

IMPLEMENTASI METODE FORWARD SELECTION PADA ALGORITMA SUPPORT VECTOR MACHINE (SVM) DAN NAIVE BAYES CLASSIFIER KERNEL DENSITY (STUDI KASUS KLASIFIKASI JALUR MINAT SMA)

Theopilus Bayu Sasongko¹, Oki Arifin²

^{1,2}Universitas Amikom Yogyakarta
Email: ¹theopilus.27@amikom.ac.id, ²okiarifin@amikom.ac.id

(Naskah masuk: 16 Agustus 2018, diterima untuk diterbitkan: 27 Mei 2019)

Abstrak

Peminatan merupakan kegiatan yang disediakan oleh pihak sekolah yang berguna untuk mengakomodasi pilihan minat, bakat, atau kemampuan peserta didik dengan orientasi pemusatan. Penentuan jalur minat umumnya melibatkan banyak *attribute*. Klasifikasi maupun prediksi pada *data mining* menggunakan fitur seleksi sangat penting untuk pemilihan *attribute* yang tepat, karena berpengaruh pada performansi model, oleh sebab itu perlu metode untuk melakukan seleksi atribut. Penelitian ini membandingkan implementasi metode *forward selection* pada Algoritma SVM dan Naive Bayes Kernel Density. Studi kasus yang digunakan adalah jalur minat pada siswa SMA pada dua sekolah yang berbeda. Proses pembentukan model klasifikasi dengan menganalisa perubahan *kernel*, faktor pinalti (C) SVM, *number of kernel Naive bayes kernel density*, dan hasil *feature subset forward selection*. Digunakan lima buah eksperimen kernel SVM yaitu *dot (linear)*, *radial (RBF)*, *polynomial*, *neural*, dan *anova*. Proses uji coba perubahan parameter menggunakan rentang 0.0-100.0. Hasil dari penelitian ini diantaranya adalah *feature subset dataset* SMA ABC yang terpilih yaitu nilai IPA, tes akademik, abstrak konseptual, analisa sintesa, dan logika numerik, sedangkan *feature subset* SMA XYZ yaitu nilai IPA, logika numerik, dan analisa sintesa. Hasil pengujian *dataset* SMA ABC menggunakan algoritma FS-SVM berbasis *kernel anova* parameter C=10.0 sebesar 99.29%. Sedangkan hasil pengujian *dataset* SMA XYZ menggunakan algoritma FS-SVM berbasis *kernel anova* parameter C=10.0 sebesar 95.17%.

Kata kunci: *Forward Selection, SVM, Naive Bayes Kernel Density (NBKD), peminatan siswa, feature selection*

IMPLEMENTATION FORWARD SELECTION METHODS FOR SUPPORT VECTOR MACHINE (SVM) AND NAIVE BAYES KERNEL DENSITY ALGORITHM (CASE STUDY STUDENT MAJOR CLASSIFICATION)

Abstract

Specialization is an activity provided by the school that is useful to accommodate the choice of interests, talents, or abilities of students with a concentration of orientation. The determination of interest generally involved many attributes. The classification and prediction on the data mining that use the selection feature is very important for the selection of the right attribute, because it affects the performance of the model, therefore a method is needed to select attributes. This study compares the implementation of the forward selection method in the SVM Algorithm and Naive Bayes Kernel Density. The case study that is used is the interest of students in high school and compared with two different schools. The process of modelling by studying kernel changes, penalty factors (C) SVM, number of kernel Naive bayes kernels, and the results of features from subset forward selection. Five SVM kernel experiments are used, namely dot (linear), radial (RBF), polynomial, neural, and anova. The trial process of changes parameters uses the range 0.0-100.0. The results of this study include features of selected ABC SMA subset data, which are IPA values, academic tests, conceptual abstracts, synthesis analysis, and numerical logic, while the XYZ SMA subset features are IPA values, numerical logic, and synthesis analysis. The test results of the ABC High School dataset that use the kernel-based FS-SVM algorithm parameter C = 10.0 is 99.29%. While the results of testing the XYZ SMA dataset that use the kernel-based FS-SVM algorithm parameter C = 10.0 for 95.17%.

Keywords: *Forward Selection, Support Vector Machine (SVM), Naive Bayes Kernel Density (NBKD), student majority, feature selection*

1. PENDAHULUAN

Di Indonesia, siswa-siswi sekolah menengah atas akan dinilai dan dikelompokkan ke dalam peminatan (penjurusan). Tujuan peminatan yang dilakukan oleh sekolah adalah mengakomodir pilihan minat, kemampuan, atau bakat peserta didik, perluasan Prosedur penentuan peminatan di Indonesia berdasarkan pada Kurikulum 2013 dimana penentuan peminatan SMA didasarkan pada nilai mata pelajaran, nilai raport, psikotest peminatan, dan masih banyak unsur lain.

Pada dasarnya sudah terdapat beberapa alat bantu yang dapat digunakan untuk melakukan prediksi peminatan secara otomatis diantaranya adalah ilmu statistika dan *data mining*.

Feature selection adalah salah satu fokus yang ada di *data mining*. *Feature selection* merupakan suatu proses pemilihan bagian dari *variable* dari semua *variable* yang ada di *dataset*. Salah satu metode *feature selection* yaitu *forward selection*.

Penelitian yang dilakukan oleh Bayu pada tahun 2016 dilakukan eksperimen dengan melakukan proses pembobotan dan seleksi *attribute* dengan menggunakan *particle swarm optimization (PSO)* pada algoritma SVM dengan studi kasus jalur minat siswa SMA. Pada penelitian ini disimpulkan bahwa algoritma PSO-SVM lebih baik dibandingkan SVM (Bayu, 2016).

Penelitian lain yang dilakukan oleh Arya Purnanditya mengenai metode *forward selection* pada algoritma *naïve bayes classifier* dengan studi kasus prediksi klasifikasi kelulusan. Penelitian ini menghasilkan tingkat akurasi sebesar 99.17% lebih baik dari pada jika tidak menggunakan metode *forward selection* yang hanya menghasilkan nilai 95.83% (Arya, 2015).

Penelitian yang dilakukan oleh Chung-Jui mengimplementasikan *feature selection* pada *PSO-Multi Class SVM*. *Dataset* yang dipakai diantaranya adalah *vowel*, *wine*, *wdbc*, *ionosphere*, dan *sonar dataset*. Kesimpulan hasil penelitian ini bahwa metode *PSO* dapat dengan baik digabungkan dengan *multiclass SVM*. Dijabarkan kekurangannya *PSO-Multi Class SVM* memiliki waktu komputasi (Chung, 2007).

Feature selection PSO pada *multi class SVM* juga diteliti oleh Fatima dan Kaddour, Algoritma PSO-SVM digunakan untuk melakukan klasifikasi *multiclass* (Ardjani, 2010).

Penelitian yang dilakukan oleh Nanja M yang berjudul Metode *K-Nearest Neighbor* Berbasis *Forward Selection* Untuk Prediksi Harga Komoditi Lada menunjukkan *forward selection* lebih baik dibandingkan *backward selection*, algoritma *FS-KNN* lebih baik dibandingkan *KNN* (Nanja, 2015)

Andi P juga melakukan penelitian mengenai Algoritma *Naïve bayes* Berbasis *Forward Selection* Pada Prediksi Kelulusan Tepat Waktu. Pada penelitiannya dapat disimpulkan bahwa penggunaan

fitur *Forward Selection* terbukti meningkatkan performansi klasifikasi sebesar 97.92% (Prasetyo, 2017).

Latar belakang masalah pada penelitian ini adalah menemukan *subset attribute* terbaik dengan melakukan *feature selection* pada *dataset* peminatan sehingga didapatkan performansi model klasifikasi yang optimal. Peneliti juga melakukan komparasi dan analisis penerapan metode *forward selection* pada algoritma SVM dan *Naïve bayes classifier kernel density*. Hasil dari penelitian ini adalah perbandingan performansi akurasi antara *FS-SVM*, *SVM*, *FS-NBK* dan *NBK (Naïve bayes kernel density)*. Selain itu fitur *subset* hasil seleksi fitur dari *dataset* peminatan.

2. METODOLOGI PENELITIAN

A. Alur Penelitian

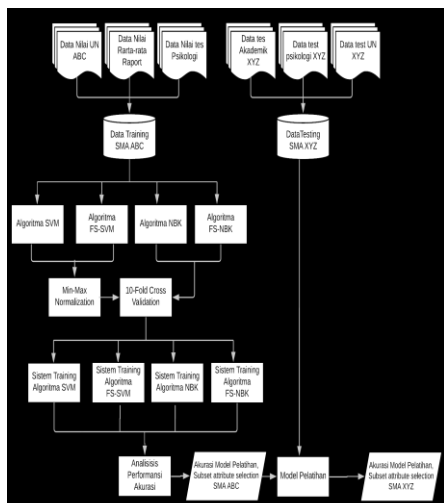
Penelitian ini berfokus membangun model klasifikasi jalur peminatan dengan membandingkan performansi algoritma yang menggunakan *forward selection* dan tanpa *forward selection*. Tujuan dari penelitian ini adalah mengembangkan model pelatihan algoritma yang digunakan untuk melakukan klasifikasi jalur minat siswa SMA, mengetahui perbandingan performansi nilai akurasi antara algoritma SVM dengan *Naïve bayes kernel density*, menganalisa pengaruh nilai parameter *kernel* terhadap performansi model, dan *feature subset* dari hasil seleksi fitur *forward selection*. Alur penelitian diawali dengan pengumpulan data peminatan di dua sekolah. Data peminatan sekolah yang pertama dinamakan *dataset ABC*. *Dataset* ini memiliki 11 atribut yang meliputi nama, nilai matematika, ipa, mipa, ips tes psikologi iq, logika rasional, konkrit operasional, abstrak konseptual, analisa sintesa, logika verbal, logika numerik, dan daya ingat dengan jumlah *records* adalah 280 siswa. Siswa IPA berjumlah 150 orang dan siswa IPS berjumlah 138 siswa.

Proses *training* algoritma SVM didahului dengan normalisasi. Normalisasi menggunakan *min-max normalization*. *Min-max normalization* dilakukan dengan mengubah dari data asli kedalam data *linear* dengan rentang +1 dan -1. (Agustinus, 2012). Sedangkan proses *training* pada algoritma *Naïve bayes classifier kernel density* tidak menggunakan normalisasi *dataset* karena *Naïve bayes kernel density* dapat menangani nilai/data diskrit dengan baik. Kedua proses perlakuan dimasukkan ke dalam sistem *training SVM* dan sistem *training Naïve bayes kernel density*. Digunakan beberapa fungsi kernel SVM diantaranya *radial (RBF)*, *dot (linear)*, *polynomial*, dan *anova* dengan parameter C (faktor pinalti) antara nilai 0.0-100. Proses *training* pada algoritma *Naïve bayes classifier kernel density* digunakan *estimation mode greedy* dengan *number of kernel* 0.0-100.0 dengan rentang 10.0. Digunakan 10 *k-fold cross validation* guna pembagian *data training* dan *testing*. Kriteria pengujian dan pemilihan model terbaik

menggunakan metode penghitungan akurasi dari setiap eksperimen yang ada.

Model dengan nilai akurasi paling tinggi selanjutnya digunakan untuk melakukan klasifikasi pada *dataset* peminatan sekolah kedua. *Dataset* peminatan sekolah kedua ini dinamakan *dataset XYZ*. *Dataset XYZ* memiliki jumlah records sebanyak 288 records dengan 11 variabel diantaranya yaitu nilai matematika, ipa, tes akademik, iq, logika rasional, konkrit operasional, abstrak konseptual, analisa sintesa, logika verbal, logika numerik, dan daya ingat. Pengujian model klasifikasi pada *dataset XYZ* menggunakan metode akurasi dan nilai AUC.

Berikut gambaran proses penelitian dari pengumpulan dataset hingga pembentukan model klasifikasi peminatan pada Gambar 1.



Gambar 1. Skema Metodologi Penelitian

B. Algoritma SVM

Support Vector Machine merupakan algoritma untuk klasifikasi maupun prediksi (Han, 2013). Penerapan algoritma SVM dengan mencari jarak atau ruang pemisah dari suatu *dataset*, pada kenyataannya dikehidupan sehari-hari banyak yang tidak dapat dipisahkan dengan cara *linear*. *Dataset non-linear* adalah suatu kondisi *dataset* yang tidak dapat dipisahkan oleh garis atau bidang yang memisahkan antar kelas data yang satu dengan kelas data yang lain. Permasalahan ini dapat diselesaikan dengan dua cara, yaitu:

1. *Soft margin hyperplane*.

Soft margin hyperplane yaitu suatu metode yang berupaya mentransformasi data *non-linear* ke bentuk *linear* dengan garis atau bidang tetap bersifat fleksibel.

Formulasi pada *soft margin hyperplane* yang menggunakan variabel slack (ξ) dirumuskan dengan Persamaan (1)(Han, 2013).

$$x_i \cdot w + b \geq 1 - \xi \text{ untuk } y_i = \text{ kelas 1} \quad (1)$$

$$x_i \cdot w + b \leq -1 + \xi \text{ untuk } y_i = \text{ kelas 2}$$

Pada rumus persamaan (1), proses menemukan bidang pemisah dengan menggunakan persamaan (2)(Han, 2013).

$$\min \frac{1}{2} \|w\|^2 + C (\sum_{i=1}^n \xi_i) \quad (2)$$

Dengan fungsi pemisah $y_i(x_i \cdot w + b) \geq 1 - \xi \forall i = 1, \dots, n$

Nilai C adalah nilai variabel pinalti yang menentukan besarnya toleransi kesalahan klasifikasi terhadap suatu data.

2. *Feature Space*

Feature space adalah suatu metode pengubahan dari *input space (dot product)* yang hanya dapat memisahkan data linear kedalam bentuk dimensi tinggi (*feature space*). *Feature space* adalah *kernel trick* $\Phi(\bar{x}_i) \cdot \Phi(\bar{x}_j)$ yang selanjutnya berubah menjadi fungsi Kernel $K(x_i, x_j)$ (Han, 2013). Proses pengubahan *dot product* ke *feature space* mengakibatkan waktu komputasi yang lama, karena itu dibuat Proses Kernel.

Digunakan lima buah fungsi kernel diantaranya *radial (RBF)*, *polynomial*, *dot (linear)*, *anova* dan *neural*. Berikut adalah formulasi dari kelima kernel tersebut :

Tabel 1. Pengujian Kernel SVM

Nama Kernel	Kernel $K(x, y), t=1, 2, \dots, N$
Dot (Linear)	$K(x, y) = x^T y + c$
Radial (RBF)	$K(x, y) = \exp(-g \ x - y\ ^2)$
Polynomial	$K(x, y) = (\alpha x^T y + c)^d$
Neural (NN)	$K(x, y) = \tanh(\alpha x^T y + c)$
Anova	$K(x, y) = \sum_{k=1}^n \exp(-g(x - y)^2)^d$

C. *Naive Bayes Classifier Kernel Density*

Naive Bayes Classifier (NBC) adalah model algoritma probabilitas sederhana yang didasarkan pada teorema *bayes*. Hukum *bayes* yang digabungkan dengan '*naive*' mengakibatkan setiap variabel memiliki sifat yang independen. *Naive Bayes Classifier* dapat dilatih dengan efisien dalam *supervised learning*, proses *Naive Bayes Classifier* menganalogikan ada atau tidak suatu variabel pada suatu kelas tidak memiliki korelasi dengan ada atau tidak variable lain dikelas yang sama (Muhamad, 2017). *Naive Bayes Classifier* termasuk algoritma yang tergolong *supervised learning*, sehingga pada prosesnya dibutuhkan data training untuk melakukan klasifikasi. Alur proses klasifikasi dengan menghitung kemungkinan (*probability*) data yang masuk pada masing-masing *class label*. *Class label* dengan tingkat probability terbesar yang akan dijadikan *class label* data masukan. *Naive Bayes Classifier* mengadopsi rumus dari teorema bayes, sehingga kompleksitas komputasi menjadi sederhana. Kelebihan lain dari *Naive bayes classifier* dapat menangani *dataset* yang beratribut banyak (Muhamad, 2017). Persamaan (3) menunjukkan rumus dari *Naive bayes classifier*

$$P(C | X) = \frac{P(X|C)P(C)}{P(X)} \quad (3)$$

Keterangan:

X : Prediksi data masukan

C : Hipotesa data masuk pada suatu kelas label

$P(C|X)$: Probabilitas hipotesa berdasarkan kondisi
 $P(C)$: Probabilitas hipotesa
 $P(X|C)$: Probabilitas atas dasar kondisi pada hipotesa
 $P(X)$: Probabilitas dari C

$P(X|C_i)P(C_i)$ dengan memilih yang terbesar sebagai kelas yang dipilih sebagai hasil prediksi.

Kelemahan dari Algoritma NBC adalah ketika menjalankan proses *training* dengan *dataset* yang memiliki bobot *attribute* yang tidak seragam (*non-normal distribution / non gaussian distribution*) akan berakibat performansi klasifikasi yang tidak sesuai. Oleh sebab itu dibutuhkan cara untuk menangani masalah ini yaitu dengan penerapan fungsi *kernel density*. Algoritma ini yang dinamakan dengan *Naïve bayes classifier kernel density estimation*. Persamaan (4) menunjukkan rumus *kernel density estimation*.

$$f(x; H) = n^{-1} \sum_{i=0}^n K_H(x - x^{(i)})$$

Pendekatan *Kernel* digunakan untuk memindahkan fungsi *dot ke hyper space*. Fungsi *kernel* yang ada dirumuskan dengan rumus pada persamaan (5) (Perez, 2009):

$$K_H(x) = |H|^{-1/2} K(H^{-1/2}x)$$

Fungsi *kernel* diatas dimana H adalah $d \times d$ bandwidth atau bentuk *matrix (BM)* $x = (x_1 \dots x_d)$ adalah sebuah bentuk dimensi dari X . N adalah estimator dari *model learning*. Fungsi $f(\cdot; H)$ digunakan untuk menentukan rata-rata kepadatan n *kernel* pada setiap observasi. Sehingga Fungsi *kernel Density estimation naïve Bayes* seperti persamaan (6).

$$P_i(x_i|C=c) = \frac{1}{N_{ch}} \sum_{j=1}^{N_c} K(x_i, X_j|i|c) \quad (6)$$

Dimana k adalah fungsi *kernel Gaussian* dengan *mean 0* dan *variance 1*, N_c adalah jumlah input data X yang diprediksi masuk kelas c , $X_j|i|c$ adalah nilai fitur yang terletak pada posisi i dari j dengan *input* $X=(x_1, x_2, x_3, \dots, x_n)$ pada kelas c dan h adalah *bandwidth* atau *smoothing parameter*. Untuk mengoptimalkan nilai klasifikasi, parameter h harus dicari nilai optimalnya. Tujuan dari *Kernel Density Estimation (KDE)* adalah dapat memanipulasi atribut yang bersifat kuantitas, sehingga bobot tiap *attribute* dapat sesuai.

D. Feature Selection

Feature Selection atau yang sering disebut juga sebagai *attribute selection* merupakan proses menemukan *subset* hasil seleksi fitur dari suatu *dataset*. *Feature selection* dipakai pada bidang statistika, dan *data mining* (Han, 2013). *Feature selection* adalah metode yang penting untuk menghasilkan klasifikasi yang baik. Tanpa *feature selection* proses komputasi dan performansi model menjadi buruk. Tujuan dari *feature selection* adalah membuang atribut yang tidak

relevan. Pada umumnya *feature selection* dibedakan menjadi 3, yaitu :

1. Filter Based Feature Selection

Filter based feature selection menggunakan metode statistika untuk menilai setiap atribut. Konsep kerjanya adalah atribut akan dirangking dan dibuat matriks terlebih dahulu untuk mengidentifikasi *attribute* yang relevan atau tidak relevan yang akan dibuang dari *dataset*. Contoh dari *filter selection* yaitu *Chi squared test*, *information gain*, dan *pearson correlation*.

2. Embedded Feature Selection

Embedded feature selection menjalankan seleksi fitur selama proses *training* dari parameter-parameter hingga optimal. *Embedded feature selection* mengadopsi prinsip *regularization methods*. *Regularization methods* memiliki faktor pinalti yang membatasi. Contoh *embedded feature selection* adalah *CART*, *C4.5*, *multinomial logistic regression*.

3. Wrapper Feature Selection

Metode *Wrapper feature selection* melakukan pencarian dari sekumpulan kombinasi *attribute*. Setiap kombinasi *attribute* dibandingkan satu dengan yang lainnya. Digunakan *model* algoritma untuk mengevaluasi kumpulan kombinasi yang ada. Contoh *Wrapper feature selection* adalah *random-hill climbing*, *forward* dan *backward selection yang berbasis heuristic*. Pada penelitian ini digunakan *forward selection* karena memiliki hasil *precision* yang lebih baik dibandingkan dengan tipe *filter* (Vinita, 2015)

Metode *Forward Selection* mengadopsi prinsip *regresi Linear*. *Forward Selection* adalah salah satu *model wrapper* yang digunakan mereduksi atribut *dataset* (Han, 2013).

Proses pencarian *attribute* dengan *forward selection* diawali dengan *empty model*, selanjutnya tiap variabel dimasukan hingga kriteria kombinasi model *attribute* terpenuhi dengan baik. Berikut adalah *pseudo code* dari *forward selection*:

1. Membuat *empty set*: $Y_k = \{\emptyset\}, k = 0$
2. Memilih *feature* terbaik:

$$X^+ = \arg \max_{x^+ \in Y_k} [J(Y_k + X^+)]$$
3. Jika $J((Y_k + x^+) > J((Y_k)$
 - a. Update $Y_{k+1} = Y_k + x^+$
 - b. $k = k + 1$
 - c. Kembali ke step -2

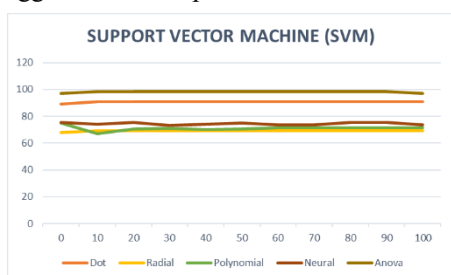
3. ANALISA HASIL DAN PEMBAHASAN

Pada penelitian ini dilakukan pengujian atau eksperimen dengan mengubah parameter yang ada pada SVM, FS-SVM, NBK, dan FS-NBK untuk menentukan model terbaik yang digunakan sebagai bahan klasifikasi data peminatan SMA ABC dan data peminatan SMA XYZ. Tabel 2 menunjukkan jumlah eksperimen beserta nilai parameter yang digunakan untuk klasifikasi model peminatan.

Tabel 2. Eksperimen Model Pelatihan Peminatan

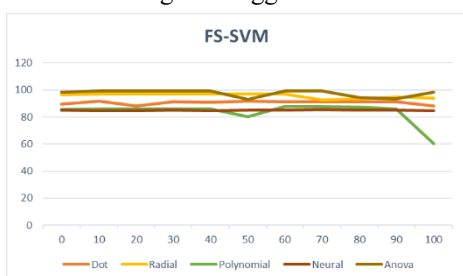
SVM dan F-SVM		NBK dan F-NBK		Estimation Mode
Jumlah Eksperimen	Parameter C	Jumlah Eksperimen	Number of Kernel	
1	0.0	1	1.0	Greedy
2	10.0	2	10.0	
3	20.0	3	20.0	
4	30.0	4	30.0	
5	40.0	5	40.0	
6	50.0	6	50.0	
7	60.0	7	60.0	
8	70.0	8	70.0	
9	80.0	9	80.0	
10	90.0	10	90.0	
11	100.0	11	100.0	

Eksperimen pertama digunakan *dataset* peminatan ABC. Hasil eksperimen memperlihatkan perbandingan nilai akurasi pengklasifikasian data menggunakan SVM pada Gambar 2.



Gambar 2. Grafik Akurasi Algoritma SVM

Pada Gambar 2 dapat dilihat bahwa nilai akurasi tertinggi diperoleh oleh *kernel anova* dengan nilai akurasi 98.21%. Grafik nilai akurasi terendah pada *kernel radial*. Pada Gambar 3 menunjukkan grafik nilai akurasi dengan menggunakan FS-SVM

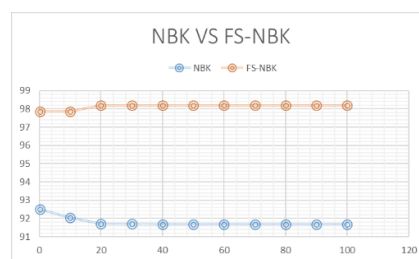


Gambar 3. Grafik Akurasi Algoritma FS-SVM

Pada Gambar 3 dilihat bahwa grafik nilai akurasi terbesar adalah pada *kernel anova* dengan nilai akurasi sebesar 99.29%.

Dilihat dari Gambar 2 dan 3 dapat disimpulkan sementara bahwa dengan menggunakan metode *Forward selection* dapat meningkatkan nilai performansi model klasifikasi algoritma SVM pada semua kernel yang diuji.

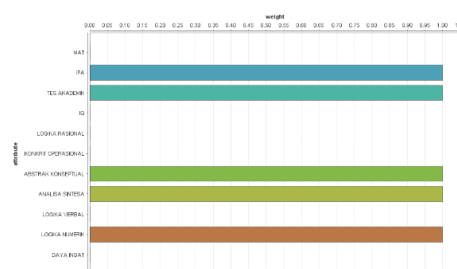
Pengujian *fitur selection* dengan *forward selection* pada *Naive bayes kernel density* dapat dilihat pada Gambar 4.



Gambar 4. Grafik Akurasi Algoritma NBK vs FS-NBK

Pada Gambar 4 dapat dilihat bahwa metode *forward selection* meningkatkan performansi yang cukup signifikan pada *Naive bayes kernel density* dengan nilai akurasi 98.21%. Nilai akurasi pada *NBK* menurun ketika *number of kernel* yang ada naik di angka 10.0, sebaliknya nilai akurasi *FS-NBK* meningkat ketika *number of kernel* yang ada naik di angka 10.0.

Kesimpulan Gambar 2,3, dan 4 hasil akurasi tertinggi diperoleh pada algoritma *FS-SVM* dengan kernel *anova* parameter C sebesar 10.0 dengan tingkat akurasi 99.29%. Berikut adalah hasil subset dari *feature selection* dataset SMA ABC menggunakan *forward selection* pada Gambar 5.



Gambar 5. Grafik Subset FS-SVM dengan C=10.0 SMA ABC

Pada Gambar 5 terlihat bahwa hasil *subset attribute selection* dengan *FS-SVM* parameter $C=10.0$ menunjukkan bahwa *attribute* yang paling berpengaruh pada performansi yaitu nilai IPA, tes akademik, abstrak konseptual, analisa sintesa, dan logika numerik.

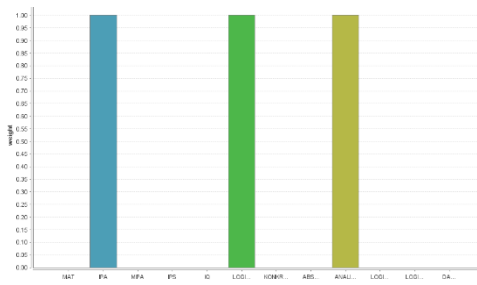
Proses selanjutnya adalah menguji model terbaik kedalam *dataset* peminatan XYZ dengan jumlah *records* 288 dan 14 *fitur*. Model klasifikasi yang diujikan adalah model *FS-SVM* kernel *Anova* dengan parameter C sebesar 10.0. Nilai pengukuran yang digunakan adalah Akurasi dan Nilai AUC.

Tabel 3 menunjukkan nilai performansi model FS-SVM kernel *anova* dengan parameter faktor pinalti (C) adalah 10.0 pada *dataset* peminatan XYZ. Sedangkan Gambar 6 adalah hasil *subset feature selection* menggunakan *feature selection* pada SMA XYZ.

Dari hasil uji coba pada dataset SMA XYZ pada Tabel 3 dan Gambar 6 menunjukkan nilai akurasi dan *subset* yang cukup berbeda dengan hasil uji coba pada SMA ABC, menurut peneliti hal ini dikarenakan *attribute* yang ada di *dataset* cukup berbeda.

Tabel 3. Performansi Model FS-SVM pada *dataset* XYZ

Nilai Performansi Model FS-SVM	
Akurasi	Nilai AUC
95.17	0.956

Gambar 7. Grafik Subset FS-SVM dengan $C=10.0$ SMA XYZ

4. KESIMPULAN DAN SARAN

Berdasar hasil penelitian dan pembahasan dapat ditarik kesimpulan beberapa, yaitu:

1. Proses pembentukan model klasifikasi peminatan dengan cara membandingkan efektifitas *forward selection* pada SVM dan *Naïve bayes kernel density*.
2. Metode *forward selection* terbukti dapat meningkatkan nilai performansi model baik itu pada algoritma SVM maupun *Naïve bayes kernel density*.
3. Nilai akurasi tertinggi *dataset* sekolah ABC diperoleh algoritma FS-SVM parameter C sebesar 10.0 dengan tingkat akurasi yaitu 99.29%
4. Pengujian model klasifikasi FS-SVM parameter C 10.0 pada *dataset* peminatan sekolah XYZ memperoleh akurasi sebesar 95.17% dan nilai AUC 0.956
5. Hasil *subset attribute selection* dengan FS-SVM parameter $C=10.0$ menunjukkan bahwa *attribute* yang paling berpengaruh pada performansi yaitu nilai IPA, tes akademik, abstrak konseptual, analisa sintesa, dan logika numerik.

Beberapa saran kedepan untuk diperbaiki adalah:

1. Diharapkan dapat menggunakan *automatic kernel selection* pada SVM.
2. Diharapkan dapat menggunakan *automatic parameter selection* pada parameter faktor pinalti (C) di algoritma SVM dan *minimum-bandwidth* pada algoritma *Naïve bayes kernel density*.

DAFTAR PUSTAKA

BAYU, T., 2016. Komparasi dan Analisis Kinerja Model Algoritma SVM dan PSO-SVM (Studi Kasus Jalur Minat SMA), Vol 2 No 2.

Jurnal Teknik Informatika dan Sistem Informasi, [online] Tersedia di: < <http://journal.maranatha.edu/index.php/jutis/article/view/627> > [Diakses 14 Agustus 2018].

ARYA, B., 2015. Penerapan Fitur Seleksi Forward Selection Menggunakan Algoritma Naïve Bayes Untuk Menentukan Atribut Yang Berpengaruh pada Klasifikasi Kelulusan Mahasiswa AKI Semarang. Universitas Dian Nuswantoro Semarang.

CHUNG, J.T., LI, Y.C., JUN, C.Y, AND CHENG, Y., 2007. Feature Selection using PSO-SVM, IAENG International Journal of Computer Science.

ARDJANI, F., SADOUNI, K., AND BENYETTOU, M., 2010. *Optimization of SVM MultiClass by Particle Swarm (PSO-SVM)*.

NANJA, M., DAN PURWANTO, 2015. Metode K-Nearest Neighbor Berbasis Forward Selection Untuk Prediksi Harga Komoditi Lada, Jurnal Pseudocode, Vol.2, No.1

PRASETYO, A., 2017. Algoritma Naïve Bayes Berbasis Forward Selection Pada Prediksi Kelulusan Tepat Waktu, Tesis, Universitas Dian Nuswantoro.

AGUSTINUS, J., 2012. Sistem Deteksi Intrusi Jaringan dengan Metode Support Vector Machine, Thesis. Jurusan Ilmu Komputer FMIPA UGM.

HAN, J., KAMBER, M., 2013. Data Mining: Concepts and Techniques. Third Edition. Morgan Kaufmann Publishers, San Fransisco.

MUHAMAD, H., ADI PRASOJO, C., AFIFAH, N., SURTININGSIH, L., DAN CHOLISSODIN, I., 2017. Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris. Jurnal Teknologi Informasi dan Ilmu Komputer, [online] Tersedia di: < <http://jtiik.ub.ac.id/index.php/jtiik/article/download/251/pdf> > [Diakses 14 Agustus 2018].

PEREZ, A., LARRANAGA, P., INZA, I., 2009. Bayesian Classifiers based on kernel density estimation Flexible classifiers. International Journal of Approximate Reasoning, [online] Tersedia di : < <https://www.sciencedirect.com/science/article/pii/S0888613X08001400> > [Diakses 10 Agustus 2018].

Vinita, C., Komparasi Algoritma Klasifikasi Machine Learning dan Feature Selection pada Analisis Sentimen Review Film, 2015. Journal of Intelligent Systems, Vol.1, No.1,