari Harvard. Kecerdasan menurut Gardner diartikan sebagai suatu kemampuan, dengan proses kelengkapannya, yang sanggup menangani kandungan masalah yang spesifik di dunia. Meskipun demikian, tidak berarti bahwa orang yang memiliki jenis kecerdasan tertentu, akan menunjukkan kemampuan tersebut dalam setiap aspek hidupnya. Dikatakan lebih lanjut bahwa setiap orang memiliki delapan jenis kecerdasan dalam tingkat yang berbeda-beda. Kedelapan jenis kecerdasan itu memiliki komponen inti dan ciri-ciri. Kehadiran ciri-ciri pada individu menentukan kadar profil kecerdasannya (Steinmayr et al., 2011).

Verbal/Linguistic Intelligence: kecerdasan ini ditunjukkan dengan kepekaan seseorang pada bunyi, struktur, makna, fungsi kata, dan bahasa.

Logical/mathematical Intelligence: kecerdasan ini ditandai dengan kepekaan pada pola-pola logis dan memiliki kemampuan mencerna pola-pola tersebut, termasuk juga numerik serta mampu mengolah alur pemikiran yang panjang.

Visual/Spatial Intelligence: kecerdasan ini ditandai dengan kepekaan mempersepsi dunia spasial-visual secara akurat dan mentransformasi persepsi awal.

Bodily/kinesthetic Intelligences: kecerdasan ini ditandai dengan kemampuan mengontrol gerak tubuh dan kemahiran mengelola objek.

Musical/Rhythmic Intelligence: kecerdasan ini ditandai dengan kemampuan menciptakan dan mengapresiasi irama pola tinada, dan warna nada; apresiasi bentuk-bentuk ekspresi musikal.

Interpersonal Intelligence: kecerdasan ini ditandai dengan kemampuan mencerna dan merespon secara tepat suasana hati, temperamen, motivasi, dan keinginan orang lain.

Intrapersonal Intelligence: kecerdasan ini ditandai dengan kemampuan memahami perasaan sendiri dan kemampuan membedakan emosi; pengetahuan tentang kekuatan dan kelemahan diri.

Naturalist Intelligence: kecerdasan ini ditandai dengan keahlian membedakan anggota-anggota suatu spesies, mengenali eksistensi spesies lain, dan memetakan hubungan antara beberapa spesies, baik secara formal maupun informal.

Learning Style

Learning style/Gaya belajar merupakan sebuah pendekatan yang menjelaskan mengenai bagaimana individu belajar atau cara yang ditempuh oleh masing-masing orang untuk berkonsentrasi pada proses dan menguasai informasi yang sulit dan baru melalui persepsi yang berbeda.

Definisi lain dikemukakan oleh Kolb yang mengatakan bahwa gaya belajar merupakan metode yang dimiliki individu untuk mendapatkan informasi, yang pada prinsipnya gaya belajar merupakan bagian integral dalam siklus belajar aktif.

Salah satu model gaya belajar yang terkenal adalah model VARK (*Visual, Aural, Read-write, Kinesthetic*). Gaya Belajar Visual lebih menitik beratkan kepada aspek visual seperti gambar, peta pikiran, memberi warna yang berbeda pada teks. Gaya belajar Aural lebih menitik beratkan pada aspek pendengaran, seperti rekaman pembelajaran, mendengarkan penjelasan dari orang lain. Gaya belajar Read/write lebih menitikberatkan pada aspek baca tulis seperti membuat rangkuman, handout, daftar istilah. Gaya belajar Kinesthetic lebih menitikberatkan pada aspek sentuhan fisik seperti praktek dan studi lapangan

III. METODE

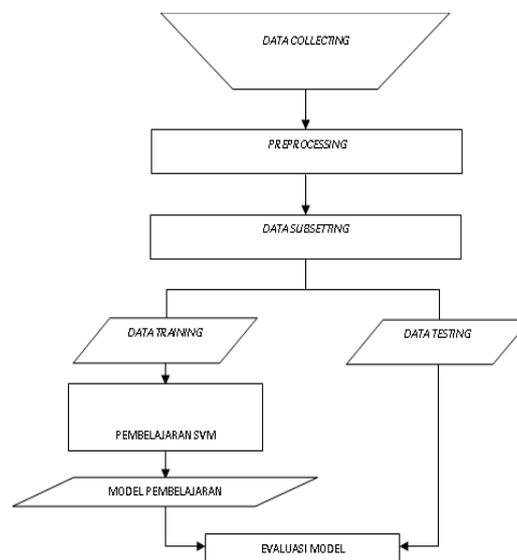
Bahan Penelitian

Data yang akan digunakan sebagai *criteria variable* adalah data sekunder berupa nilai IPK mahasiswa yang diperoleh dari Sistem Informasi Akademik (Simak) AMIK YMI Tegal untuk mengidentifikasi data nilai IPK mahasiswa dari tiga angkatan yang berbeda.

Data hasil kuisioner yang akan digunakan sebagai predictor variable, terdiri dari: (1) Data demografi mahasiswa untuk keperluan identifikasi data latar belakang mahasiswa dan untuk mengetahui pengaruh orang tua dalam proses studinya. (2) Assessment terhadap 8 aspek kecerdasan yang termasuk dalam *multiple intelligence* (kecerdasan majemuk), yaitu *linguistic intelligence* (MI 01), *logical-mathematical intelligence* (MI 02), *spatial intelligence* (MI 03), *musical intelligence* (MI 04), *bodily-kinesthetic intelligence* (MI 05), *interpersonal intelligence* (MI 06), *intrapersonal intelligence* (MI 07), dan *naturalist intelligence* (MI 08). Untuk keperluan *assessment multiple intelligence* digunakan kuisioner yang berisi materi tes talenta yang diadaptasi berdasarkan model yang bersumber dari *Learning Disabilities Resources Community*, yang diciptakan oleh Greg Gay dan J. Ivanco berupa pertanyaan yang memiliki 5 level jawaban yaitu sangat tidak setuju, tidak setuju, agak setuju, setuju, dan sangat setuju. Serta (3) *assessment* terhadap *learning style* yang diterapkan mahasiswa yang diklasifikasikan menjadi *visual learner* (V), *aural learner* (A), *read-write learner* (R), dan *kinesthetic learner* (K). Untuk *assessment learning style* mahasiswa digunakan kuisioner VARK yang diadopsi dari The VARK Questionnaire ver.7.1.

Prosedur Penelitian

Skema prosedur yang dilakukan pada penelitian ini adalah sebagai berikut :



Gambar 3. Skema Prosedur Penelitian

Identifikasi Permasalahan

Pada tahapan pertama yang dilakukan dalam penelitian berupa identifikasi permasalahan data kemahasiswaan yang terfokus kepada nilai IPK mahasiswa, yang kemudian dihubungkan dengan faktor demografi, kecerdasan dasar mahasiswa, serta gaya belajar yang diterapkan mahasiswa.

Melakukan pengumpulan data (data collection)

Data yang digunakan dalam penelitian ini adalah data mahasiswa AMIK YMI Tegal dari tiga angkatan sebanyak 204 mahasiswa.

Criteria variable berupa Nilai IPK mahasiswa diperoleh dari Simak (Sistem Informasi Akademik) AMIK YMI Tega yang direpresentasikan ke dalam 2 kelas : Potensial dan Tidak Potensial.

Predictor variable berupa: Data Demografi diperoleh dari kuisioner demografi berupa pertanyaan umur, jenis kelamin (Laki-laki/Perempuan), status marital (Menikah/Tidak Menikah), status pendidikan saat mulai kuliah (fresh graduate/tidak), domisili (bersama orang tua/tidak), pendidikan ayah dan ibu (berpendidikan tinggi [$\geq D3$]/tidak [$< D3$]), sumber biaya (dari orang tua/lainnya). Kemudian data assesment multiple intelligence diperoleh dari kuisioner yang berisi

materi tes talenta yang diadaptasi berdasarkan model yang bersumber dari Learning Disabilities Resources Community. Kuisisioner ini menghasilkan skor untuk masing-masing jenis kecerdasan. Serta data assesment learning style diperoleh dari kuisisioner VARK yang diadopsi dari The VARK Questionnaire ver.7.1, menghasilkan skor gaya belajar mahasiswa.

Melakukan representasi terhadap data yang diperoleh (preprocessing)

Representasi data dalam konsep data mining adalah upaya untuk mengkonversi/ merubah bentuk data suatu atribut menjadi bentuk yang bisa diolah oleh metode yang digunakan. Dalam SVM, nilai yang digunakan adalah nilai numerik berbasis vektor/matriks. Dataset yang digunakan memiliki atribut yang tipe nilainya *real* dan *discrete*. Agar dataset ini dapat diproses oleh SVM maka dataset tersebut harus direpresentasikan sebagai vektor dari bilangan *real*.

Selain mengkonversi nilai atribut, normalisasi / *scaling* (perubahan rentang dari nilai atribut) atribut yang bernilai bilangan real umumnya juga diperlukan agar proses perhitungan lebih mudah dan atribut yang rentang nilainya besar tidak mendominasi atribut yang rentang nilainya lebih kecil. Normalisasi akan memberikan hasil yang lebih baik. Rentang nilai atribut yang dianjurkan untuk digunakan adalah [0,1] atau [-1,+1] (Hsu, et al., 2004). SVM menyediakan kemampuan untuk melakukan normalisasi data secara otomatis. Dalam penelitian ini, rentang nilai yang digunakan adalah [-1,+1]. Langkah representasi data berupa tahapan:

Yang pertama, membagi data (*subset data*) menjadi data *training* dan data *testing*. Dataset sebanyak 204 record dibagi menjadi dua, 80% digunakan sebagai data training dan 20% digunakan sebagai data testing. Pemilihan data training dan testing dilakukan secara acak menggunakan fungsi holdout yang disediakan oleh Matlab.

Yang kedua, Implementasi Model Pembelajaran SVM terdiri dari (a) Estimasi Parameter, yaitu sebelum pelatihan dilakukan, terlebih dahulu harus ditentukan nilai parameter Alpha(α) dan Bias(b) sebagai parameter user yang digunakan untuk mengukur tingkat akurasi klasifikasi dari metode SVM, beserta fungsi kernel dan parameternya untuk pembelajaran *supervised SVM* dan kemudian (b) Pelatihan data (*data training*), karena dataset yang digunakan merupakan data nonlinier, maka dibutuhkan kernel trick. Dalam penelitian ini akan digunakan fungsi Polynomial (order 2) kernel. Penulis menggunakan metode Polinomial Kernel dengan orde dua ($p = 2$) atau biasa disebut metode Lokal Linier yang memiliki kelebihan yaitu mengestimasi fungsi di setiap titik sehingga model yang didapatkan lebih mendekati pola data yang sesungguhnya dan estimator ini tidak memerlukan data dalam jumlah banyak untuk estimasi model (Nottingham dan Cook, 2001)

Dan berikutnya yang ketiga, Evaluasi Model. Pada tahap evaluasi bertujuan untuk mengetahui tingkat akurasi dari hasil penggunaan metode SVM (*Support Vector Machine*) dengan cara menghitung jumlah data uji yang kelasnya diprediksi secara benar. Adapun cara mengukur kinerja dari sistem akan dilakukan evaluasi dengan menggunakan *confusion matrix*. *Confusion matrix* adalah matriks yang berisi data aktual dan data prediksi yang telah diklasifikasikan oleh sistem.

Tabel 2. Confusion Matrix

		Actual	
		Negative	Positive
Predictive	Negative	TN	FN
	Positive	FP	TP

True Positive (TP) adalah kasus positif yang diklasifikasikan secara benar dan hasilnya positif. *True Negative* (TN) adalah kasus negatif yang diklasifikasikan secara benar dan hasilnya negatif. *False Positive* (FP) adalah kasus dengan kelas negatif yang diklasifikasikan pada kelas positif. *False Negative* (FN) adalah kasus dengan kelas positif yang diklasifikasikan pada kelas negatif. Dari *confusion tabel* di atas kemudian dapat dihitung tingkat akurasi dan error dari klasifikasi dengan rumus berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$Error = \frac{FP+FN}{TP+TN+FP+FN} \times 100\%$$

Accuracy adalah nilai yang menunjukkan tingkat prosentase pengujian data yang menghasilkan nilai yang benar/sesuai, sedangkan error merupakan nilai kebalikannya yaitu tingkat prosentase pengujian data yang menghasilkan nilai yang salah/tidak sesuai.

IV. HASIL DAN PEMBAHASAN

Pengumpulan Data

Proses pengumpulan data (*data collection*) dilakukan menggunakan kuesioner melibatkan 204 responden mahasiswa terdiri dari 4 angkatan yang berbeda. Pertanyaan kuisisioner terdiri dari : Pertanyaan seputar faktor demografi meliputi: umur mahasiswa, jenis kelamin (laki-laki/perempuan), status marital (menikah/tidak menikah), status pendidikan pada saat mendaftar (fresh graduate/tidak), status domisili (bersama orang tua/tidak), pendidikan ayah ($\geq D3$ / $< D3$), pendidikan ibu ($\geq D3$ / $< D3$), dan Sumber biaya (dari orang tua/selain dari orang tua). Pertanyaan untuk mendeteksi kecerdasan majemuk (*multiple intelligence/MI*) mahasiswa. Terdiri dari 80 pertanyaan yang diadopsi dari *Learning Disabilities Resources Community*, Greg Gay dan J. Ivanco yang telah dimodifikasi. Setiap pertanyaan memiliki 5 pilihan jawaban yang masing-masing memiliki bobot skor : sangat tidak setuju (bobot skor=1), tidak setuju (bobot skor=2), agak setuju (bobot skor=3), setuju (bobot skor=4), sangat setuju (bobot skor=5). Soal telah diacak sedemikian rupa, dan penjumlahan bobot skor akan menunjukkan dominasi kecerdasan dasar yang dimiliki mahasiswa. Jumlah skor tertinggi menunjukkan kecerdasan paling menonjol yang dimiliki. Daftar nomor soal menunjuk kepada jenis kecerdasan tertentu ditunjukkan dalam tabel berikut:

Tabel 3. Tabel hubungan nomor soal kuisisioner *multiple intelligence* dengan jenis kecerdasan

Nomor Soal	Jenis Kecerdasan	Kode Kecerdasan
1,9,17,25,33,41,49,57,65,73	Kecerdasan Verbal/Linguistik	MI 01
2,10,18,26,34,42,50,58,66,74	Kecerdasan Logis/Matematis	MI 02
3,11,19,27,35,43,51,59,67,75	Kecerdasan Visual/Spasial	MI 03
4,12,20,28,36,44,52,60,68,76	Kecerdasan Kinestetik	MI 04
5,13,21,29,37,45,53,61,69,77	Kecerdasan Musikal	MI 05
6,14,22,30,38,46,54,62,70,78	Kecerdasan Interpersonal	MI 06
7,15,23,31,39,47,55,63,71,79	Kecerdasan Intrapersonal	MI 07
8,16,24,32,40,48,56,64,72,80	Kecerdasan Naturalis	MI 08

Pertanyaan untuk mendeteksi gaya belajar (*learning style/LS*) mahasiswa. Kuisisioner gaya belajar ini berupa kuisisioner VARK yang diadopsi dari The VARK Questionnaire ver.7.1 yang terdiri dari 16 pertanyaan dengan 4 pilihan jawaban (a,b,c,d). Setiap pilihan jawaban (boleh lebih dari satu jawaban) dalam satu soal akan menambah skor komponen VARK (Visual, Aural, Read/write, Kinestetik). Skor tertinggi menunjukkan gaya belajar mahasiswa.

Tabel 4. Penentuan kategori learning style berdasar jawaban pada setiap soal kuisisioner VARK

Pertanyaan	Jawaban a	Jawaban b	Jawaban c	Jawaban d
1	K	A	R	V
2	V	A	R	K
3	K	V	R	A
4	K	A	V	R
5	A	V	K	R
6	K	R	V	A
7	K	A	V	R
8	R	K	A	V
9	R	A	K	V

10	K	V	R	A
11	V	R	A	K
12	A	R	V	K
13	K	A	R	V
14	K	R	A	V
15	K	A	R	V
16	V	A	R	K

Data hasil jawaban dari setiap mahasiswa diolah dengan aplikasi excel. Data ini kemudian akan menjadi dataset penelitian dan menjadi dasar untuk pemodelan dengan metode SVM menggunakan aplikasi Matlab.

Representasi Data

Dataset yang digunakan memiliki atribut yang tipe nilainya *real* dan *discrete*. Untuk data *discrete* yang berupa kategori, data perlu direpresentasikan menjadi bentuk yang dapat diolah oleh metode SVM. SVM memisahkan data dalam 2 bentuk rentang nilai $[0,1]$ atau $[-1,1]$.

Normalisasi Data

Dalam penelitian ini rentang yang digunakan adalah $[-1,1]$. Proses normalisasi dilakukan secara otomatis oleh SVM dengan melakukan pengaturan rentang nilai pada aplikasi matlab.

Subsetting Data

Dataset sebanyak 204 record dibagi menjadi dua, 80% digunakan sebagai data training dan 20% digunakan sebagai data testing. Pemilihan data training dan testing dilakukan secara acak menggunakan fungsi holdout yang disediakan oleh Matlab. Berdasarkan fungsi tersebut, data training terdiri dari 164 record dan data testing terdiri dari 40 record.

Implementasi Model

Karena dataset yang digunakan memiliki karakteristik data nonlinier dan memiliki dua kelas pemisah yaitu potensial dan tidak potensial, maka penelitian ini akan menggunakan polynomial order-2 kernel yang menghasilkan parameter α (Alpha) dan b (Bias). Aplikasi yang digunakan adalah Matlab R.2010a untuk pembangunan model SVM.

Evaluasi Model

Hasil akhir yang akan dicapai adalah mengetahui tingkat akurasi prediksi berdasarkan hasil pembelajaran yang dilakukan oleh SVM.

Penelitian ini mencoba memetakan tingkat akurasi berdasarkan pembelajaran terhadap data training yang melibatkan seluruh atribut, atribut demografi, atribut skor kecerdasan majemuk, dan atribut skor gaya belajar.

Berdasarkan hasil pembelajaran terhadap data training yang kemudian diujikan kepada data testing, diperoleh hasil sebagai berikut :

Tabel 5. Tingkat akurasi berdasarkan 4 kriteria atribut

Atribut data	Confusion				Accuracy
	TP	FP	TN	FN	
Semua (20)	35	2	2	1	0.9250 (92.5%)
Demografi (8)	35	2	0	3	0.8750 (87.5%)
MI (8)	36	1	2	1	0.9500 (95%)
LS (4)	35	2	0	3	0.8750 (.5%)

Berdasarkan hasil penelitian, tingkat akurasi yang diperoleh dengan semua atribut data (sebanyak 20 atribut yang terdiri dari 8 atribut demografi, 8 atribut skor kecerdasan majemuk, dan 4 atribut skor gaya belajar) adalah 92.5%.

Namun ketika Atribut data dipecah, terdapat kondisi menarik, dimana tingkat akurasi yang diperoleh ketika hanya menggunakan 8 atribut skor kecerdasan majemuk ternyata menghasilkan

nilai yang lebih tinggi yaitu 95%. Dapat diartikan bahwa kecerdasan majemuk mahasiswa ternyata memberikan pengaruh yang signifikan terhadap keberhasilan studi mahasiswa. Dan apabila dilihat dari skor confusion ternyata mampu memprediksi secara baik mahasiswa yang potensial maupun yang tidak potensial.

Sedangkan dua kriteria atribut yang lain, yaitu 8 atribut demografi dan 4 atribut skor gaya belajar, ternyata menghasilkan tingkat akurasi yang lebih rendah yaitu masing-masing hanya 87.5%. Dapat diartikan bahwa kedua kriteria atribut ini ternyata tidak memiliki pengaruh yang signifikan terhadap keberhasilan studi mahasiswa. Dan apabila dilihat dari skor confusion ternyata kedua kriteria atribut ini hanya mampu memprediksi secara baik mahasiswa yang potensial, namun tidak mampu memprediksi secara baik mahasiswa yang tidak potensial karena skor TN (*true negative*) selalu 0.

Berdasarkan kondisi tersebut, penelitian ini kemudian mencoba lebih spesifik mengukur kecerdasan majemuk yang memiliki tingkat pengaruh paling tinggi terhadap prediksi kemampuan belajar mahasiswa. Hasil yang diperoleh adalah:

Tabel 4.4. Hasil Pengukuran Tingkat Akurasi Berdasarkan Atribut Skor Kecerdasan Majemuk

Atribut data	B	Confusion				Accuracy
		TP	FP	TN	FN	
MI 01	-0.7625	37	0	0	3	0.9250 (92.5%)
MI 02	-1.8182	37	0	2	1	0.9750 (97.5%)
MI 03	-0.9541	37	0	0	3	0.9250 (92.5%)
MI 04	-0.9853	37	0	1	2	0.9500 (95%)
MI 05	-0.8949	37	0	0	3	0.9250 (92.5%)
MI 06	-0.9083	35	2	1	2	0.9000 (90%)
MI 07	-1.0486	36	1	0	3	0.9000 (90%)
MI 08	-0.8402	37	0	0	3	0.9250 (92.5%)

Tabel di atas menunjukkan bahwa ternyata atribut MI 02 (Kecerdasan Logis/Matematis) memiliki tingkat akurasi tertinggi yaitu 97.5%, yang dapat diartikan bahwa mahasiswa yang memiliki kecerdasan logis/matematis lebih dominan dibanding kecerdasan yang lain, bisa diprediksi cenderung memiliki kemampuan akademik yang baik/potensial berhasil.

Berdasarkan hasil penelitian di atas dan guna penelitian selanjutnya, peneliti merekomendasikan agar konsep kuisisioner demografi, kecerdasan majemuk, dan gaya belajar ini bisa diterapkan sebagai salah satu materi uji pada saat seleksi calon mahasiswa sehingga sudah bisa memetakan tingkat kemampuan akademik mahasiswanya sejak awal. Harapan lebih jauh adalah dengan terpetakannya kemampuan mahasiswa, institusi perguruan tinggi bisa merencanakan langkah strategis dalam proses pendidikan yang akan diterapkan agar tetap bisa mengoptimalkan kemampuan mahasiswa.

Kemudian peneliti juga merekomendasikan agar dilakukan penelitian yang lebih mendalam dengan metode yang berbeda mengenai kemampuan akademis/*academic performance* menggunakan variabel yang lebih kompleks sehingga bisa menghasilkan tingkat akurasi yang lebih baik, dengan sasaran akhir adalah tercapainya keseimbangan antara pembelajaran yang diterapkan oleh institusi pendidikan dengan kemampuan akademik mahasiswa.

V. KESIMPULAN

Dalam penelitian ini, prediksi kemampuan mahasiswa dilakukan dengan pemodelan menggunakan metode *support vector machine* berdasarkan faktor demografi, *multiple intelligence*, dan *learning style* yang ia miliki. Data yang akan digunakan sebagai *criteria variable* adalah data sekunder berupa nilai IPK mahasiswa yang diperoleh dari Sistem Informasi Akademik (Simak) AMIK YMI Tegal untuk mengidentifikasi data nilai IPK mahasiswa dari tiga angkatan yang berbeda. Proses pengumpulan data (*data collection*) dilakukan menggunakan kuesioner melibatkan 204 responden mahasiswa terdiri dari 4 angkatan yang berbeda.

Dalam penelitian ini digunakan fungsi kernel polynomial ordinal-2, dan hasil yang diperoleh menghasilkan nilai akurasi yang cukup tinggi yaitu 92.5% yang menunjukkan hubungan yang signifikan antara faktor demografi, *multiple intelligence*, dan *learning style* terhadap kemampuan akademik mahasiswa yang diklasifikasikan ke dalam dua kelas, potensial dan tidak potensial. Hasil lainnya menunjukkan bahwa di antara ketiga faktor tersebut, faktor *multiple intelligence* memiliki tingkat signifikansi tertinggi terhadap kemampuan mahasiswa yaitu memiliki tingkat akurasi sampai dengan 95%. Dan di antara 8 jenis *multiple intelligence* yang memiliki pengaruh tertinggi terhadap keberhasilan studi mahasiswa AMIK YMI Tegal adalah *logical-mathematical intelligence* (MI 02) dengan tingkat akurasi mencapai 97.5%

VI. REFERENSI

- Bakare, C. G. M. (1975). Some psychological correlates of academic success and failure. *African Journal of Educational Research*, 2(1), 11–22.
- Beceren, B. Ö. (2010). Determining multiple intelligences pre-school children (4-6 age) in learning process. *Procedia - Social and Behavioral Sciences*, 2(2), 2473–2480. <https://doi.org/10.1016/j.sbspro.2010.03.356>
- Christianini, N, and Taylor, J. S. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods* (1st editio). Cambridge University Press. <https://doi.org/https://doi.org/10.1017/CBO9780511801389.012>
- Janan, F., & Ghosh, S. K. (2021). Prediction of student's performance using support vector machine classifier. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, 7078–7088.
- Jun, Z. (2021). The Development and Application of Support Vector Machine. *Journal of Physics: Conference Series*, 1748(5). <https://doi.org/10.1088/1742-6596/1748/5/052006>
- Khajehpour, M. (2011). Relationship between emotional intelligence, parental involvement and academic performance of high school students. *Procedia - Social and Behavioral Sciences*, 15, 1081–1086. <https://doi.org/10.1016/j.sbspro.2011.03.242>
- Maya, C., Audilla, P., Riyadi, S., Yogyakarta, U. M., & Brawijaya, J. (2020). *Prediction of Student Study Period Based on Admission Pathways Using Support Vector Machine Algorithm*. 1(4), 155–160. <https://doi.org/https://doi.org/10.18196/eist.v1i4.16598>
- Mirvald, M., & Tománková, M. (2018). the Determinant of Success in Basic Economics Courses Taught By the Department of Economics At the University of Economics in Prague. *International Journal of Economic Sciences*, VII(1), 82–90. <https://doi.org/10.20472/es.2018.7.1.005>
- Mlambo, V. (2011). An analysis of some factors affecting student academic performance in an introductory biochemistry course at the University of the West Indies. *Caribbean Teaching Scholar*, 1(2), 79–92.
- Nugroho, A. S. (2003). Support Vector Machine Teori dan Aplikasinya dalam Bioinformatika. *Kuliah Umum IlmuKomputer.Com*. IlmuKomputer.com
- Oloruntoba, S. A., & Akinode, J. L. (2017). STUDENT ACADEMIC PERFORMANCE PREDICTION USING SUPPORT VECTOR MACHINE. *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY*, 6(12), December. <https://doi.org/10.5281/zenodo.1130905>
- Prasetyo, E. (2012). Data mining konsep dan aplikasi menggunakan satscan. In *CV Andi Offset* (1st ed.). CV Andi Offset. <https://elibrary.bsi.ac.id/readbook/200350/data-mining-konsep->

dan-aplikasi-menggunakan-matlab

Steinmayr, R., Bipp, T., & Spinath, B. (2011). Goal orientations predict academic performance beyond intelligence and personality. *Learning and Individual Differences, 21*(2), 196–200. <https://doi.org/10.1016/j.lindif.2010.11.026>

TL, D. I., Widowati, A. I., & Surjawati, S. (2017). Faktor-Faktor Yang Mempengaruhi Prestasi Akademik : Studi Kasus Pada Mahasiswa Program Studi Akuntansi Universitas Semarang. *Jurnal Dinamika Sosial Budaya, 18*(1), 39. <https://doi.org/10.26623/jdsb.v18i1.557>