

Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Menggunakan Metode Support Vector Machine

Mochamad Fadillah Abdullah*, Kusri*, M Rudyanto Arief*

Magister Teknik Informatika, Universitas Amikom Yogyakarta

Correspondence Author: fadil.abdullah97pps@gmail.com

Info Artikel :	ABSTRACT
<p>Sejarah Artikel : Menerima : 17 Februari 2021 Revisi : 31 Agustus 2021 Diterima : 30 September 2021 Online : 05 Januari 2021</p> <p>Keyword : graduation, data mining, klasifikasi, SVM, boosting</p>	<p><i>Student graduation is one that must be considered because it is included in the Internal Quality Assurance Standards of a university. The Faculty of Information Technology is one of the faculties at the KH A Wahab Hasbullah University Jombang. For graduation, there are standards that will be achieved by the faculty, namely a study time of 4 years and a minimum GPA of 3.00. To be able to achieve the quality of graduation, it is necessary to predict the graduation rate with the standards that have been set for students who are still carrying out studies so that anticipation can be made from the start so that they can overcome problems in the academic field. To predict the graduation rate and standard GPA used data mining methods with classification functions. The classification method used is the SVM method The classification method consists of predictor variables and one target variable. The predictor variables consisted of gender, achievement index from semester 1 to 4 and semester credits from semester 1 to semester 4. The device used to process data was using RapidMiner. The results of the SVM algorithm and using the boosting method are used to predict student graduation grades and times and GPA 3 with accuracy (86.36%), error (13.64%), and AUC (0.876). And it can be concluded that the SVM algorithm and the boosting method can perform the classification well.</i></p>
	INTISARI
<p>Kata Kunci : kelulusan, data mining, klasifikasi, SVM, boosting</p>	<p><i>Kelulusan mahasiswa merupakan salah satu yang harus diperhatikan karena masuk dalam Standar Penjaminan Mutu Internal suatu perguruan tinggi. Fakultas Teknologi Informasi merupakan salah satu fakultas yang di universitas KH A Wahab Hasbullah Jombang. Untuk kelulusan terdapat standar yang akan dicapai oleh fakultas tersebut yaitu waktu studi selama 4 tahun dan IPK minimal 3,00. Untuk dapat mencapai mutu kelulusan tersebut dibutuhkan suatu prediksi tingkat kelulusan dengan standar yang telah ditetapkan untuk mahasiswa yang masih menjalankan studi sehingga dapat dilakukanantisipasi dari awal sehingga dapat menanggulangi terjadinya permasalahan dalam bidang akademik. Untuk memprediksi tingkat kelulusan dan IPK standar tersebut digunakan metode data mining dengan fungsi klasifikasi. Metode klasifikasi yang digunakan menggunakan metode SVM. Metode klasifikasi terdiri dari variabel-variabel prediktor dan satu variabel target. Variabel-variabel prediktor terdiri dari jenis kelamin, indeks prestasi dari semester 1 sampai 4 dan sks semester 1 sampai semester 4. Perangkat yang digunakan untuk mengolah data menggunakan</i></p>

	<i>RapidMiner. Hasil dari algoritma SVM dan menggunakan metode boosting untuk prediksi nilai dan waktu kelulusan mahasiswa dan IPK ≥ 3 dengan nilai accuracy (86,36%), error (13,64%), dan AUC (0,876). Dan dapat disimpulkan bahwa algoritma SVM dan metode boosting dapat melakukan pengklasifikasian dengan baik.</i>
--	--

1. PENDAHULUAN

Kelulusan mahasiswa merupakan salah satu hal yang harus diperhatikan karena termasuk keuntungan kedua belah pihak antara mahasiswa dan pihak universitas. Pihak mahasiswa akan mendapatkan pekerjaan dengan mudah karena perusahaan cenderung mencari yang baru lulus. Sedangkan dari pihak universitas karena seiring tepatnya waktu kelulusan mahasiswa hal itu akan membantu memajukan kualitas universitas tersebut seperti peningkatan akreditasi, dengan begitu suatu perguruan tinggi dapat mencapai standard penjaminan mutu internal (SPMI). Di Universitas KH.A Wahab Habullah Fakultas Teknologi Informasi (FTI) telah ditetapkan untuk standard kelulusan tepat waktu yaitu maksimal delapan semester dengan IPK minimal 3,00.

Dari data mahasiswa baru dan lulusan yang ada pada data database yang tersimpan di kampus pada tahun 2013 sampai 2019, terlihat tidak stabil karena mengalami penurunan disetiap tahunnya seperti contoh pada database jumlah mahasiswa baru FTI ditahun 2013 meningkat dengan jumlah 348 mahasiswa masuk, tetapi pada akhir ketepatan waktu kelulusan hanya ada 134 mahasiswa yang lulus tepat waktu dengan IPK minimal 3,00, Tahun 2014 mengalami penurunan, jumlah mahasiswa baru FTI hanya 98 mahasiswa dan pada ketepatan waktu kelulusan hanya berjumlah 52 mahasiswa, Kemudian pada angkatan tahun 2015 mahasiswa baru berjumlah 106 mahasiswa masuk dan data mahasiswa lulus tepat waktu hanya berjumlah 58 mahasiswa keluar. Ketidak stabilan jumlah pada saat mahasiswa masuk dengan jumlah mahasiswa keluar ini disebabkan dari beberapa faktor internal dan faktor eksternal yang dialami mahasiswa, karena mahasiswa tidak selalu dapat menuntaskan studinya selama waktu normal yang telah ditentukan. Sehingga diperlukan penerapan prediksi kelulusan mahasiswa yang dapat mengklasifikasikan data prediksi kelulusan secara tepat berdasarkan parameter-parameter yang telah ditentukan, yaitu menggunakan data mining, dimana akan dicari pola yang terdapat pada database lulusan untuk memprediksi waktu kelulusan empat tahun dengan nilai IPK 3,00.

Tujuan penelitian ini adalah prediksi nilai dan waktu kelulusan menggunakan Metode SVM, Dengan menggunakan Metode SVM ini dapat memberikan keuntungan dalam membuat data menjadi optimal, mendapatkan akurasi terbaik dan sebagai masukan untuk mengantisipasi mahasiswa-mahasiswa yang tidak sesuai dengan target standar lulusan.

2. LANDASAN TEORI

2.1 Prediksi

Digunakan untuk memperkirakan atau forecasting suatu kejadian sebelum kejadian-kejadian atau peristiwa tertentu terjadi. Misalnya, bagaimana Badan Metereologi Dan Geofisika (BMKG) memperkirakan tanggal tertentu bagaimana cuacanya apakah hujan, panas dan lain sebagainya. Metode yang sering digunakan salah satunya adalah Roug set.

Data mining juga sama hanya dengan konsep neural network mengandung 2 (dua) pengelompokan yaitu:

1. *Supervised learning* merupakan pembelajaran menggunakan guru dan biasanya ditandai dengan adanya class/label/target pada himpunan data. Adapun metode-metode yang digunakan bersifat *Supervised learning* seperti metode prediksi dan klasifikasi algoritma SVM *support vector machine*, metode *roug set* dan lain-lain.
2. *Unsupervised learning* merupakan pembelajaran tanpa menggunakan guru dan biasanya ditandai pada himpunan datanya dan tidak memiliki atribut keputusan atau class/label/target. Metode-metode yang bersifat *Unsupervised learning* meliputi metode *estimasi*, *clustering*, *asosiasi*, *regresi linier*, *analytical clustering* dan lain-lain.

2.2 Algoritma SVM

Algoritma SVM (Support Vector Machine) adalah metode learning supervised dimana dalam tahap training kita ingin menemukan dua parameter diantara garis dan bias b . setelah ditemuakn parameter yang optimal melalui optimasi pemrograman kuadrat (quadratic programming). Jika ukurannya problem cukup besar, solver program akuadrat biasanya lambar (Syamsiah dan Agus Darmawan, 2020).

Karakteristik dari Support Vector Machine adalah Support Vector Machine adalah linier classifier (Suhardjono, Wijaya, Ganda, Hamid, Abdul., 2019). SVM merupakan salah satu metode terbaik yang bisa dipakai dalam permasalahan klasifikasi (Pratama, Arif, Wihandika, C, R, Ratnawati, E, D., 2018). Metode SVM adalah suatu Teknik lama. Teknik lama ini pada tahun 1995, dapat dilakukan pemecahan masalah prediksi baik dalam kasus klasifikasi ataupun regresi yang sangat populer saat itu. Support vector machine (SVM) adalah algoritma pembelajaran mesin yang banyak digunakan untuk masalah klasifikasi dan prediksi (Pang, Yulie, Judd, Nicolas, Dkk., 2017).

Support Vector Machine (SVM) merupakan metode salah satu metode machine learning yang masuk dalam kategori supervised learning (Miftrahurrohmah, dkk 2019). Metode SVM berada dalam satu kelas dengan ANN dalam hal fungsi dan kondisi permasalahan yang bisa diselesaikan keduanya masuk dalam kelas supervised (Syamsiah dan Agus Darmawan, 2020).

2.3 RapidMiner

RapidMiner merupakan perangkat lunak yang bersifat terbuka (open source). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi. RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. RapidMiner ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi.

Rapid Miner merupakan perangkat lunak yang dibuat oleh Dr. Markus Hofmann dari Institute of Technology Blanchardstown dan Raif Klinkenberg dari rapid-i.com dengan tampilan GUI (Graphical User Interface) sehingga memudahkan pengguna dalam menggunakan perangkat lunak ini (Srisulistiowati Dwi Budi, Dkk., 2021).

RapidMiner adalah aplikasi data mining berbasis open-source yang terkemuka dan ternama. Didalamnya terdapat aplikasi yang berdiri sendiri untuk analisis data dan sebagai mesin data mining seperti untuk loading data, transformasi data, pemodelan data, dan metode visualisasi data (Nofitri Rika, Irawati Novica, 2019).

RapidMiner sebelumnya bernama YALE (Yet Another Learning Environment), dimana versi awalnya mulai dikembangkan pada tahun 2001 oleh RalfKlinkenberg, Ingo Mierswa, dan Simon Fischer di Artificial Intelligence Unit dari University of Dortmund. RapidMiner didistribusikan di bawah lisensi AGPL (GNU Affero General Public License) versi 3. Hingga saat ini telah ribuan aplikasi yang dikembangkan menggunakan RapidMiner di lebih dari 40 negara. RapidMiner sebagai software open source untuk data mining tidak perlu diragukan lagi karena software ini sudah terkemuka di dunia. RapidMiner menempati peringkat pertama sebagai Software data mining pada polling oleh KDnuggets, sebuah portal data-mining pada 2010-2011 (Ihaka dan Gentleman 1996).

2.4 Metode Boosting (Adaboost)

Boosting adalah metode ensemble yang digunakan untuk meningkatkan akurasi dari model klasifikasi dengan cara membangkitkan kombinasi dari suatu model (Luluk Susiana, lut Tri Utami, dan Juanidi 2019). Adaboost metode pembelajaran ensemble yang dapat mengurangi varian, hal ini terjadi karena efek bias rata-rata ensemble untuk mengurangi varian dari satu set pengkalsifikasian (Bisri Ahmad, Wahono Satrio romi, 2015).

Berbeda halnya dengan bagging dan random forest yang mendapatkan hasil prediksi dari proses bootstrap, boosting mengacu pada kumpulan algoritma yang dapat mengkombinasikan pengklasifikasian lemah menjadi pengklasifikasian yang kuat sehingga mendapatkan tingkat keakuratan yang tinggi dalam klasifikasinya. Boosting juga di kenal dengan sebutan Adaboost (Adaptive Boosting).

2.5 Metode CRISP-DM

CRISP-DM merupakan metodologi yang digunakan dalam mengembangkan data mining adalah CRISP-DM, CRISP-DM adalah metode yang dibentuk oleh komisi Eropa pada tahun 1996 yang menerapkan standar dalam proses data mining. Dalam CRISP-DM terdapat enam fase yang akan dilakukan dalam penelitian pengembangan data mining sebagai berikut:

1. Business Understanding
Tahap pertama adalah memahami tujuan dan kebutuhan dari sudut pandang bisnis, kemudian menterjemahkan pengetahuan ini ke dalam pendefinisian masalah pada data mining.
2. Data Understanding
Tahap ini dimulai dengan pengumpulan data yang kemudian akan dilanjutkan dengan proses untuk mendapatkan pemahaman yang mendalam tentang data, mengidentifikasi masalah kualitas data, atau untuk mendeteksi adanya bagian yang menarik dari data yang dapat digunakan untuk hipotesa untuk informasi yang tersembunyi.
3. Data Preparation
Tahap ini meliputi semua kegiatan untuk membangun dataset akhir (data yang akan diproses pada tahap pemodelan) dari data mentah.
4. Modelling
Dalam tahap ini akan dilakukan pemilihan dan penerapan berbagai teknik pemodelan dan beberapa parameternya akan disesuaikan untuk mendapatkan nilai yang optimal.
5. Evaluation
Pada tahap ini, model sudah terbentuk dan diharapkan memiliki kualitas baik jika dilihat dari sudut pandang analisa data.
6. Deployment
Pada tahap ini, pengetahuan atau informasi yang telah diperoleh akan diatur dan dipresentasikan dalam bentuk khusus sehingga dapat digunakan oleh pengguna.

3. METODE PENELITIAN

3.1 Desain Penelitian

Metode yang dilakukan dalam penelitian ini menggunakan metode CRISP-DM. Data yang digunakan untuk penelitian ini diperoleh dari database PDDIKTI Admin terpadu Universitas KH A Wahab Hasbullah Jombang Fakultas Teknologi Informasi untuk penggalan kekayaan terhadap data tersebut dan dimanfaatkan dalam menentukan prediksi nilai dan waktu kelulusan mahasiswa menggunakan data mining dan menggali data tersebut menggunakan metode klasifikasi dengan algoritma SVM. Kemudian data tersebut diolah dengan RapidMiner untuk mengimplementasikan algoritma SVM pada program data mining.

3.2 Pengumpulan Data

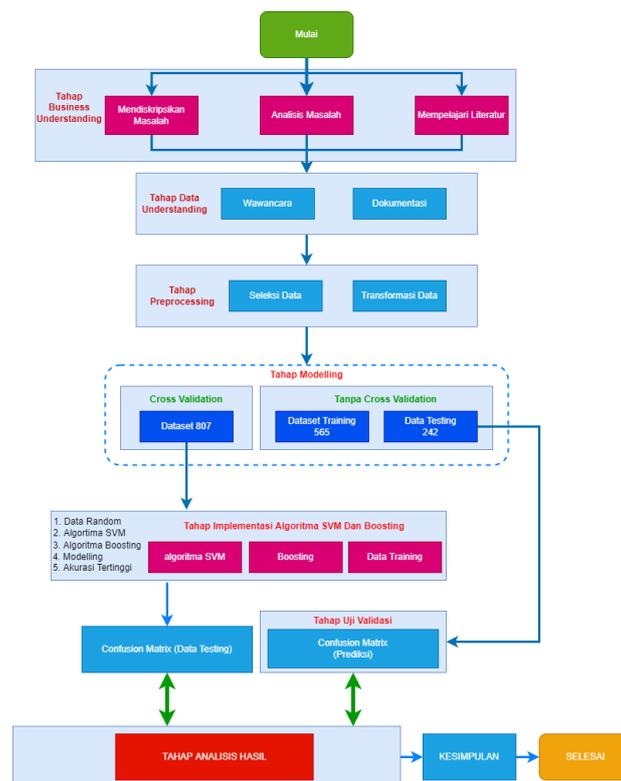
Dalam penelitian ini menggunakan observasi dan wawancara. Tempat penelitian adalah Universitas KH A Wahab Hasbullah Jombang. Pengambilan data pada bulan November 2020 secara observasi melakukan pengamatan secara langsung terhadap sistem administrasi data yang ada dan telah diterapkan di Universitas KH A Wahab Hasbullah Jombang untuk memperoleh data valid yang dibutuhkan membantu menyelesaikan penelitian ini. di admin Terpadu dan Wawancara Melakukan wawancara secara langsung dengan pihak terkait dibidang masing-masing mengenai mekanisme di Universitas KH A Wahab Hasbullah Jombang. Untuk memprediksi nilai dan waktu kelulusan mahasiswa Universitas KH A Wahab Hasbullah Jombang. Dengan demikian mereka dapat menghitung jumlah *class* LULUS dan TIDAK LULUS.

3.3 Data Preprocessing

Data penelitian ini dilakukan 2 teknik *preprocessing*, yaitu seleksi data dan transformasi data. Seleksi data dilakukan secara manual dengan kriteria atribut yang dipilih meliputi hal-hal yang bersifat akademis dan erat hubungan dengan nilai dan waktu kelulusan mahasiswa. Transformasi data dilakukan untuk memperbaiki data-data yang bernilai terlalu Panjang dengan menyederhanakan nilai atribut sehingga memudahkan nantinya dalam pembuatan model *support vector machine* pada RapidMiner.

3.4 Alur penelitian

Alur penelitian yang digunakan dalam penerapan algoritma SVM dalam memprediksi nilai dan waktu kelulusan mahasiswa menggunakan metode SVM, dalam proses data mining metode ini terdiri dari 6 fase tahapan (Iarose, 2005), Adapun tahapan-tahapannya seperti yang ditunjukkan pada gambar 1 berikut.



Gambar 1. Diagram Penelitian

Penelitian ini dibuat untuk memprediksi faktor-faktor yang mempengaruhi nilai dan waktu kelulusan mahasiswa dengan menerapkan algoritma SVM, metode boosting dan *k-folds cross validation* serta untuk uji validasi tanpa *cross validation* menggunakan 242 data. Hasil akhir dari penelitian ini adalah berupa model klasifikasi yang menerangkan faktor-faktor utama yang mempengaruhi data metode boosting (Adabost) dan *k-fold cross validation* yang menerangkan untuk melakukan evaluasi klasifikasi dengan teknik *cross validation*. Dalam pengujian ini, penulis menggunakan *5 fold*, *7 fold*, dan *10 fold*. Selanjutnya setelah data dikelompokkan menjadi beberapa kelompok sesuai nilai *fold*, maka Langkah selanjutnya menghitung akurasi dataset PDDIKTI di Admin Terpadu Universitas KH A WAHAB HASBULLAH Jombang.

4 HASIL DAN ANALISA

Penelitian ini menggunakan pendekatan kualitatif. Studi kasus penelitian ini bertempat di Universitas KH A Wahab Hasbullah Jombang. Februari sampai maret 2021. Target penelitian ini data mahasiswa sejumlah 807 dari tahun 2013-2019.

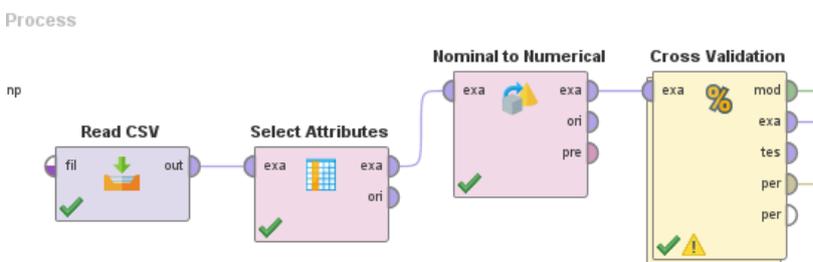
Hasil penelitian ini menggunakan tahap CRISP-DM. Untuk memprediksi nilai dan waktu kelulusan mahasiswa Universitas KH A Wahab Hasbullah Jombang. Dengan demikian mereka dapat di prediksi nilai dan waktu kelulusan mahasiswa dengan menghitung jumlah klasifikasi 2 class untuk membutuhkan training set positif dan negative dan menghasilkan hyperplane (pemisah) terbaik untuk memisahkan ke dalam 2 kelas dari 9 atribut jenis kelamin, ip semester 1, ip semester 2, ip semester 3, ip semester 4, sks semester 1, sks semester 2, sks semester 3, sks semester 4 dan 1 atribur hasil yaitu kelulusan. Dengan menggunakan algoritma SVM yang diimplementasikan kedalam rapid miner.

Pada penelitian ini, analisa menggunakan sebuah sistem yaitu data mining dengan menggunakan metode algoritma svm data sebanyak 807 data mahasiswa FTI Universitas KH A Wahab Hasbullah. Berikut ini adalah data sampel berupa tabel yang akan dilakukan proses melalui tool rapid miner sebagai berikut:

1	1	1355201...	Abdur R...	S1 Infor...	L	Peserta ...	2013	3,00	3,70	3,90
2	2	1355201...	Ahmad F...	S1 Infor...	L	Peserta ...	2013	1,56	2,50	2,50
3	3	1355201...	Ahmad S...	S1 Infor...	L	Peserta ...	2013	2,30	2,70	2,50
4	4	1355201...	Ahmad Z...	S1 Infor...	L	Peserta ...	2013	2,75	2,80	2,60
5	5	1355201...	Aris Suy...	S1 Infor...	L	Peserta ...	2013	2,56	2,30	2,40
6	6	1355201...	Ari Iswah...	S1 Infor...	L	Peserta ...	2013	3,20	3,10	3,50
7	7	1355201...	Ali Mahfud	S1 Infor...	L	Peserta ...	2013	3,50	3,50	3,20
8	8	1355201...	Dewi Nu...	S1 Infor...	P	Peserta ...	2013	3,10	3,70	3,60
9	9	1355201...	Dian Wa...	S1 Infor...	P	Peserta ...	2013	3,50	3,40	3,40
10	10	1355201...	Imam M...	S1 Infor...	L	Peserta ...	2013	3,45	3,60	3,50

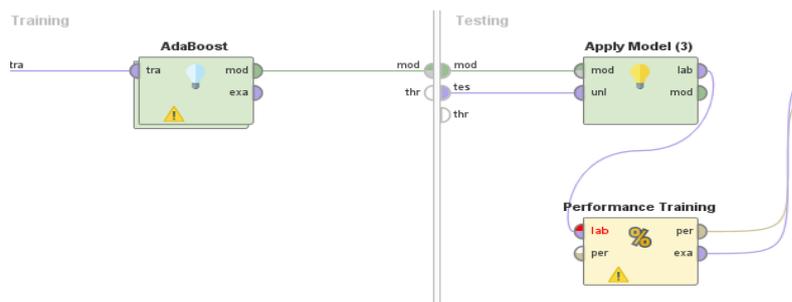
Gambar 2. data training

Pada proses selanjutnya proses validasi dilakukan penerapan algoritma data mining yang digunakan pada penelitian ini yaitu algoritma SVM dan metode boosting seperti gambar berikut:



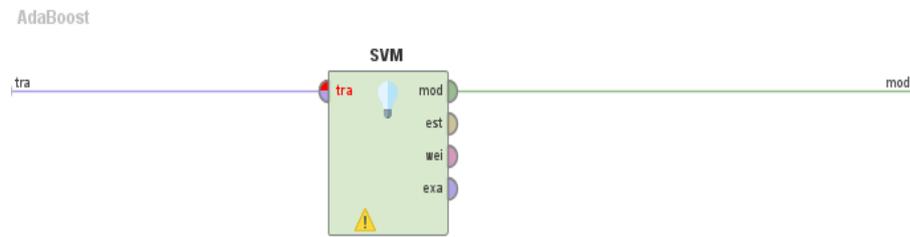
Gambar 4.3 Proses pengambilan data untuk algoritma svm dan metode boosting

Berdasarkan gambar 4.3 proses pengambilan data, data dimasukkan dalam bentuk format CSV kemudian menyeleksi atribut yang dipakai kemudian atribut yang sudah diseleksi masuk pada nominal to numerical dan data nya di ubah menjadi numerical fungsinya untuk diproses ke algoritma svm dan metode boosting (AdaBoost), kemudian masuk ke cross validation.



Gambar 3. Susunan Operator AdaBoost, dan Algoritma SVM

Selanjutnya pada gambar 4.4 didalam cross validation terdapat susunan operator adaboost yang didalamnya ada algoritma SVM yang ada pada gambar 4.5 dibawah ini. Setelah itu masuk pada apply model kemudian performance training.



Gambar 4. algoritma SVM

4.1. Analisis Tahap Pengujian

Implementasi algoritma SVM dan metode boosting dilakukan dengan bantuan perangkat lunak RapidMiner. Pembuatan model svm dan metode boosting untuk menghasilkan nilai akurasi menggunakan Uji *K-fold cross validation* dengan k = 5, 7, 10 dapat dilihat pada proses SVM untuk C= 0.0 yang digunakan bernilai yaitu 86.36%, artinya dari 807 data, 86.36% terklasifikasi dengan benar oleh model. Oleh karena itu model klasifikasi yang dari beberapa C yang ditentukan pada Tabel 1 berikut:

Tabel 1. Nilai akurasi dan model klasifikasi

Accuracy (%)		K-Fold Cross Validation dan Adaboost		
		K=5	K=7	K=10
SVM (kernel: dot)	C = 0.0	86.87% +/- 1.09%	86.62% +/- 1.29%	86.36% +/- 3.52%
	C = 0.5	83.28% +/- 4.86%	80.19% +/- 4.93%	80.54% +/- 4.95%
	C = 1.0	72.48% +/- 11.68%	72.99% +/- 11.98%	79.43% +/- 5.68%

Dari tabel 1. Diatas dapat diketahui dari hasil klasifikasi dan pengukuran pada data validasi dengan jumlah 807 data yang diambil dengan menggunakan 10 fold cross validation dengan metode boosting dan algoritma svm diperoleh hasil dari tabel 2 berikut:

Tabel 2. Confusion matrix dari Rapid Miner

Prediksi	Ya	Tidak	Class precision
Ya	359	49	87.99%
Tidak	61	338	84.71%
Class recall	85.48%	87.34%	

Keterangan Tabel 4.2 adalah:

- a. Jumlah data sebenarnya yang SESUAI dan diprediksi SESUAI adalah 359.
- b. Jumlah data sebenarnya yang TIDAK SESUAI dan diprediksi TIDAK SESUAI adalah 338.
- c. Jumlah data sebenarnya yang TIDAK SESUAI dan diprediksi SESUAI adalah 49.
- d. Jumlah data sebenarnya yang SESUAI dan diprediksi TIDAK SESUAI adalah 61.

Analisis hasil pengujian dilakukan dengan melakukan perhitungan menggunakan Rapid Miner berupa *confusion matrix*. Berikut ini merupakan hasil dari perhitungan *Confusion matrix* pada algoritma SVM dan metode boosting (Adaboost).

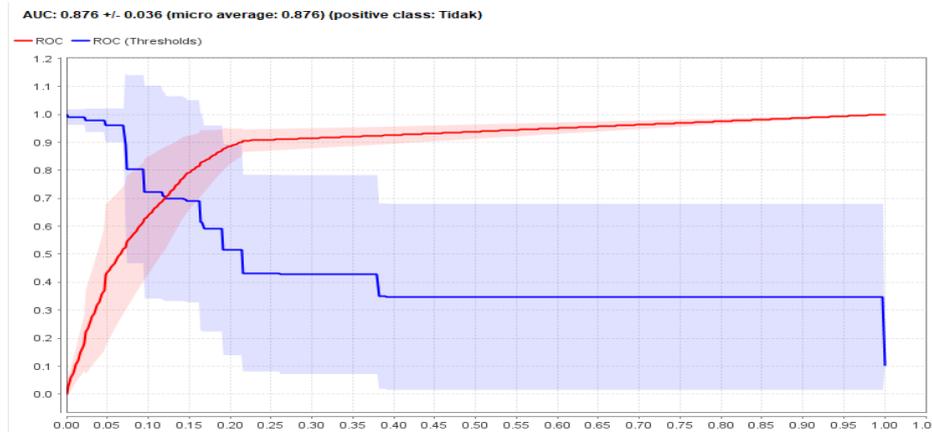
$$\text{Accuracy} = \left(\frac{359+338}{807} \right) * 100\% = 86,36\%$$

$$\text{Precision} = \left(\frac{359}{359+49} \right) * 100\% = 87,99\%$$

$$\text{Recall} = \left(\frac{359}{359+61} \right) * 100\% = 85,48\%$$

Dari perhitungan diatas, dapat disimpulkan hasil dari perhitungan accuracy, precision dan recall tersebut sama dengan hasil perhitungan yang ditampilkan pada tabel 2.

AUC dari model yaitu:



Gambar 5. AUC Support Vector Machine

Berdasarkan gambar diatas bahwa metode SVM termasuk klasifikasi sangat baik karena memiliki nilai AUC 0.876.

4.2. Validasi dan Pengujian

Dari hasil klasifikasi dan pengukuran pada data validasi dengan jumlah 807 data yang diambil dengan menggunakan 10 fold cross validation dengan metode boosting dan algoritma svm diperoleh hasil sebagai berikut:

Tabel 3. *Confusion matrix*

Aktual	Prediksi	
	Ya	Tidak
Ya	359	49
Tidak	61	338

Hasil *Confusion matrix* pada tabel 3. Algoritma svm dan metode boosting mampu mengidentifikasi 870 data yang sesuai dengan data uji. Dari hasil data uji, 359 data bernilai berhasil dan 61 data tidak berhasil, sehingga didapat algoritma svm dan metode boosting mengidentifikasi berhasil sebanyak 359 data dan tidak berhasil sebanyak 61 data, kesalahan mengidentifikasi sebanyak 387 data.

Analisis hasil pengujian dilakukan dengan melakukan perhitungan secara manual dengan *confusion matrix*. Berikut ini merupakan hasil dari perhitungan *confusion matrix* pada algoritma svm dan metode boosting.

$$\begin{aligned}
 \text{Akurasi} &= \frac{359+338}{359+49+61+338} = 86.36\% \\
 \text{Error} &= \frac{61+49}{359+49+61+338} = 13.64\% \\
 \text{AUC} &= 0.876
 \end{aligned}$$

Dari perhitungan diatas dapat disimpulkan bahwa hasil dari perhitungan accuracy, error dan AUC tersebut sama dengan hasil perhitungan manual. Berdasarkan pengujian dan Analisa hasil pengujian yang dilakukan dengan tingkat akurasi 86.36%, error 13,64%, AUC 0.876% menunjukkan nilai akurasi yang masih dalam kategori baik error dan auc bernilai baik karena tingkat kesalahan sedikit, bahwa peneliti berhasil dalam mengimplementasikan algoritma svm dengan metode boosting dengan baik untuk prediksi nilai dan waktu kelulusan mahasiswa.

5 KESIMPULAN

5.1. Kesimpulan

Penutup Dari penelitian prediksi nilai dan waktu kelulusan mahasiswa Universitas KH A Wahab Hasbullah Jombang dengan total 807 data, dimana pada penelitian ini menggunakan skenario uji k fold cross validation dengan nilai k = 5, 7 dan 10 diperoleh beberapa kesimpulan yaitu

Dari hasil implementasi data mining dengan menggunakan algoritma SVM dan metode boosting mampu menghasilkan rule guna memprediksi nilai dan waktu kelulusan mahasiswa berdasarkan riwayat akademik terdahulu. Pengujian decision system dengan menggunakan Aplikasi RapidMiner sangat dirasakan dapat mempermudah proses decision system dalam menghasilkan rule keputusan sebagai dasar melakukan prediksi. Dan berdasarkan hasil uji coba yang sudah dilakukan dari uji validasi K = 5, K= 7 jika dibandingkan dengan hasil yang di peroleh pada 10 fold Cross Validation, adaboost dan svm yaitu akurasi 86,36% error 13,64%, AUC 0.876% yang memiliki nilai akurasi tertinggi. Yang artinya algoritma SVM dan boosting (AdaBoost) mampu melakukan pengklasifikasian dengan baik.

5.2 Saran

Bersadarkan kesimpulan di atas peneliti menyarankan beberapa hal, yaitu

1. Hasil penelitian berdasarkan analisa dengan pemrograman rapid miner dapat digunakan sebagai *Decision Support System* bagi pihak Universitas KH A Wahab Hasbullah guna merancang strategi atau landasan dalam menentukan kelulusan waktu bagi mahasiswa khususnya memprediksi nilai dan waktu kelulusan mahasiswa.
2. Untuk pengembangan penelitian selanjutnya, peneliti menyarankan untuk dikembangkan menjadi *sistem pendukung keputusan (SPK)*.
3. Pada penelitian ini menggunakan 807 data mahasiswa. Pada penelitian selanjutnya untuk mengestimasi akurasi yang digunakan pada sebuah algoritma akan lebih baik jika data yang digunakan lebih banyak sehingga kemungkinan akurasi akan lebih akurat dalam sebuah perhitungan algoritma.

ACKNOWLEDGEMENTS

Terima kasih kepada Segenap Civitas Akademik Universitas KH A Wahab Hasbullah Jombang yang telah memberikan izin kepada peneliti untuk mengambil sample data mahasiswa sebanyak 807 data dari tahun 2013-2019, untuk bahan penelitian. Penelitian ini ditujukan sebagai salah satu syarat kelulusan program studi magister teknik informatika pada universitas AMIKOM Yogyakarta.

DAFTAR PUSTAKA

- Bisri Ahmad, Wahono Satrio romi, 2015 Penerapan Adaboost untuk Penyelesaian Ketidakseimbangan Kelas pada Penentuan Kelulusan Mahasiswa dengan Metode Decision Tree, Fakultas Teknik, Universitas Pamulang.
- Larose, D. T. (2005) *Discovering Knowledge in Data: An Introduction to Data Mining*, 81 81 *Discovering Knowledge in Data: An Introduction to Data Mining*. doi: 10.1002/0471687545.
- Luluk Susiana, Dkk., 2019 Penerapan Metode Boosting Pada Cart Untuk Mengklasifikasikan Korban Kecelakaan Lalu Lintas Di kota Palu, Universitas Tadulako, Sulawesi Tengah.
- Miftrahurrohman, B, dkk., 2019 Analisis Prediksi Mahasiswa Mengundurkan Diri Dari Universitas Xyz Dengan Metode Support Vector Machine, Program Studi Sistem Informasi, Fakultas Teknologi Informasi dan Kreatif, Universitas Internasional Semen Indonesia, Kompleks PT. Semen Indonesia (Persero) Tbk. Jl. Veteran, Gresik Jawa Timur, 61122.
- Nofitri Rika, Irawati Novica, 2019 Integrasi Metode Neive Bayes Dan Software Rapidminer Dalam Analisis Hasil Usaha Perusahaan Dagang, Sekolah Tinggi Manajemen Infromatika dan Komputer Royal, Sumatera Utara.

- Pang, Yulie, Judd, Nicolas, Dkk., 2017 Predicting Students' Graduation Outcomes through Support Vector, Department of Mathematics, Southern Connecticut State University, USA.
- Pratama, Arif, Wihandika, C, R, Ratnawati, E, D., 2018 Implementasi Algoritme Support Vector Machine (SVM) untuk Prediksi Ketepatan Waktu Kelulusan Mahasiswa, Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya, Malang.
- Srisulistiowati Dwi Budi, Dkk., 2021 Sistem Informasi Prediksi Penjualan Alat Tulis Kantor Dengan Metode FP-Growth (Studi Kasus Toko Koperasi Sekolah Bina Mulia), Universitas Bhayangkara Jakarta, Jakarta.
- Suhardjono, Wijaya, Ganda, Hamid, Abdul., 2019 Prediksi Waktu Kelulusan Mahasiswa Menggunakan SVM Berbasis PSO, Universitas Bina Sarana Informatika, Yogyakarta.
- Syamsiah, Dkk., 2020 Analisa Particle Swarm Optimization Terhadap Kepuasan Taman Dadap Merah Dengan Model SVM, Universitas Indraprasta PGRI, Jagakarsa Jakarta Selatan.